International Journal of Pedagogy and Teacher Education

Journal Homepage: jurnal.uns.ac.id/ijpte

Assessing Teachers' Pedagogical Performance under the Merdeka Curriculum in Indonesian Vocational School

Ismiyati Ismiyati[1]*, Tusyanah Tusyanah[1], Mar'atus Sholikah[2], Teguh Hadi Raharjo[1], Sulistyawati Sulistyawati[1], Djoko Legowo[1]

[1]Office Administration Education Study Program, Faculty of Economics and Business, Universitas Negeri Semarang, Semarang, Indonesia

[2]International Business Administration Study Program, Politeknik Balekambang Jepara, Jepara, Indonesia

ARTICLE INFO

Article History

Received: October 16, 2025

1st Revision: March 10, 2026

Accepted: April 22, 2026

Available Online: April 30, 2026

Keywords:

Pedagogical Performance; Merdeka Curriculum; Cluster Analysis; Vocational High School Teachers; Teaching Competence.

***Corresponding Author**

Email address:

ismiyati@mail.unnes.ac.id

ABSTRACT

The implementation of the Merdeka Curriculum in Indonesia requires teachers to demonstrate strong pedagogical performance aligned with student-centered learning, flexible instructional design, and competency-based assessment. However, empirical evidence on variations in teachers' pedagogical performance, particularly in vocational education, remains limited. This study aims to identify clusters of teachers' pedagogical performance and to describe the characteristics of each cluster within the context of the Merdeka Curriculum. A quantitative approach was employed involving 99 vocational high school teachers from the Office Management and Business Services (OMBS) program in Central Java, Indonesia. Data were collected using a structured questionnaire covering lesson planning, instructional implementation, classroom management, and assessment practices. K-means cluster analysis was applied to classify teachers based on similarities in pedagogical performance. The findings reveal three distinct clusters of pedagogical performance: low (35%), moderate (38%), and high (27%). The high-performance cluster is characterized by systematic lesson planning, active learning strategies, and consistent use of formative assessment. The moderate cluster demonstrates relatively adequate instructional planning but limited innovation in instructional practice. The low-performance cluster reflects teacher-centered practices and restricted assessment variation. These findings highlight the heterogeneous nature of teachers' pedagogical performance in implementing the Merdeka Curriculum. These findings suggest the need for targeted professional development, particularly for teachers in the moderate and low clusters, to support the adoption of more student-centered and competency-based practices.

How to cite: Ismiyati, I., Tusyanah, T., Sholikah, M., Raharjo, T. H., Sulityawati, S., & Legowo, D. (2026). Assessing Teachers' Pedagogical Performance under the Merdeka Curriculum in Indonesian Vocational School. *International Journal of Pedagogy and Teacher Education,* 10(1), 150–157. https://doi.org/10.20961/ijpte.v10i1.110020

1. INTRODUCTION

Curriculum reform places increasing demands on teachers to demonstrate strong pedagogical performance, particularly in designing and implementing student-centered learning. Effective reform depends not only on policy design but also on teachers' capacity to translate curricular principles into classroom practice. Within this study, pedagogical performance is understood as teachers' ability to plan instruction, implement learning activities, manage classrooms, and conduct assessment practices that support meaningful student learning outcomes. Empirical research consistently identifies teachers as a key determinant of student achievement and long-term academic success (Burroughs et al., 2019; Darling-Hammond et al., 2017). Contemporary frameworks of effective teaching emphasize core dimensions such as cognitive activation, classroom management, and the creation of a supportive learning environment (Baumert et al., 2013; König et al., 2021; Kunter et al., 2013). Recent developments in education have also expanded the meaning of pedagogical performance through the growing importance of technology-integrated instruction. Frameworks such as Technological Pedagogical Content Knowledge (TPACK) highlight that effective teaching increasingly requires the integration of pedagogy, content, and technology (Koehler et al., 2013). These developments suggest that pedagogical performance should be understood as a multidimensional and context-dependent construct.

Global curriculum reforms have increasingly shifted toward competency-based and student-centered approaches intended to foster critical thinking, creativity, and lifelong learning (Voogt et al., 2019). Successful

p-ISSN: 2597-7792 / e-ISSN: 2549-8525

DOI: https://doi.org/10.20961/ijpte.v10i1.110020

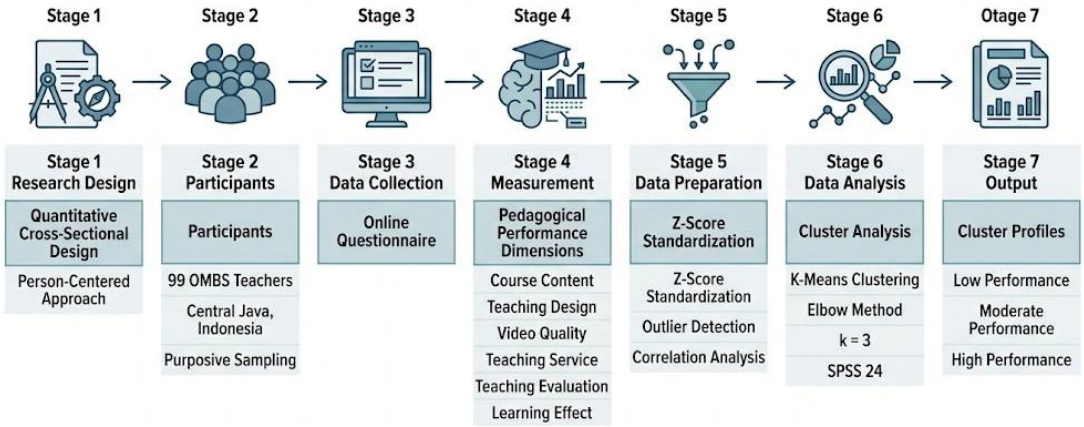
implementation of these reforms depends on teachers' ability to interpret curricular change and enact it in pedagogically meaningful ways (Pearson, 2013; Priestley et al., 2021; Puad & Ashton, 2023). Insufficient pedagogical performance may therefore hinder the translation of curriculum policy into effective classroom practice. Indonesia's Merdeka Curriculum represents a major reform that emphasizes flexible learning pathways, differentiated instruction, and competency-based assessment. Implementation of this curriculum requires teachers to design adaptive learning experiences that respond to students' needs, interests, and learning profiles (Cholilah et al., 2023). Variations in teachers' pedagogical performance are therefore likely to emerge during curriculum enactment. Identification of these variations is important because differences in pedagogical performance may shape the extent to which teachers are able to implement student-centered and competency-based learning in practice.

