

DUAL-SOURCE Z-SCORE INTEGRATION FOR COMPETENCY ASSESSMENT OF IPC TEACHERS IN HIGHER VOCATIONAL EDUCATION: EVIDENCE FROM COLLEGE H

Liu Huan¹, Somkiet Tunkeaw², Somyos Chanboon³

Corresponding Author E-mail: 1669167222@qq.com

Received: Nov 19, 2025; Revised: Feb 24, 2026; Accepted: Mar 10, 2026

Abstract

Many higher vocational colleges lack a practical and comparable way to diagnose ideological and political theory (IPC) teachers' competency needs using both teacher and student perspectives. This study employed a dual-source survey design in one Chinese higher vocational college, based on Zhou and Xiang's (2019) four-dimension, 22-item competency framework. Data were collected from 21 IPC teachers (self-assessment) and 414 students (evaluation). The instruments showed acceptable reliability and construct validity. To enable cross-source comparison, teacher and student ratings were standardized to a common scale and integrated into a composite demand index. Findings indicate broad agreement between teachers and students, but clear differences in priority areas: teachers emphasized teaching ability, research competence, and knowledge literacy, while students placed greater weight on pedagogical foundations and teacher charisma. Overall, the results indicated an imbalance between explicit competencies (knowledge and skills) and implicit competencies (attitudes and personal attributes). The competency levels derived from the composite index can support tiered professional development, targeted mentoring and appraisal plans, and resource allocation aligned with demand levels. This dual-source, standardized approach offers a practical tool for competency-based teacher development in vocational education.

Keywords : Ideological and Political Theory (IPC) Teachers, Teacher Competency, Higher Vocational Education, Dual-Source Evaluation, Z-score Standardization

¹Doctoral student, Faculty of Education , Chiang Rai Rajabhat University

²lecturer, Faculty of Education , Chiang Rai Rajabhat University

³lecturer, Faculty of Faculty of Humanities and Social Sciences, Chiang Rai Rajabhat University

Introduction

China has continued to promote the development and reform of ideological and political theory (IPC) courses in higher vocational education, positioning IPC teaching as an important vehicle for students' moral development and value formation. In this context, IPC teachers are the key agents in fostering virtue through education, and their professional competency directly affects teaching quality and educational outcomes (Bian, & Siththad., 2024: 3796 – 3807; Wang, 2018: 81 – 88). At the same time, vocational colleges are characterized by substantial differences in students' academic foundations, a pronounced practice orientation of curricula, and relatively constrained teaching resources. These features require IPC teachers to translate ideological and political content into implementable and actionable teaching, and to effectively align it with vocational learning tasks. In this study, “teacher competency” is defined as a multidimensional ability structure that supports effective instructional design and classroom implementation, encompassing dimensions such as knowledge literacy, teaching ability, professional attitudes, and personal attributes (Zhou, & Xiang. 2019).

Although teacher development has received extensive attention, competency assessment for IPC teachers still often relies on single-source evidence (e.g., teacher self-assessment or managerial judgments), thereby weakening diagnostic accuracy and the usability of management decisions. Teacher self-assessment may be affected by self-enhancement bias, whereas student evaluation may overemphasize observable classroom performance and relatively underestimate “invisible” competencies such as theoretical mastery and instructional design (Antera. 2021: 459 – 479). Therefore, a key research gap is that, in the context of IPC courses in vocational colleges, there remains a lack of a standardized diagnostic pathway that can integrate teachers' (internal perspective) and students' (external perspective) inputs while achieving cross-source comparability, so as to reveal structural competency gaps and translate results into actionable priorities for teacher development.

To address this gap, this study takes College H as a case and adopts a “dual-source evaluation” design that integrates teacher self-assessment and student evaluation. Methodologically, Z-score standardization is first used to align the two types of ratings onto a common measurement scale to enhance cross-source comparability; a composite demand index is then constructed through weighted integration. As an operational starting point, the weights follow a transparent, a priori rule (teachers = 0.6; students = 0.4): on the one hand, this reflects that teachers are closer to professional knowledge and instructional planning processes; on the other hand, it retains students' crucial evidence regarding instructional implementation and learning experiences. Meanwhile, we emphasize that the weighting scheme can be adjusted according to different evaluation purposes and policy orientations. The specific objectives are to: (1) describe dual-source demand characteristics at the overall, dimension, and indicator levels; (2) test distributional assumptions to support Z-score standardization and integrative analyses; (3) compute the composite demand index D and classify the four dimensions and 22 items into high/medium/low demand levels to identify priority areas; and (4) provide empirical evidence, based on the demand-level results, to inform managerial decisions such as tiered professional development and resource allocation.

Research Objectives

- 1 To analyze the differences between teachers and students across each competency dimension and indicator.
- 2 To verify the reliability, validity, and distributional suitability of the dual-source data.
- 3 To develop a weighted integrated composite demand model, thereby providing empirical support for teacher training and development strategies.