Research on pedagogical performance has grown substantially, yet several important gaps remain. First, many previous studies have relied primarily on variable-centered approaches, such as regression or structural equation modeling, to examine relationships among teaching-related variables (Fauth et al., 2019; Praetorius et al., 2018). Although these approaches are valuable for identifying general associations, they are less suited to capturing heterogeneous patterns of pedagogical performance across teachers. Second, empirical evidence on pedagogical performance in vocational education remains limited, despite the distinctive demands placed on vocational teachers to integrate theoretical knowledge, practical competencies, and workplace relevance (Gessler & Siemer, 2020; Rothwell et al., 2025). Third, limited research has examined pedagogical performance within the implementation of the Merdeka Curriculum using a person-centered analytical perspective. This gap is particularly important in the Indonesian context, where curriculum implementation may vary considerably across teachers and school settings.

This study addresses these gaps by applying cluster analysis to identify distinct profiles of teachers' pedagogical performance. Rather than focusing only on relationships among variables, this approach enables the classification of teachers into relatively similar groups based on shared performance characteristics. Focus on Office Management and Business Services (OMBS) teachers is particularly relevant because this vocational field requires the integration of administrative competence, digital skills, and workplace-oriented practices. Such characteristics make OMBS teachers especially responsive to curriculum changes that emphasize flexibility, technology integration, and competency-based learning. This study therefore aims to identify clusters of pedagogical performance among vocational teachers implementing the Merdeka Curriculum and to describe the characteristics of each cluster. Contribution of the study lies in three areas: it offers a person-centered perspective on pedagogical performance, extends empirical evidence from the vocational education context, and provides an analytical basis for more targeted professional development to support effective curriculum implementation.

2. MATERIAL AND METHOD

Research Design



p-ISSN: 2597-7792 / e-ISSN: 2549-8525
DOI: https://doi.org/10.20961/ijpte.v10i1.110020

Figure 1. Flowchart of the Research Design

This study employed a quantitative cross-sectional research design using a person-centered analytical approach to examine variations in teachers' pedagogical performance under the implementation of the Merdeka Curriculum. Focus of the design was not to test causal relationships among variables, but to identify relatively homogeneous groups of teachers who shared similar patterns across key pedagogical dimensions. Data were collected through an online questionnaire administered to vocational high school teachers, followed by measurement across six domains of pedagogical performance, namely course content, teaching design, video quality, teaching service, teaching evaluation, and learning effect. Data preparation involved score standardization, multivariate outlier screening, and correlation checking prior to clustering. K-means cluster analysis was then applied to classify teachers into distinct pedagogical performance profiles, with the optimal number of clusters determined using the elbow method. Figure 1 presents the overall research design and analytical flow of the study, from research design and participant selection to data collection, measurement, data preparation, clustering procedure, and cluster output.

Participants and Data Sources

The target population comprised teachers of the Office Management and Business Services (OMBS) program in vocational high schools in Central Java, Indonesia. Based on official data obtained from the provincial education authority, 32 vocational schools offering the OMBS program constituted the sampling frame. A purposive sampling technique was used to recruit teachers who were actively involved in implementing the Merdeka Curriculum in their instructional practice.

Data were collected through an online questionnaire distributed via professional teacher networks and subject-teacher association groups. The survey was administered during the period from March to June 2024. A total of 185 invitations were distributed, yielding 143 responses (response rate: 77.3%). Responses were screened for completeness, and questionnaires with missing responses on one or more core instrument items were excluded from further analysis. After this screening process, 99 valid questionnaires were retained. The inclusion criteria required respondents to:

(1) be actively teaching in the OMBS program,

(2) have direct experience implementing the Merdeka Curriculum in classroom practice, and

(3) complete all questionnaire items.

Although efforts were made to include teachers from multiple districts, the use of voluntary participation may have introduced self-selection bias, as teachers who were more engaged or more confident in their pedagogical practices may have been more likely to participate. This limitation should therefore be considered when interpreting the findings.

Measurement

Teachers' pedagogical performance was measured using a structured questionnaire adapted from Hua and Ren (2020). In the present study, the instrument was organized into six dimensions: course content, teaching design, video quality, teaching service, teaching evaluation, and learning effect. The structure of the adapted questionnaire is summarized in Table 1, including the six pedagogical-performance dimensions and their measurement focus. All items were rated on a five-point Likert scale ranging from 1 (never) to 5 (always). This frequency-based format was used to capture the extent to which pedagogical practices were enacted in teaching, rather than teachers' subjective perceptions of competence. The instrument adaptation process involved several stages. First, the original questionnaire was translated from English into Indonesian. Second, item wording was contextually modified to align with the characteristics of vocational education and the implementation of the Merdeka Curriculum. Third, the adapted instrument was reviewed by three experts in vocational education and educational measurement to evaluate item relevance, clarity, and representativeness. Fourth, the instrument was pilot-tested with a small group of OMBS teachers to assess clarity and contextual suitability. Feedback from the expert review and pilot testing was used to refine several item labels and wording prior to the main data collection.

Content validity was established through expert judgment focusing on item relevance, clarity, and representativeness. Item validity was subsequently examined using corrected item–total correlations, with all

p-ISSN: 2597-7792 / e-ISSN: 2549-8525
DOI: https://doi.org/10.20961/ijpte.v10i1.110020

items exceeding the acceptable threshold of 0.30. Internal consistency reliability was assessed using Cronbach's alpha, and all dimensions demonstrated coefficients above 0.70, indicating acceptable reliability (Hair et al., 2019). Several indicator labels were also refined to improve conceptual clarity in reporting. For example, terms such as *Scientific* and *Text expression* were reformulated into clearer pedagogical descriptors to better reflect instructional clarity, interaction quality, and digitally supported teaching practices. Inclusion of video quality as one dimension of pedagogical performance was theoretically justified by the increasing role of technology-integrated instruction in contemporary teaching, particularly within the TPACK framework, which emphasizes the integration of pedagogy, content, and technology.

Table 1. presents the variable measurement used in this study.

No	Variable	Indicators	Validity	Cronbach's Alpha
1	*Course content*	*a. Teaching goal*	0.63	0.84
		b. Time mismatch	0.58	
		c. Content preparation	0.66	
		d. Scientific	0.61	
		e. Knowledge	0.64	
2	*Teaching design*	*a. Interest stimulation*	0.69	0.86
		b. Teaching method	0.72	
		c. Example displays	0.65	
		d. Exercise design	0.67	
3	*Video quality*	*a. Clarity*	0.62	0.81
		b. Text expression	0.60	
		c. Teacher teaching Attitude	0.64	
4	*Teaching service*	*a. Communication*	0.68	0.79
		b. Question Answering	0.66	
5	*Teaching evaluation*	*a. Class Practice*	0.70	0.83
		b. Final Exam	0.67	
6	*Learning effect*	*a. Question Answering*	0.61	0.85
		b. 1. Assignment Completion	0.65	
		c. 2. Assignment Completion	0.63	
		d. 3. Assignment Completion	0.66	

Data Preapration

Several data preparation procedures were conducted prior to clustering. First, all variables were standardized using z-scores to ensure comparability across dimensions and to prevent variables with larger numerical ranges from dominating the clustering solution. Second, multivariate outliers were examined using Mahalanobis distance. Observations exceeding the critical threshold were evaluated and excluded when necessary to improve the robustness of the cluster structure. Third, correlation analysis was performed to examine the relationships among variables and to ensure that no variables were excessively redundant. The correlation coefficients indicated moderate associations among dimensions, suggesting that each variable contributed meaningful and non-duplicative information to the clustering procedure.

Data Analysis

Cluster analysis was conducted using the K-means algorithm in SPSS version 24.0 to classify teachers into relatively similar groups based on their pedagogical performance. This approach was selected because it enables the identification of distinct performance profiles by grouping individuals with similar characteristics while maximizing differences across clusters. The optimal number of clusters was determined using the elbow method based on the total within-cluster sum of squares (SSE). Several solutions ranging from k = 2 to k = 5 were evaluated. The elbow point was identified at k = 3, where the reduction in SSE began to level off, indicating that

p-ISSN: 2597-7792 / e-ISSN: 2549-8525
DOI: https://doi.org/10.20961/ijpte.v10i1.110020

additional clusters did not substantially improve the clustering structure. Therefore, a three-cluster solution was retained for the final analysis.

The K-means procedure was implemented using random initial centroid selection with a maximum of 100 iterations. A fixed random seed was applied to improve the reproducibility of the analysis. Convergence was reached when no substantial changes in cluster membership were observed between iterations. The adequacy of the clustering solution was assessed by examining within-cluster similarity and between-cluster separation. Lower within-cluster SSE values indicated a relatively high degree of similarity among teachers within the same cluster, while clear differences between final cluster centers indicated meaningful separation across clusters. Cluster stability was further examined by repeating the K-means procedure with different initial centroid seeds. The consistency of the resulting cluster memberships across repeated runs suggested that the three-cluster solution was sufficiently stable for interpretation.

3. RESULTS

Preliminary Data Screening

Table 2. Correlation Matrix of Research Variables

Variables	1	2	3	4	5	6
1. Course Content	1.000					
2. Teaching Design	0.642	1.000				
3. Video Quality	0.514	0.587	1.000			
4. Teaching Service	0.468	0.532	0.495	1.000		
5. Teaching Evaluation	0.553	0.601	0.472	0.546	1.000	
6. Learning Effect	0.498	0.573	0.521	0.588	0.615	1.000

Preliminary screening indicated that the dataset was sufficiently clean and analytically appropriate for cluster modeling. From the initial 143 responses, 100 complete questionnaires remained after listwise screening, and multivariate outlier detection using Mahalanobis distance identified one extreme case that was excluded, leaving a final sample of 99 teachers for the cluster analysis. This pattern suggests that most responses were suitable for analysis, whereas one case showed an atypical multivariate configuration that could have distorted the distance-based grouping if it had been retained. Sensitivity to extreme observations is a well-recognized issue in multivariate classification, which is why Mahalanobis-based screening remains an important procedure for improving the robustness and interpretability of cluster solutions (Cabana et al., 2021).

Correlation analysis further showed that the six pedagogical dimensions were meaningfully related, yet not redundant, which supports their suitability for clustering. Values reported in Table 2 ranged from 0.468 to 0.642, with the strongest association found between Course Content and Teaching Design (r = 0.642), followed by Teaching Evaluation and Learning Effect (r = 0.615), whereas the weakest relationship appeared between Teaching Service and Course Content (r = 0.468). This result indicates that the variables moved in the same general direction but still represented distinguishable aspects of pedagogical performance, so the data did not collapse into a single undifferentiated construct. Substantively, the matrix suggests that stronger instructional planning tended to accompany stronger instructional design, while evaluation practices were closely connected with perceived learning outcomes. Person-centered and cluster-based studies in education similarly show that meaningful profiles are more likely to emerge when indicators are theoretically connected but still capture different facets of professional practice (Chahuán-Jiménez et al., 2025).

Outlier Handling

Table 3. Mahalanobis Distance Calculation Results

	N	*Minimum*	*Maximum*	*Mean*	*SD*
Mahalanobis Distance	*99*	*1.11963*	*19.65616*	*5.7165356*	*4.06106544*
Valid N (listwise)	*99*				

p-ISSN: 2597-7792 / e-ISSN: 2549-8525
DOI: https://doi.org/10.20961/ijpte.v10i1.110020

Multivariate outlier handling showed that the dataset required only minimal refinement before clustering. One observation exceeded the predefined Mahalanobis distance threshold and was therefore classified as an extreme multivariate case. Its exclusion indicates that the overall data structure was generally well behaved, with only a single response displaying an unusual combination of values across the six pedagogical dimensions. This step is methodologically important because distance-based procedures such as K-means are sensitive to extreme observations, and Mahalanobis-based screening remains a recommended approach for reducing distortion in multivariate classification and improving the robustness of subsequent clustering results (Cabana et al., 2021).

Robustness checking further showed that the general three-cluster pattern remained stable both before and after the removal of the outlier, suggesting that the overall cluster structure was not artificially created by a single extreme case. A more important result, however, was that within-cluster similarity improved after exclusion, which means that the final solution became more internally coherent and easier to interpret. Values reported in Table 3 also indicate that the retained dataset had a Mahalanobis distance range from 1.11963 to 19.65616, with a mean of 5.7165 and a standard deviation of 4.0611, supporting the conclusion that the final analysis based on 99 observations was more stable for profile interpretation. Recent cluster-based educational studies likewise emphasize that interpretable profiles depend not only on statistical separation, but also on the stability and internal consistency of the resulting cluster solution (Chahuán-Jiménez et al., 2025).