Literature Review

Teacher professional competency is generally defined as an integrated capability structure composed of knowledge, skills, attitudes, and values, serving as the foundation for effective teaching and continuous professional development (OECD, 2005). From the perspective of the knowledge base, Shulman's concept of pedagogical content knowledge (PCK) emphasizes the integration of subject knowledge and instructional strategies, highlighting teachers' ability to transform knowledge into teachable content (Shulman, 1986: 4 – 14). From a competency perspective, related studies further distinguish between explicit competencies (observable knowledge and skills) and implicit competencies (professional attitudes and personality traits that are less directly observable but exert significant influence on teaching) (Boyatzis, 1982). In IPC teaching, competency requirements are further intensified due to the dual mission of value cultivation and learning facilitation, while also being shaped by differences in vocational students' academic foundations and practice-oriented learning environments. To better capture this context, Zhou, & Xiang (2019) developed a four-dimension, 22-item competency framework for IPC teachers (knowledge literacy, teaching ability, professional attitudes, and personal attributes), which provides the theoretical basis for variable selection in this study as well as for interpreting structural gaps between explicit and implicit competencies (Zhou, & Xiang, 2019).

At the assessment level, different information sources tend to capture different aspects of teaching. Teacher self-assessment supports reflective practice and can cover professional knowledge and instructional planning processes, but it is susceptible to self-enhancement bias and halo effects (Novotová, & Nišpanská, 2017: 54 – 73). Student evaluation more effectively reflects observable dimensions such as classroom interaction, instructional implementation behaviors, and teaching charisma, yet it may underestimate “invisible competencies,” including theoretical mastery and instructional design (Tait-McCutcheon, & Knewstubb, 2018: 772 – 785). Consequently, multi-source assessment emphasizes the integration of complementary perspectives to enhance diagnostic validity and decision usability, and teacher-student dual-source evaluation can be regarded as a feasible and contextually appropriate simplified form of broader multi-source assessment (Cheong, 2017: 1 – 20). However, when focusing on IPC teachers in higher vocational education, existing studies remain constrained by single-source measurement, insufficient attention to the higher vocational IPC context, and analyses that largely remain at descriptive or correlational levels, with limited translation of evidence into outputs that can be directly embedded in management decision-making (e.g., operational competency levels or graded support intensity). Therefore, this study adopts a dual-source design and introduces an operational classification logic to transform competency evidence into management priorities for the development of IPC teachers in higher vocational education.

Research Methodology

1 Research Design

This study adopts a quantitative, cross-sectional research design and collects data using dual-source, structurally equivalent questionnaires. Within the applied context of identifying professional competency needs of IPC teachers in vocational colleges, a comprehensive diagnostic framework integrating teacher self-assessment and student evaluation was constructed. The research procedure is illustrated as follows:

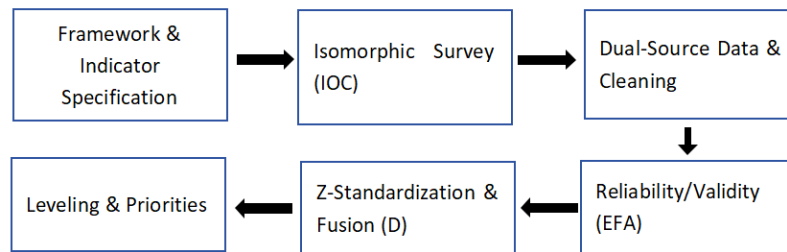


Figure 1 Research Workflow Diagram

2 Research Instrument Design and Reliability and Validity Testing

2.1 Research Instrument Design

Based on established theories of teacher professional competency, this study constructed the assessment indicator system according to the four-dimension, 22-item model proposed by Zhou and Xiang (2019). A structurally equivalent dual-source questionnaire was developed, in which the teacher version and student version differed only in the referenced subject, with item content, dimensional structure, and rating scales kept consistent. The questionnaire contained a total of 22 items measured using a five-point Likert scale (1 = very low need, 5 = very high need), covering four dimensions: knowledge literacy, teaching ability, professional attitudes, and personal attributes.

To verify content validity, this study employed the Item–Objective Congruence (IOC) index. Five experts in educational management were invited to independently evaluate the degree of alignment between each item and the research variables (−1 = inappropriate, 0 = neutral, +1 = appropriate). The calculated IOC values ($IOC = \sum \text{expert ratings} / \text{number of experts}$) all exceeded 0.60, meeting the criterion proposed by Lynn (1986). No items required deletion; several items were slightly revised in wording based on expert suggestions and subsequently reconfirmed as acceptable.

2.2 Research Sample, Data Collection, and Data Cleaning

The study was conducted in one Chinese higher vocational college (College H). The teacher sample was obtained through a full-coverage survey including all full-time IPC teachers currently employed at College H. The student sample was selected using stratified random sampling, with faculties/majors as the primary strata and proportional sampling within strata based on grade level and gender. Eligible participants were students who were currently enrolled in the relevant courses during the semester and had sufficient classroom experience.

Data were collected through an online questionnaire following the principles of anonymity, voluntariness, and informed consent. The first page of the questionnaire explained the research purpose and participants' rights, and respondents proceeded only after providing consent. No identifiable information such as names, student identification numbers, or contact details was collected, and all data

were used solely for academic research. Considering potential pressure arising from teacher–student power relations, student questionnaires were distributed and collected centrally by the research team rather than by course instructors. Participants were explicitly informed that non-participation or withdrawal would not affect grades or any rights, and results were reported only in aggregated statistical form to prevent individual identification. A total of 21 teacher questionnaires and 420 student questionnaires were distributed. All teacher questionnaires were returned and valid ($n = 21$, response rate = 100%). Among the student questionnaires, 420 were returned, of which 414 were valid (valid response rate = 98.6%).

2.3 Reliability and Validity Testing

To ensure that the dual-source, structurally equivalent questionnaire was suitable for subsequent cross-source comparison on a common scale and integrative analysis, this study first examined measurement quality and factor-analysis suitability. Specifically, Cronbach's α and corrected item–total correlation (CITC) were used to assess internal consistency and item discrimination. In addition, the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity were conducted on the student sample to determine whether the data satisfied the prerequisites for exploratory factor analysis (EFA). The relevant results are summarized in Table 1.

Table 1 Summary of Measurement Quality and Factorability Diagnostics

Indicator	Teacher Group (n=21)	Student Group (n=414)
Cronbach's α (total scale)	0.959	0.942
CITC range	0.353–0.921	0.507–0.746
KMO	-	0.955
Bartlett's Test of Sphericity	-	$\chi^2(231)=5178.827$, $p<0.001$

Note: Due to the relatively small teacher sample size ($n = 21$), EFA/CFA was not conducted for the teacher group. Under the isomorphic design, IOC and reliability indicators were used as acceptable substitute evidence.

As shown in Table 1, the scale demonstrated high internal consistency in both the teacher and student groups (teacher group Cronbach's $\alpha = 0.959$; student group Cronbach's $\alpha = 0.942$). The CITC values for both groups reached acceptable thresholds (teacher group: 0.353–0.921; student group: 0.507–0.746), indicating satisfactory item homogeneity and discrimination. For the student group, the KMO value was 0.955 and Bartlett's test of sphericity was significant ($\chi^2(231) = 5178.827$, $p<0.001$), suggesting that the correlation matrix was suitable for factor analysis and providing sufficient statistical prerequisites for subsequent EFA.

On the basis of meeting the assumptions required for factor analysis, this study further conducted exploratory factor analysis (EFA) using the student sample, and the key results concerning construct validity are summarized in Table 2.

Table 2 Structural Validity Summary from Student-Sample EFA (PAF–Oblimin)

Factor	Extraction Eigenvalue	Variance Explained (%)	Cumulative Variance Explained(%)	Pattern Matrix Loading Range
1	9.655	43.888	43.888	0.627–0.818
2	1.423	6.469	50.357	0.724–0.841
3	0.933	4.240	54.597	0.552–0.846

Factor	Extraction Eigenvalue	Variance Explained (%)	Cumulative Variance Explained(%)	Pattern Matrix Loading Range
4	0.787	3.576	58.173	0.473–0.838

Note: Factor correlation range (oblique rotation): 0.423–0.599. EFA was conducted using principal axis factoring (PAF) with Oblimin oblique rotation. Under oblique rotation, the squared loadings in the rotated solution cannot be directly summed to obtain total variance; therefore, the explained variance reported here is based on the Extraction (sum of squared loadings) results.

As shown in Table 2, after applying principal axis factoring (PAF) with Kaiser-normalized Oblimin oblique rotation, the theoretically proposed four-factor structure was supported. The extracted eigenvalues for the four factors were 9.655, 1.423, 0.933, and 0.787, corresponding to variance explained rates of 43.888%, 6.469%, 4.240%, and 3.576%, respectively, with a cumulative explained variance of 58.173%. The Pattern Matrix indicated that the primary loading ranges for the four factors were 0.627–0.818, 0.724–0.841, 0.552–0.846, and 0.473–0.838, demonstrating satisfactory convergent validity of the items within their respective factors. Meanwhile, factor correlations ranged from 0.423 to 0.599, consistent with the expectation that related constructs may be interrelated in practice. Because oblique rotation was employed, the rotated sums of squared loadings cannot be directly aggregated to calculate total variance; therefore, this study consistently reports the eigenvalues and explained variance based on the Extraction (sum of squared loadings) results.

In summary, this study constructed a dual-source, structurally equivalent questionnaire based on the four-dimension, 22-item framework. Through internal consistency testing in both the teacher and student groups, content validity evaluation, and the KMO/Bartlett tests and EFA results from the student sample, the scale was verified to possess sound reliability and validity. The student-sample EFA supported the theoretically specified four-factor structure, and both cumulative explained variance and primary factor loadings reached acceptable levels, thereby providing a reliable measurement foundation for subsequent cross-source comparison on a common scale and weighted integrative analysis.