Cluster Formation and Final Centers

Table 4. Final Cluster Centers

	Cluster		
	1 (low)	***2 (high)***	***3 (moderate)***
CourseContent	*45*	*57*	*51*
TeachingDesign	*21*	*27*	*24*
VideoQuality	*19*	*25*	*21*
TeachingService	*25*	*31*	*28*
TeachingEvaluation	*28*	*35*	*31*
LearningEffect	*29*	*37*	*31*

K-means clustering was employed to classify teachers based on six dimensions of pedagogical performance, namely Course Content, Teaching Design, Video Quality, Teaching Service, Teaching Evaluation, and Learning Effect. Evaluation of several candidate solutions showed that the decline in the within-cluster sum of squares began to level off at k = 3, indicating that a three-cluster solution provided the most reasonable balance between simplicity and meaningful differentiation. This result suggests that the data were not best represented as either a very small number of overly broad groups or a large number of overly fragmented groups. From an analytical perspective, such a solution is consistent with the primary objective of cluster analysis, which is to identify relatively homogeneous subgroups while preserving meaningful distinctions between them (Gao et al., 2023). The corresponding elbow plot should therefore be presented in the revised manuscript to visually support the selection of the three-cluster model.

Centroid values reported in Table 4 further show that the three-cluster solution was substantively interpretable across all six dimensions. Cluster 2 consistently displayed the highest composite scores, Cluster 1 the lowest, and Cluster 3 intermediate values, which justified labeling the groups as high, low, and moderate pedagogical performance, respectively. This consistent ordering indicates that the clusters were not formed randomly, but reflected a coherent gradient of pedagogical performance across the measured dimensions. Substantively, the data suggest that teachers in the high-performance cluster demonstrated stronger and more integrated pedagogical practices across planning, delivery, support, evaluation, and learning-related outcomes, whereas the low-performance cluster showed weaker performance across the same domains. Recent teacher-related clustering studies likewise show that centroid-based profiles can provide meaningful and interpretable distinctions among professional competence groups when the dimensions display consistent relative ordering across clusters (Chahuán-Jiménez et al., 2025).

p-ISSN: 2597-7792 / e-ISSN: 2549-8525
DOI: https://doi.org/10.20961/ijpte.v10i1.110020

Cluster Membership Distribution

Table 5. Cluster Total

Number of Cases in each Cluster		
Cluster	*1*	*35.000*
	2	*27.000*
	3	*37.000*
Valid		*99.000*
Missing		*0.000*

Cluster membership distribution indicates that teachers were not concentrated in a single dominant performance profile, but were spread across three relatively distinct groups. Values reported in Table 5 show that the moderate-performance cluster constituted the largest group with 37 teachers (37.4%), followed closely by the low-performance cluster with 35 teachers (35.4%), whereas the high-performance cluster comprised only 27 teachers (27.3%). This pattern suggests that most teachers were located in the middle-to-lower range of pedagogical performance, while a smaller proportion demonstrated consistently stronger performance across all six dimensions. Substantively, the data imply that implementation of the Merdeka Curriculum had not yet produced uniformly high pedagogical performance among teachers in the sample, but instead reflected meaningful heterogeneity in readiness and practice. From a person-centered perspective, such membership patterns are important because they show the relative prevalence of each profile within the sample and help identify which groups may require greater developmental support (Gao et al., 2023). Similar cluster-based studies in teacher competence also report that professional groups often distribute across lower, intermediate, and higher profiles rather than forming one homogeneous category, thereby providing a more nuanced basis for interpretation and intervention (Chahuán-Jiménez et al., 2025).

Cluster Differentiation Accros Dimenssion

Table 6. Means ANOVA

	Cluster		***Error***		***F***	***Sig.***
	Mean Square	***df***	***Mean Square***	***df***		
CourseContent	*1166.187*	*2*	*8.482*	*96*	*137.493*	*.000*
TeachingDesign	*247.317*	*2*	*2.844*	*96*	*86.961*	*.000*
VideoQuality	*280.026*	*2*	*5.950*	*96*	*47.060*	*.000*
TeachingService	*339.086*	*2*	*3.422*	*96*	*99.089*	*.000*
TeachingEvaluation	*448.442*	*2*	*4.332*	*96*	*103.525*	*.000*
LearningEffect	*570.191*	*2*	*5.573*	*96*	*102.320*	*.000*

Values reported in Table 6 indicate how strongly each pedagogical dimension contributed to the differentiation of the three cluster profiles. These ANOVA statistics should be interpreted descriptively, not as confirmatory inferential tests, because they were generated after the clusters had already been formed as part of the K-means procedure. In other words, the table is useful for identifying which variables most strongly separated the profiles, but it should not be used to make formal post hoc significance claims about population-level group differences (IBM Corp., 2024). From this perspective, the results provide an internal indicator of cluster distinctiveness rather than an independent test of the model.

Patterns in Table 6 further show that all six variables made meaningful contributions to the three-cluster solution, although the magnitude of their contribution varied considerably. Course Content displayed the largest F value (F = 137.493), indicating that it was the strongest discriminator among the clusters, followed by Teaching Evaluation (F = 103.525), Learning Effect (F = 102.320), and Teaching Service (F = 99.089). Teaching Design (F =

p-ISSN: 2597-7792 / e-ISSN: 2549-8525
DOI: https://doi.org/10.20961/ijpte.v10i1.110020

86.961) and Video Quality (F = 47.060) also contributed to cluster separation, although to a lesser extent. Substantively, this pattern suggests that the three-cluster solution was differentiated across the full set of pedagogical dimensions, with instructional content, assessment, and learning-related outcomes playing the most prominent role in shaping the profiles. Cluster-analytic literature similarly emphasizes that variables with larger between-cluster relative to within-cluster variation tend to be more influential in defining interpretable subgroup structures, while recent profile studies in teacher competence likewise use these differences to explain why some dimensions are more central than others in distinguishing professional groups (Chahuán-Jiménez et al., 2025; Gao et al., 2023).

Cluster Profile by Pedagogical Dimensions
Course Content Dimension

Table 7. Comparative Analysis of Indicators on the Content Course Variable

Variable	Indicators	Items	Group		
			1	2	3
Course content	Teaching Goals (Teaching Objectives)	A	113	106	133
		B	107	102	126
		C	108	102	127
	Time Mismatch (Consistency between Content and Objectives)	A	104	105	127
		B	106	103	123
		C	99	101	117
		D	101	97	113
	Content Preparation	A	112	105	139
		B	103	105	127
		C	95	101	114
	Scientific	A	108	107	132
		B	103	103	124
	Knowledge.	A	99	95	124
		B	107	106	129
		C	105	106	134

The Course Content dimension revealed meaningful variation across the three groups, although the indicator-level pattern should be interpreted with caution. Values in Table 7 show that Group 3 consistently obtained the highest scores across indicators related to Teaching Goals, Time Mismatch, Content Preparation, Scientific, and Knowledge, whereas Group 1 generally showed the lowest values and Group 2 tended to occupy an intermediate position. At the same time, the aggregate centroid pattern in Table 4 identified Cluster 2 as the high-performance cluster, Cluster 1 as the low-performance cluster, and Cluster 3 as the moderate-performance cluster. This discrepancy suggests that the item-level totals in Table 7 should be treated primarily as descriptive evidence of within-dimension variation rather than as a basis for redefining the overall cluster labels. In cluster analysis, profile interpretation should remain anchored in the overall centroid structure, because cluster membership is determined by the combined multivariate pattern rather than by isolated indicator counts within a single dimension (Gao et al., 2023).

Substantively, the pattern in Table 7 indicates that differences in pedagogical performance were closely associated with the quality of instructional preparation, coherence between content and learning objectives, and adequacy of subject-related knowledge. These aspects are central to effective teaching because stronger alignment between objectives, learning content, and instructional design supports clearer pedagogical direction and more coherent learning experiences (Pérez-Álvarez et al., 2020). Relevance of this dimension becomes even

p-ISSN: 2597-7792 / e-ISSN: 2549-8525
DOI: https://doi.org/10.20961/ijpte.v10i1.110020

stronger in the context of the Merdeka Curriculum, where teachers are expected to adapt content flexibly while still maintaining coherence among objectives, materials, and students' learning needs (Cholilah et al., 2023). Taken together, the results suggest that the course content dimension functioned as an important marker of pedagogical differentiation, even though the item-level group ordering requires careful interpretation against the broader cluster solution.

Teaching Design Dimension

Table 8. Comparison of Indicators on the Teaching Design Variable

Variable	Indicators	Items	Group		
			1	**2**	**3**
Teaching Design	Interest Stimulation	A	107	103	128
		B	102	101	115
	Teaching Method	A	92	99	106
		B	103	103	122
		C	108	104	122
	Example Display	A	111	106	135
		B	113	107	137
	Exercise Design	A	108	104	131
		B	107	103	132

The Teaching Design dimension also revealed meaningful variation across the three groups. Values in Table 8 show that Group 3 generally obtained the highest scores on Interest Stimulation, Example Display, and Exercise Design, whereas the Teaching Method indicators displayed a more uneven pattern across groups. This result suggests that teachers were relatively more consistent in presenting examples and designing exercises than in varying instructional methods. Interpretation of the item-level pattern should nevertheless remain cautious, because the broader centroid structure in Table 4 identifies Cluster 2 as the high-performance group, Cluster 1 as the low-performance group, and Cluster 3 as the moderate-performance group. For that reason, the distribution in Table 8 is better read as descriptive evidence of variation within the teaching design dimension rather than as a basis for redefining the overall cluster labels. Cluster interpretation should remain anchored in the multivariate centroid pattern, since cluster membership is determined by the combined contribution of all dimensions rather than by a single set of item frequencies (Gao et al., 2023).

Substantively, the pattern in Table 8 indicates that stronger pedagogical performance was associated with more developed instructional design, especially in terms of stimulating student interest, selecting appropriate teaching methods, and structuring meaningful learning activities. Importance of this dimension is consistent with recent literature emphasizing that effective instructional design depends on coherence between learning goals, activities, and pedagogical strategies, as well as the teacher's ability to organize learning experiences that actively engage students (Pérez-Álvarez et al., 2020). Relevance of these findings is particularly strong in the context of the Merdeka Curriculum, which expects teachers to design flexible, student-centered learning that responds to learners' needs while maintaining pedagogical coherence (Cholilah et al., 2023). Taken together, the results suggest that differences in pedagogical performance extended beyond content preparation and were also reflected in how teachers designed and structured learning experiences in practice.

Video Quality Dimension

The Video Quality dimension also contributed meaningfully to cluster differentiation. Values reported in Table 9 show variation across the indicators of Clarity, Text Expression, and Teachers' Teaching Attitude, with Group 3 generally obtaining higher item-level scores and Group 1 showing lower scores across most items. This pattern indicates that differences among teachers were not limited to content planning or instructional design,

p-ISSN: 2597-7792 / e-ISSN: 2549-8525
DOI: https://doi.org/10.20961/ijpte.v10i1.110020

but also extended to the quality of their digitally mediated instructional presentation. Interpretation of the item-level results should nevertheless remain cautious, because the broader centroid structure in Table 4 still identifies Cluster 2 as the high-performance group, Cluster 1 as the low-performance group, and Cluster 3 as the moderate-performance group. For that reason, the distribution in Table 9 is better understood as descriptive evidence of variation within the video quality dimension rather than as a basis for changing the overall cluster labels.

Substantively, the pattern in Table 9 suggests that stronger pedagogical performance was associated with clearer presentation, better textual support, and a stronger quality of instructional delivery in video-based learning materials. Relevance of this finding is consistent with recent literature showing that effective technology-integrated teaching depends not only on access to digital tools, but also on the teacher's ability to combine pedagogical and technological knowledge in ways that enhance clarity, interaction, and learning quality (Çebi et al., 2022; Jiménez Sierra et al., 2023). Seen from this perspective, the video quality dimension should not be reduced to a purely technical feature. A more appropriate interpretation is that digitally mediated instructional quality formed one meaningful component of broader pedagogical performance under the Merdeka Curriculum, where teachers are increasingly expected to present learning in flexible and pedagogically coherent ways.

Table 9. Comparison Analysis of Indicators in the Video Quality Variable

Variable	Indicators	Items	Group		
			1	2	3
Video Quality	Clarity	A	87	88	101
		B	92	92	109
		C	95	99	112
	Text expression	A	97	98	110
		B	96	99	113
	Teachers' teaching attitude	A	99	98	113
		B	100	101	111

Teaching Service Dimension

Table 10. Comparison Analysis of Indicators on the Teaching Service Variable

Variable	Indicators	Items	Group		
			1	2	3
Teaching Service	Communication	A	111	108	140
		B	109	107	130
		C	105	105	126
		D	107	104	129
		E	112	108	132
	Question answering	A	106	107	128
		B	103	103	119
		C	111	104	133

The Teaching Service dimension showed comparatively strong performance across the three groups, particularly on the Communication indicators reported in Table 10. Scores for Question Answering were also relatively positive, although several items remained slightly lower than those for general communication. Item-level values again indicate that Group 3 tended to obtain higher scores than the other groups, while Group 1

p-ISSN: 2597-7792 / e-ISSN: 2549-8525
DOI: https://doi.org/10.20961/ijpte.v10i1.110020

generally showed lower values. Interpretation of this pattern should nonetheless remain cautious, because the broader centroid structure in Table 4 still identifies Cluster 2 as the high-performance group, Cluster 1 as the low-performance group, and Cluster 3 as the moderate-performance group. For that reason, the distribution in Table 10 is better understood as descriptive evidence of variation within the teaching service dimension rather than as a basis for redefining the overall cluster labels. Overall, the results suggest that teacher–student interaction was a relatively strong aspect of pedagogical practice, while responsiveness to students' questions still appeared less consistent in some groups.