3. Data Analysis

Data analysis was primarily conducted using SPSS 29.0. Statistical notation was standardized, and all numerical values were reported to two decimal places. The analytical procedures included the following :

3.1 Normality Testing

Considering differences in sample size and the characteristics of Likert-scale data, the Shapiro–Wilk (S–W) test was applied to the teacher sample ($n = 21$), while the Kolmogorov–Smirnov (K–S) test was used for the student sample ($n = 414$). Graphical inspections using Q–Q plots and histograms were conducted as supplementary verification. In addition, the sensitivity of the K–S test to minor deviations under large-sample conditions was explicitly acknowledged to avoid misinterpreting statistical significance as substantive non-normality.

3.2 Z-score Standardization

To reduce the influence of differences in mean levels, rating tendencies, and dispersion between the two data sources on comparative results, and to achieve cross-source comparability, Z-score transformation was conducted separately within the teacher group and the student group for all 22

indicators. Z-score standardization converts raw scores into standardized scores relative to the distribution of each source, aligning data from different sources onto a common scale (mean = 0, standard deviation = 1) without altering relative ranking information. This procedure provides a common measurement basis for subsequent integration and classification (Field, 2024). The calculation formula is as follows:

$$Z = \frac{X - \mu}{\sigma}$$

where X represents the raw score, μ denotes the mean of the corresponding dataset, and σ represents the standard deviation of that dataset. The resulting standardized scores, Z_t (teachers), Z_s (students), were used for subsequent cross-source integration and structural difference comparisons.

3.3 Weighted Integration and Composite Demand Score Calculation

After completing scale alignment, to comprehensively reflect both “professional self-knowledge” (teacher perspective) and “learner experience” (student perspective) as mutually corroborating sources, this study employed weighted integration to calculate the composite demand score D :

$$D = W_t \times Z_t + W_s \times Z_s$$

where the weights were set as $W_t = 0.6$, $W_s = 0.4$, that is:

$$D = 0.6 \times Z_t + 0.4 \times Z_s$$

The weighting scheme followed the principles of information accessibility, role relevance, and decision purpose. Teachers are closer to aspects that students cannot directly observe, such as curriculum goal setting, instructional design, and mastery of professional knowledge; therefore, teacher self-assessment contains relatively higher informational value for diagnosing professional competency needs. Student evaluation, in contrast, more effectively reflects learning-process experiences, including classroom interaction, instructional implementation, and personal influence, and thus serves as indispensable external evidence. Given the applied context of this study, an a priori weighting scheme of “slightly higher weight for teachers while retaining student input” was adopted as a transparent and replicable operational starting point, balancing the two sources of evidence while reducing the risk of single-source bias. It should be emphasized that the 0.6/0.4 weighting is not a universal constant but may be adjusted according to institutional evaluation purposes, policy orientations, or data quality requirements. In this study, the weights were predetermined prior to analysis to avoid hypothesis risks associated with post hoc adjustment, and the decision rules were applied consistently throughout the study.

3.4 Threshold Classification and Priority Formation

To transform the composite demand score D into demand levels that are interpretable and applicable for management purposes, this study established a three-level threshold system based on the interval interpretation of Z-scores under the standard normal distribution, with reference to the commonly used classification logic distinguishing scores “around the average level” from those “deviating from the average” (Gravetter, et.al. 2021). Specifically, ± 0.5 was adopted as the boundary of the “moderate range.” On the one hand, this threshold differentiates items that are close to the overall mean (moderate demand) from those that are clearly above or below the mean (high/low demand). On the other hand, the threshold setting is simple and transparent, facilitating cross-indicator comparison and management communication. The classification criteria are as follows:

$D > 0.5$: classified as high current demand;

$-0.5 \leq D \leq 0.5$: classified as moderate current demand;

$D \leq -0.5$: classified as low current demand.

Based on these criteria, level classifications and priority rankings were established for all 22 indicators, and aggregated comparisons were conducted at the dimensional level across the four primary constructs to support subsequent decision-making discussions regarding teacher professional development priorities, support intensity, and resource allocation.

Research Results

1 Descriptive Statistics

Table 3 Basic Characteristics of Participants

Group	Teacher Group	Student Group
Sample size (n)	21	414
Gender (% female/male)	76.2 / 23.8	50.7 / 49.3
Grade level (% 2023/2024 cohort)	—	48.8 / 51.2
Position (% teacher/administrator)	66.7 / 33.3	—
Teaching experience (% 1–3 years / 4–6 years / 7–10 years / 10+ years)	57.1 / 4.8 / 9.5 / 28.6	—
Educational level (% bachelor's/master's)	19.0 / 81.0	—
Academic rank (% assistant lecturer/lecturer/associate professor)	71.4 / 19.0 / 9.5	—
Coverage of secondary College/faculties	All teachers from the College of Marxism	Coverage across 11 secondary College/faculties (individual coverage: 8.9%–9.2%)

Note: “—” indicates that the variable is not applicable to the corresponding group.

As shown in Table 3, after data cleaning, the valid teacher self-assessment sample was $n = 21$ (covering all full-time IPC teachers from the College of Marxism), and the valid student evaluation sample was $n = 414$ (covering 11 secondary College/faculties). The student sample was relatively balanced in terms of gender and grade distribution.