Substantively, the pattern in Table 10 indicates that pedagogical performance under the Merdeka Curriculum involved not only instructional planning and delivery, but also the quality of interpersonal support provided during learning. Relevance of this dimension is supported by recent literature showing that teachers' interpersonal communication skills are closely associated with students' perceived teacher support and classroom engagement, both of which are important for effective learning environments (Le et al., 2022). Importance of communicative and supportive interaction is also consistent with the student-centered orientation of the Merdeka Curriculum, which expects teachers to create responsive and meaningful learning experiences rather than rely only on one-way instruction (Cholilah et al., 2023). Taken together, the results suggest that stronger pedagogical performance was reflected not only in what teachers taught, but also in how they communicated, responded, and supported students during the learning process.

Teaching Evaluation Dimension

Table 11. Comparative Analysis of Indicators for the Teaching Evaluation Variable

Variable	Indicators	Items	Group		
			1	2	3
Teaching Evaluation	Class practice	A	106	105	123
		B	99	98	116
		C	108	108	130
		D	110	107	125
		E	109	107	125
		F	108	106	131
	Final exam.	A	109	106	136
		B	107	106	126
		C	107	107	129

The Teaching Evaluation dimension displayed a clear descriptive pattern across the three groups. Values reported in Table 11 show that the indicators under Class Practice and Final Exam tended to be higher in Group 3 and lower in Group 1, suggesting that evaluative practices were implemented more consistently in some groups than in others. Interpretation of this pattern should nevertheless remain cautious, because the broader centroid structure in Table 4 identifies Cluster 2 as the high-performance group, Cluster 1 as the low-performance group, and Cluster 3 as the moderate-performance group. For that reason, the item-level distribution in Table 11 is better treated as descriptive evidence of variation within the teaching evaluation dimension rather than as a basis for relabeling the overall clusters. A more defensible interpretation is that assessment-related practices contributed meaningfully to differentiating the three pedagogical profiles, even though the item-level ordering does not fully mirror the aggregate cluster labels.

Substantively, the pattern in Table 11 indicates that evaluation practices formed an important part of the cluster structure, particularly through differences in the consistency of classroom practice and final-exam-related assessment. Relevance of this finding is consistent with recent literature emphasizing that effective evaluation depends on coherence between instructional objectives, learning activities, and assessment practices, because stronger alignment allows teachers to monitor learning more meaningfully and support instructional

p-ISSN: 2597-7792 / e-ISSN: 2549-8525
DOI: https://doi.org/10.20961/ijpte.v10i1.110020

decision-making (Pérez-Álvarez et al., 2020). Importance of this dimension is also compatible with recent work showing that teachers' assessment practices shape the quality of classroom teaching and students' learning experiences, especially when assessment is implemented as an integrated part of instruction rather than as a separate administrative task (Bahati et al., 2024). In the context of the Merdeka Curriculum, these results suggest that stronger-performing teachers were more able to connect teaching and evaluation in a coherent way, which may reflect better alignment between instruction and assessment in classroom practice (Cholilah et al., 2023).

Learning Effect Dimension

The Learning Effect dimension also differentiated the three cluster profiles, although the item-level pattern should be interpreted with caution. Values reported in Table 12 show variation across indicators related to the rationality of question frequency, homework participation, homework completion, and passing the exam, with Group 3 generally displaying higher item-level scores and Group 1 showing lower values across most indicators. This pattern suggests that perceptions of learning-related outcomes varied in line with broader differences in pedagogical performance. A cautious reading is still necessary, however, because the broader centroid structure in Table 4 identifies Cluster 2 as the high-performance group, Cluster 1 as the low-performance group, and Cluster 3 as the moderate-performance group. For that reason, the distribution in Table 12 is better treated as descriptive evidence of variation within the learning effect dimension rather than as a basis for changing the overall cluster labels. A more defensible interpretation is that the three clusters differed not only in planning and instructional delivery, but also in the learning-related effects associated with pedagogical practice.

Substantively, the pattern in Table 12 indicates that stronger pedagogical performance was associated with more favorable scores in student questioning, homework participation, homework completion, and exam-related outcomes. Interpretation of this dimension should remain closely tied to the reported indicators themselves, rather than extended to broader constructs not explicitly measured in the table. Relevance of these findings is supported by recent literature showing that homework engagement and completion are meaningfully related to students' academic achievement and learning development, particularly when learning activities are organized consistently and supported by teachers (Valle et al., 2024). Related evidence also shows that teacher support and classroom interaction are closely linked to student engagement, which helps explain why learning-related outcomes may vary across pedagogical profiles (Le et al., 2022). Taken together, the results suggest that the cluster solution captured differences not only in teaching processes, but also in the learning effects associated with those pedagogical practices.

Table 12. Comparison Analysis of Indicators for the Learning Effect Variable

Variable	*Indicators*	*Items*	*Group*		
			1	*2*	*3*
Learning Effect	*The rationality of the frequency of questions*	*A*	*104*	*102*	*118*
		B	*105*	*101*	*116*
	The level of homework participation	*A*	*93*	*95*	*103*
		B	*101*	*100*	*115*
		C	*99*	*102*	*111*
	The level of homework completion	*A*	*96*	*97*	*110*
		B	*101*	*98*	*118*
		C	*107*	*103*	*116*
	Moreover, passing the exam.	*A*	*95*	*100*	*112*
		B	*107*	*106*	*118*

p-ISSN: 2597-7792 / e-ISSN: 2549-8525
DOI: https://doi.org/10.20961/ijpte.v10i1.110020

4. DISCUSSION

Teacher Readiness in Implementing the Merdeka Curriculum

The findings reveal variation in teachers' pedagogical performance in implementing the Merdeka Curriculum, as reflected in the distribution of clusters: low (35.4%), moderate (37.4%), and high (27.3%). This distribution indicates that teacher readiness was not uniform across the sample and suggests that implementation of the Merdeka Curriculum occurred through different levels of pedagogical performance.

Previous studies have similarly shown that curriculum reform is often enacted unevenly across educational settings and teacher groups (Darling-Hammond et al., 2017; König et al., 2021). The present findings support that general observation by showing that vocational teachers in this study did not form a single pedagogical profile, but instead were distributed across three distinguishable performance groups. Related evidence also suggests that teachers may encounter difficulties when adapting instructional practices during periods of educational change (Yulin & Danso, 2025). In the present study, such variation appears in the cluster structure rather than in a uniform pattern of readiness. The findings may also be interpreted conceptually through the TPACK framework, which emphasizes the interrelationship among content, pedagogy, and technology in teaching (Koehler et al., 2013). The cluster differences observed in this study suggest that pedagogical performance under the Merdeka Curriculum involved multiple dimensions rather than a single teaching capability. Recent studies also indicate that teachers often encounter challenges in integrating technology into pedagogically meaningful practice (Jiménez Sierra et al., 2023). For that reason, teacher readiness in this context is more appropriately understood as a multidimensional construct.

Course Content Dimension and Pedagogical Content Knowledge

The analysis of the course content dimension indicates that differences across clusters were closely associated with instructional goal formulation, alignment between content and objectives, content preparation, and knowledge-related indicators. These aspects are particularly important in the Merdeka Curriculum, where teachers are expected to organize learning content flexibly while still maintaining pedagogical coherence and relevance to students' needs. Although the item-level pattern in Table 7 should be interpreted cautiously, the overall centroid structure still shows that stronger-performing clusters achieved better scores on course content than lower-performing clusters. This pattern suggests that variation in pedagogical performance was not merely a matter of delivery style, but was also rooted in how effectively teachers prepared, structured, and aligned instructional content before classroom implementation.

Conceptually, this finding can be understood through the lens of pedagogical content knowledge, which emphasizes the integration of subject matter and pedagogy in effective teaching. Previous studies have shown that stronger pedagogical content knowledge is associated with better instructional design and higher teaching quality (Hattie & Donoghue, 2016). Present results similarly suggest that teachers differed in the extent to which they were able to prepare and organize content in pedagogically meaningful ways, especially in relation to coherence between objectives, materials, and knowledge representation. Related evidence has also shown that teachers may experience difficulty when adapting subject-related knowledge to changing educational contexts and emerging instructional demands (Chiu, 2021; Li et al., 2025). These findings therefore imply that performance differences in the course content dimension reflect not only differences in planning quality, but also differences in teachers' capacity to translate subject knowledge into instruction that is coherent, adaptive, and educationally relevant.

Digital Pedagogy and the Role of Video Quality

The findings on video quality indicate that digitally mediated instructional presentation formed a meaningful component of pedagogical performance in this study. Differences across the clusters were visible in clarity, text presentation, and teaching attitude, suggesting that teachers varied in how effectively they designed and delivered video-based learning materials. Performance in this dimension was weaker in the low-performance cluster and stronger and more consistent in the high-performance cluster, which implies that digital presentation quality was not merely an auxiliary feature of teaching, but one of the dimensions through which broader pedagogical differences became visible. This pattern suggests that teachers who demonstrated stronger overall pedagogical performance were also more capable of producing instructional materials that were clearer, more accessible, and more supportive of student learning. Conceptually, this pattern is consistent with literature

p-ISSN: 2597-7792 / e-ISSN: 2549-8525
DOI: https://doi.org/10.20961/ijpte.v10i1.110020

emphasizing the importance of digital competence and technology-integrated pedagogy in effective instruction (Çebi et al., 2022). Seen through the TPACK perspective, the differences across clusters may be interpreted as reflecting variation in how teachers combined pedagogical and technological knowledge in practice (Koehler et al., 2013). Present findings therefore suggest that digital pedagogy should be understood as one component of broader pedagogical performance rather than as a purely technical attribute, because the quality of video-based instruction depends not only on technical presentation, but also on the teacher's ability to integrate clarity, pedagogy, and communication into meaningful learning experiences.

Implications

The cluster structure identified in this study provides an empirical basis for differentiated professional development strategies. Since the findings show heterogeneous levels of pedagogical performance across six dimensions, a uniform approach to teacher development may be less effective than a profile-based approach. Teachers in the low-performance cluster may require more foundational support in instructional preparation, teaching design, digital pedagogy, and evaluation practice. Teachers in the moderate-performance cluster may benefit from programs that strengthen the integration of content-related preparation, adaptive instructional design, and digitally supported teaching. Teachers in the high-performance cluster, who demonstrated stronger performance across dimensions, may represent a potential resource for collaborative learning within schools, although this implication should be interpreted cautiously because mentoring roles were not directly examined in the study. This interpretation is broadly aligned with the argument that effective professional development should be context-specific, sustained, and responsive to teachers' existing competencies rather than based on a one-size-fits-all model (Desimone & Garet, 2015). Practice-based and collaborative approaches have also been discussed as promising ways to strengthen technology, pedagogy, and content integration in teaching (Scherer et al., 2023; Tondeur et al., 2017; Geletu, 2026; Wang et al., 2026)

Limitations and Future Research

This study has several limitations that should be considered when interpreting the findings. First, the data were collected using a voluntary online survey distributed through WhatsApp, which may introduce sampling bias. Participants who chose to respond may differ systematically from those who did not participate. Similar limitations have been noted in prior studies using self-selected survey samples in educational research (Hiratsuka, 2025). Second, the study relies on self-reported measures of pedagogical performance. Although self-report instruments are widely used in educational research, they may be subject to social desirability bias and individual perception differences. Previous research suggests that combining self-reported data with observational or performance-based measures can improve the validity of findings (Baldwin et al., 2017; Kislaya et al., 2021). Third, the use of cluster analysis provides a useful framework for identifying patterns of pedagogical performance; however, the results are sensitive to the selected variables and clustering parameters. As noted in recent methodological studies, cluster solutions should be interpreted as exploratory rather than definitive representations of population structure (Sarstedt et al., 2019).

Finally, this study focuses on teachers within a specific vocational education context in Central Java, which may limit the generalizability of the findings to other educational settings. Future research is recommended to validate these cluster patterns in different regions and educational levels, as well as to incorporate longitudinal designs to examine how teacher performance evolves over time (Scherer et al., 2023; Tondeur et al., 2017). Despite these limitations, the study provides a structured empirical contribution by demonstrating that teacher pedagogical performance is not homogeneous but varies across distinct clusters, each with specific developmental needs.

5. CONCLUSION

This study contributes to understanding teacher readiness for implementing the Merdeka Curriculum by identifying distinct profiles of pedagogical performance through a cluster-based approach. The findings reveal three pedagogical performance clusters among vocational teachers: low-performance, moderate-performance, and high-performance. This pattern indicates that teachers' pedagogical performance varies across the sample and that readiness for curriculum implementation is not uniform. The results further show that the identified clusters differ across multiple dimensions, including course content, teaching design, video quality, teaching

p-ISSN: 2597-7792 / e-ISSN: 2549-8525
DOI: https://doi.org/10.20961/ijpte.v10i1.110020

service, teaching evaluation, and learning effect. These differences suggest that pedagogical performance in the context of the Merdeka Curriculum is multidimensional and reflects variation in how teachers plan, implement, support, and evaluate learning. In particular, the low-performance cluster was characterized by comparatively weaker performance across dimensions, the moderate-performance cluster showed a balanced but less consistent profile, and the high-performance cluster demonstrated stronger and more integrated pedagogical practices.