Table 4 Means and Standard Deviations of Teacher Self-Assessment and Student Evaluation

Indicator	Teacher Group		Student Group	
	Mean	SD	Mean	SD
1. Knowledge Literacy	3.84	0.83	3.83	0.77
1.1 Marxist theoretical knowledge literacy	3.76	1.04	2.97	1.02
1.2 Humanities and social sciences knowledge literacy	4.00	0.78	4.24	0.88
1.3 Educational theory knowledge literacy	3.76	0.83	4.28	0.84
2. Teaching Ability	4.00	0.64	3.82	0.76
2.1 Learning ability	4.10	0.77	3.65	1.01
2.2 Innovation ability	3.48	0.93	3.64	1.06
2.3 Research competence	4.29	0.72	3.77	0.98

Indicator	Teacher Group		Student Group	
	Mean	SD	Mean	SD
2.4 Communication ability	4.10	0.94	3.91	0.94
2.5 Information acquisition ability	4.10	0.89	3.68	1.04
2.6 Teaching method application ability	4.10	0.77	3.19	1.07
2.7 Multimedia expression ability	4.00	0.78	4.10	0.93
2.8 Language expression ability	4.10	0.77	4.22	0.80
2.9 Teaching evaluation ability	3.95	0.81	4.14	0.87
2.10 Instructional design ability	4.14	0.73	3.87	1.01
2.11 Classroom management and organization ability	3.76	0.94	4.01	0.93
2.12 Curriculum system integration ability	3.95	0.97	3.64	1.04
3. Professional Attitudes	3.35	0.90	3.07	0.82
3.1 Dedication to teaching	3.14	0.85	3.21	0.96
3.2 Caring for students	3.43	0.87	2.96	0.93
3.3 Rigorous scholarship	3.48	1.12	3.03	1.04
4. Personal Attributes	3.14	0.78	3.27	0.83
4.1 Professional ethics	3.43	1.03	2.98	1.04
4.2 Teacher charisma	2.81	0.81	3.34	0.97
4.3 Approachability	3.00	0.78	3.29	1.04
4.4 Social insight	3.33	0.91	3.48	0.89
Overall	3.58	0.69	3.50	0.63

As shown in Table 4, the overall means of the teacher and student groups were close (teacher group $M = 3.58$, $SD = 0.69$; student group $M = 3.50$, $SD = 0.63$), indicating a high level of agreement between the two sources regarding overall competency development needs.

At the dimensional level, both groups exhibited similar structural distribution patterns: compared with implicit competency dimensions such as professional attitudes and personal attributes, explicit competency dimensions—namely knowledge literacy and teaching ability—received higher scores. Specifically, the dimensional means of the teacher group, from highest to lowest, were teaching ability ($M = 4.00$, $SD = 0.64$), knowledge literacy ($M = 3.84$, $SD = 0.83$), professional attitudes ($M = 3.35$, $SD = 0.90$), and personal attributes ($M = 3.14$, $SD = 0.78$). In the student group, knowledge literacy ($M = 3.83$, $SD = 0.77$) and teaching ability ($M = 3.82$, $SD = 0.76$) were relatively higher, followed by personal attributes ($M = 3.27$, $SD = 0.83$), while professional attitudes were the lowest ($M = 3.07$, $SD = 0.82$). These results indicate that dual-source evidence shows consistency in a demand structure characterized by the prioritization of explicit competencies and the relative weakening of implicit competencies.

2. Normality Testing

To support the statistical assumptions required for subsequent common-scale processing (Z-score standardization), normality tests were conducted on both dimensional and overall scores. The Shapiro–Wilk test was applied to the teacher group ($n = 21$), and the Kolmogorov–Smirnov test was applied to the student group ($n = 414$), supplemented by graphical inspection using histograms and Q–Q plots.

Table 5 Shapiro–Wilk Normality Test Results for Teacher Self-Assessment

Dimension	S–W Statistic	df	p value	Normality Decision
Knowledge Literacy	0.937	21	0.188	Yes
Teaching Ability	0.968	21	0.693	Yes
Professional Attitudes	0.977	21	0.873	Yes
Personal Attributes	0.967	21	0.675	Yes
Overall	0.978	21	0.894	Yes

Table 6 Kolmogorov–Smirnov Normality Test Results for Student Evaluation

Dimension	K–S Statistic	df	p value	Normality Decision
Knowledge Literacy	0.154	414	<0.001	No
Teaching Ability	0.073	414	<0.001	No
Professional Attitudes	0.149	414	<0.001	No
Personal Attributes	0.079	414	<0.001	No
Overall	0.043	414	0.067	Yes(p > 0.05)

The test results indicate that none of the four dimensions or the overall score in the teacher group significantly deviated from normal distribution ($p > 0.05$). As shown in Table 6 (student-group K–S test), all four dimensions in the student group reached statistical significance under the K–S test ($p < 0.001$), whereas the overall score did not significantly deviate from normality ($p = 0.067$). Considering the high sensitivity of the K–S test to minor deviations under large-sample conditions, as well as the common boundary effects associated with five-point Likert-scale data, this study further incorporated graphical inspections (histograms and Q–Q plots). These visual checks indicated that the overall distribution was approximately normal, with only slight deviations at the distribution tails. Therefore, the data from both groups were considered suitable for Z-score standardization and subsequent integrative analysis.

3. Integration of Z-scores and Level Classification

To achieve cross-source comparability, this study first conducted Z-score standardization within the teacher and student groups, obtaining Z_t (teachers) and Z_s (students). A weighted integration was then applied to compute the composite demand index $D = 0.6 \times Z_t + 0.4 \times Z_s$, where D represents the “composite relative demand intensity” of each competency relative to the overall distribution across both sources. Demand levels were classified using a three-level division based on standard score intervals (± 0.50 as the boundary of the moderate range), categorizing D into three groups: high, moderate, and low, as shown in Table 7.

Table 7 Integration of Teacher and Student Competency Data and Level Classification

Indicator	Teacher		Student		D Value	Evaluation
	Mean	Z Value	Mean	Z Value		
1. Knowledge Literacy	3.84	+0.328	3.83	+0.551	+0.417	Moderate
1.1 Marxist theoretical knowledge literacy	3.76	+0.129	2.97	-1.474	-0.512	Low
1.2 Humanities and social sciences knowledge literacy	4.00	+0.727	4.24	+1.516	+1.043	High
1.3 Educational theory knowledge literacy	3.76	+0.129	4.28	+1.611	+0.722	High

Indicator	Teacher		Student		D Value	Evaluation
	Mean	Z Value	Mean	Z Value		
2. Teaching Ability	4.00	+0.727	3.82	+0.528	+0.647	High
2.1 Learning ability	4.10	+0.976	3.65	+0.127	+0.636	High
2.2 Innovation ability	3.48	-0.568	3.64	+0.104	-0.299	Moderate
2.3 Research competence	4.29	+1.449	3.77	+0.410	+1.033	High
2.4 Communication ability	4.10	+0.976	3.91	+0.739	+0.881	High
2.5 Information acquisition ability	4.10	+0.976	3.68	+0.198	+0.665	High
2.6 Teaching method application ability	4.10	+0.976	3.19	-0.956	+0.203	Moderate
2.7 Multimedia expression ability	4.00	+0.727	4.10	+1.187	+0.911	High
2.8 Language expression ability	4.10	+0.976	4.22	+1.469	+1.173	High
2.9 Teaching evaluation ability	3.95	+0.602	4.14	+1.281	+0.874	High
2.10 Instructional design ability	4.14	+1.076	3.87	+0.645	+0.904	High
2.11 Classroom management and organization ability	3.76	+0.129	4.01	+0.975	+0.467	Moderate
2.12 Curriculum system integration ability	3.95	+0.602	3.64	+0.104	+0.403	Moderate
3. Professional Attitudes	3.35	-0.892	3.07	-1.238	-1.030	Low
3.1 Dedication to teaching	3.14	-1.415	3.21	-0.909	-1.213	Low
3.2 Caring for students	3.43	-0.693	2.96	-1.497	-1.015	Low
3.3 Rigorous scholarship	3.48	-0.568	3.03	-1.332	-0.874	Low
4. Personal Attributes	3.14	-1.415	3.27	-0.767	-1.156	Low
4.1 Professional ethics	3.43	-0.693	2.98	-1.450	-0.996	Low
4.2 Teacher charisma	2.81	-2.237	3.34	-0.603	-1.583	Low
4.3 Approachability	3.00	-1.764	3.29	-0.720	-1.346	Low
4.4 Social insight	3.33	-0.942	3.48	-0.273	-0.674	Low
Overall	3.58	-0.319	3.50	-0.226	-0.282	Moderate

3.1 Integrated Results at the Dimensional Level

At the dimensional level, the composite demand index (D) exhibited a clear structural distribution across the four primary dimensions. Teaching ability was classified as a high-demand dimension ($D = +0.647$), knowledge literacy as a moderate-demand dimension ($D = +0.417$), while both professional attitudes ($D = -1.030$) and personal attributes ($D = -1.156$) fell into the low-demand range. The overall composite score was also categorized as moderate demand ($D = -0.282$). These results indicate that, after integrating evidence from teachers' internal perspectives and students' external perspectives, professional development needs of IPC teachers are primarily concentrated in explicit competency dimensions such as knowledge and teaching, whereas implicit competency dimensions related to attitudes and personal attributes demonstrate relatively weaker demand intensity within the current sample.

3.2 Distribution of Demand Levels at the Indicator Level

The results further show a clear hierarchical distribution of demand intensity across the 22 competency indicators. First, a total of 10 indicators were classified as high demand ($D > 0.50$), mainly concentrated in the explicit competency dimensions of knowledge literacy and teaching ability, particularly

reflecting competencies related to classroom presentation, instructional implementation, and academic support. Among these, language expression ability ($D = +1.173$), humanities and social sciences knowledge literacy ($D = +1.043$), and research competence ($D = +1.033$) represented the highest-demand items. Second, four indicators were categorized as moderate demand ($-0.50 \leq D \leq 0.50$), including innovation ability, teaching method application ability, classroom management and organization ability, and curriculum system integration ability. This suggests that although these competencies still have room for improvement, their relative urgency is lower than that of high-demand items. Finally, eight indicators were classified as low demand ($D < -0.50$), primarily concentrated in implicit competency dimensions such as professional attitudes and personal attributes, while also including Marxist theoretical knowledge literacy ($D = -0.512$), indicating that these competencies show comparatively weaker demand intensity after dual-source integration.

Research Conclusions

Based on data collected through the dual-source, structurally equivalent questionnaires, this study first provided a descriptive analysis of teacher self-assessment and student evaluation within the four-dimension, 22-item framework. The results indicated that the overall means of the two groups were similar, and both exhibited a consistent structural distribution pattern across dimensions: explicit competency dimensions, such as knowledge literacy and teaching ability, received relatively higher scores, whereas implicit competency dimensions, including professional attitudes and personal attributes, received comparatively lower scores.

Second, to support the statistical assumptions required for cross-source comparison on a common scale, normality tests were conducted on both dimensional and overall scores and supplemented with graphical inspections. None of the dimensions or the overall score in the teacher group significantly deviated from normal distribution. Although the K-S tests at the dimensional level in the student group were statistically significant, the overall score did not significantly deviate from normality. Considering the sensitivity of large samples and the graphical inspection results, the data overall supported the feasibility of implementing Z-score standardization and subsequent integrative analysis.

On this basis, dual-source ratings were standardized using Z-scores, and the composite demand index (D) was calculated. According to the predetermined threshold rules, the four dimensions and 22 indicators were classified into high, moderate, and low demand levels. At the dimensional level, teaching ability was identified as high demand, knowledge literacy as moderate demand, and both professional attitudes and personal attributes as low demand, while the overall composite score indicated moderate demand. At the indicator level, the classification further revealed a clear hierarchical distribution: high-demand items were mainly concentrated in indicators related to knowledge literacy and teaching ability; moderate-demand items were associated with competencies related to instructional process optimization and classroom organization; and low-demand items were primarily located within the dimensions of professional attitudes and personal attributes, while also including certain knowledge indicators such as Marxist theoretical knowledge literacy.

Overall, the demand for improving teacher competency demonstrates a stronger emphasis on practical teaching skills and knowledge expansion, alongside relatively insufficient attention to professional

commitment and personal cultivation. This finding suggests that future teacher development strategies should continue to strengthen teaching ability while also promoting professional identity, sense of responsibility, and comprehensive personal development, thereby fostering more balanced and holistic professional growth among teachers.

Discussion

1. Core Findings and Mechanism Interpretation

This study shows that while the two sources converged in their overall judgments, the demand structure exhibited a stable pattern characterized by the prioritization of explicit competencies and the relative weakening of implicit competencies. This pattern can be interpreted through the competency “iceberg model” and the distinction between explicit and implicit competencies: knowledge and skills are more visible, trainable, and assessable, whereas attitudes and personal traits are more latent and less easily captured by conventional evaluation systems (Boyatzis. 1982). From the perspective of evaluation mechanisms, teacher self-assessment is susceptible to self-enhancement and social desirability effects and tends to focus on competency domains more closely associated with performance accountability and quantifiable indicators. In contrast, student evaluation is driven by observability and classroom interaction cues, relying more heavily on instructional implementation and interpersonal influence when forming judgments (Novotová, & Nišponská. 2017: 54 - 73). Furthermore, students’ stronger emphasis on pedagogical theoretical foundations can be aligned with Shulman’s PCK perspective: learners often perceive whether teachers possess the ability to transform disciplinary knowledge into learnable content through classroom processes such as explanation, demonstration, and feedback (Shulman. 1986: 4 – 14). Therefore, the phenomenon of “explicit agreement but implicit divergence” should not be regarded as noise; rather, it reflects differences in attention focus and evidence sources among evaluation subjects when assessing the same competency structure.

2. Academic Contributions

The primary contribution of this study lies in proposing a replicable diagnostic pathway characterized by “dual-source input—common-scale alignment—level-based output.” By applying Z-score standardization to enhance cross-source comparability and using a composite index to translate multi-source evidence into interpretable managerial signals, the study provides a practical analytical framework. Its theoretical foundation aligns with the logic of multi-source feedback (360-degree feedback), which suggests that single-source evaluations often contain systematic bias, whereas integrating multiple sources improves diagnostic robustness and decision effectiveness (London, & Smither. 1995: 803 – 839). Unlike prior studies that typically remain at parallel comparisons between self-assessment and course evaluation, this study further compresses dual-source information into operational level-based outputs, enabling integration into management cycles such as resource allocation, training prioritization, and support-intensity design. Theoretically, the findings also extend the four-dimension framework proposed by Zhou, and Xiang (2019): the same framework may activate different priority structures under teacher versus student evaluation perspectives, indicating that competency models in practical applications must account for evaluator differences and mechanisms of evidence integration (Zhou, & Xiang. 2019).

Recommendations

1. Practical Implications

In the context of IPC courses in higher vocational education, the practical implications of this study are reflected in two main aspects:

1.1 Application of Dual-Source Evaluation

The composite index can be used to determine overall priority rankings, while discrepancies between teacher and student evaluations may help identify risks of perceptual misalignment. Periodic reassessment can be conducted to monitor level transitions and evaluate intervention effectiveness, thereby forming a closed-loop process of “diagnosis–intervention–review.”

1.2 Application in Management Practice

Based on high-, moderate-, and low-demand classifications, institutions may adopt a tiered governance approach. High-demand items should be prioritized through targeted training programs and short-cycle improvement indicators. Moderate-demand items may be enhanced through peer collaboration, classroom observation, and instructional coaching to support iterative improvement. Low-demand dimensions should not be interpreted as unimportant; rather, they may represent implicit developmental areas that are easily overlooked. These should be institutionally embedded through mentoring systems, reflective practice, and mechanisms supporting professional identity development.

2. Research Limitations and Future Research Directions

Although this study demonstrates methodological and analytical innovation, several limitations remain and should be addressed in future research.

First, the sample size and structure limit external validity. Although the teacher sample covered the entire institution ($n = 21$), the number remained relatively small. While the student sample was sufficient ($n = 414$), no stratified analysis by gender or academic major was conducted, which may have obscured group differences. Future studies should expand the sample scope to include teachers from different regions, disciplines, and institutional levels, and employ stratified sampling or hierarchical linear modeling (HLM) to enhance representativeness and robustness.

Second, the scale relied on self-reported data, which may be influenced by social desirability and response bias. Although reliability and validity were tested, subjective evaluations may still overestimate or underestimate actual competencies. Future research is encouraged to integrate multiple sources of evidence—such as classroom observation, interviews, peer evaluation, and teaching portfolios—to construct a triangulated validation framework combining self-assessment, external evaluation, and empirical evidence.

Finally, the cross-sectional design limits the ability to capture the dynamic evolution of teacher competency. Future studies may adopt longitudinal tracking or quasi-experimental designs to examine competency changes across professional development programs, instructional reforms, and career-stage transitions, thereby revealing developmental trajectories and key influencing factors.

References

- Antera, S. (2021). Professional competence of vocational teachers: A conceptual review. *Vocations and Learning*, 14(3), 459 - 479.
- Bian, X., & Siththada, T.(2024). Teacher competency development model of higher vocational schools in Shandong province, China. *Edelweiss Applied Science and Technology*, 8(6), 3796-3807. <https://doi.org/10.55214/25768484.v8i6.2823>
- Boyatzis, R. E. (1982). *The Competent Manager: A Model for Effective Performance*. New York: John Wiley & Sons.
- Cheong, G. (2017). Developing an instrument for self-evaluation of teaching and learning competencies: A review of faculty professional development and the changing higher education landscape in Singapore. *Education Research and Perspectives*, 44, 1 - 20.
- Field, A. (2024). *Discovering statistics using IBM SPSS statistics*. Sage publications limited.
- Gravetter, F. J., Wallnau, L. B., Forzano, L. A. B., & Witnauer, J. E. (2021). *Essentials of statistics for the behavioral sciences* (p. 648). Boston, MA, USA: Cengage.
- London, M., & Smither, J. W. (1995). Can multi-source feedback change perceptions of goal accomplishment, self-evaluations, and performance-related outcomes? *Theory-based applications and directions for research. Personnel psychology*, 48(4), 803-839.
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing research*, 35(6), 382-386.
- Novotová, J., & Nišpanská, M. (2017). Self-assessment and assessment of professional competencies of trainee teachers during the block teaching practice. *Lifelong Learning - celoživotní vzdělávání*, 7(1), 54-73. <https://doi.org/10.11118/lifele2017070154>
- OECD (2005), *Teachers Matter: Attracting, Developing and Retaining Effective Teachers*, Education and Training Policy. OECD Publishing. Paris, <https://doi.org/10.1787/9789264018044-en>.
- Shulman, L. S. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher*, 15(2), 4-14. <https://doi.org/10.3102/0013189X015002004>
- Tait-McCutcheon, S., & Knewstubb, B. (2018). Evaluating the alignment of self, peer and lecture assessment in an Aotearoa New Zealand pre-service teacher education course. *Assessment & Evaluation in Higher Education*, 43(5), 772-785.
- Wang, G. (2018). Cultivation of Teachers' Ideological and Political Quality in Colleges and Universities in the New Era. *International Conference on Contemporary Education, Social Sciences and Ecological Studies (CESSES 2018)* (pp. 81-88). Atlantis Press.
- Zhou, L., & Xiang, Q. (2019). *Research on the construction of teaching competence of ideological and political theory course teachers*. Nanchang: Jiangxi Higher Education Press. [in Chinese]