From a methodological perspective, the study shows that a cluster-based approach can provide a differentiated picture of pedagogical performance by identifying groups of teachers with relatively similar performance profiles. This perspective extends existing research on teacher performance, particularly in vocational education, by moving beyond average tendencies and highlighting meaningful variation among teachers. From a practical perspective, the findings suggest that professional development may be more effective when aligned with the characteristics of different teacher profiles. Teachers in the low-performance cluster may require more foundational support in instructional design, digital pedagogy, and evaluation practices, whereas teachers in the moderate cluster may benefit from further support in strengthening adaptive and integrated teaching practices. Teachers in the high-performance cluster may also represent a valuable resource for collaborative professional learning within schools. Several limitations should be acknowledged. The use of voluntary sampling through an online WhatsApp-based survey may have introduced self-selection bias, and the focus on one regional vocational context may limit the broader transferability of the findings. In addition, the study relied on self-reported data, which may be influenced by response bias. Future research is therefore recommended to examine pedagogical performance using more diverse samples, complementary data sources, and broader methodological approaches, particularly in relation to digital pedagogy and technology-integrated teaching. Overall, the study suggests that teacher readiness for curriculum implementation may be meaningfully interpreted through distinct pedagogical performance profiles, highlighting the importance of more targeted and context-responsive professional development strategies.

6. ACKNOWLEDGMENTS

This research was funded by the DPA FEB Universitas Negeri Semarang under Budget Number DPA 023.17.2.690645/2024.07, based on the Research Assignment Agreement for the Implementation of Research Funded by DPA FEB UNNES 2024, Number 55.1.4/UN37/PPK.07/2024, dated 1 April 2024. The authors gratefully acknowledge the financial support provided by Faculty of Economics and Business, Universitas Negeri Semarang. This work is original and has not received funding from any other institution or funding source.

7. REFERENCES

Bahati, B., Alam, G. M., & Said, J. (2024). A systematic review on the factors affecting classroom formative assessment practices. Asia Pacific Education Review, 25, 161–176. https://doi.org/10.1007/s12564-023-09881-0

Baldwin, J. N., McKay, M. J., Hiller, C. E., Moloney, N., Nightingale, E. J., & Burns, J. (2017). Relationship between physical performance and self-reported function in healthy individuals across the lifespan. Musculoskeletal Science and Practice, 30. https://doi.org/10.1016/j.msksp.2017.05.001

Baumert, J., Kunter, M., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (2013). Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project. In Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project. https://doi.org/10.1007/978-1-4614-5149-5

Burroughs, N., Gardner, J., Lee, Y., Guo, S., Touitou, I., Jansen, K., & Schmidt, W. (2019). A review of the literature on teacher effectiveness and student outcomes. In IEA Research for Education (Vol. 6). https://doi.org/10.1007/978-3-030-16151-4_2

Cabana, E., Lillo, R. E., & Laniado, H. (2021). Multivariate outlier detection based on a robust Mahalanobis distance with shrinkage estimators. Statistical Papers, 62, 1583–1609. https://doi.org/10.1007/s00362-019-01148-1

p-ISSN: 2597-7792 / e-ISSN: 2549-8525
DOI: https://doi.org/10.20961/ijpte.v10i1.110020

Çebi, A., Bahçekapılı Özdemir, T., Reisoğlu, İ., & Çolak, C. (2022). From digital competences to technology integration: Re-formation of pre-service teachers' knowledge and understanding. International Journal of Educational Research, 113, 101965. https://doi.org/10.1016/j.ijer.2022.101965

Chahuán-Jiménez, K., Lara-Yergues, E., Garrido-Araya, D., Salum-Alvarado, E., Hurtado-Arenas, P., & Rubilar-Torrealba, R. (2025). Cluster analysis of digital competencies among professors in higher education. Frontiers in Education, 10. https://doi.org/10.3389/feduc.2025.1499856

Chiu, T. K. F. (2021). A holistic approach to the design of artificial intelligence (AI) education for K–12 schools. TechTrends, 65(5), 796–807. https://doi.org/10.1007/s11528-021-00637-1

Cholilah, M., Tatuwo, A. G. P., Rosdiana, S. P., & Fatirul, A. N. (2023). Pengembangan kurikulum merdeka dalam satuan pendidikan serta implementasi kurikulum merdeka pada pembelajaran abad 21. Sanskara Pendidikan dan Pengajaran, 1(2), 57–66. https://doi.org/10.58812/spp.v1i02.110

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). Effective teacher professional development (Research brief). https://doi.org/10.54300/122.311

Desimone, L. M., & Garet, M. S. (2015). Best practices in teachers' professional development in the United States. Psychology, Society & Education, 7(3). https://doi.org/10.25115/psye.v7i3.516

Fauth, B., Decristan, J., Decker, A.-T., Büttner, G., Hardy, I., Klieme, E., & Kunter, M. (2019). The effects of teacher competence on student outcomes in elementary science education: The mediating role of teaching quality. Teaching and Teacher Education, 86, 102882. https://doi.org/10.1016/j.tate.2019.102882

Gao, C. X., Dwyer, D. E., Zhu, Y., Smith, C. L., Du, L., Filia, K. M., Bayer, J., Menssink, J. M., Wang, T., Bergmeir, C., Wood, S. J., & Cotton, S. M. (2023). An overview of clustering methods with guidelines for application in mental health research. Psychiatry Research, 327, 115265. https://doi.org/10.1016/j.psychres.2023.115265

Geletu, G. M. (2026). The effects of pedagogical mentoring and coaching on primary school teachers' professional development practices and students' learning engagements in classrooms in Oromia regional state: Implications for professionalism. Education 3–13, 54(1). https://doi.org/10.1080/03004279.2023.2293209

Gessler, M., & Siemer, C. (2020). Umbrella review: Methodological review of reviews published in peer-reviewed journals with a substantial focus on vocational education and training research. International Journal for Research in Vocational Education and Training, 7(1). https://doi.org/10.13152/IJRVET.7.1.5

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. European Business Review, 31(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Hattie, J. A. C., & Donoghue, G. M. (2016). Learning strategies: A synthesis and conceptual model. NPJ Science of Learning, 1(1). https://doi.org/10.1038/npjscilearn.2016.13

Hiratsuka, T. (2025). The volunteer participation paradox: Ethical tensions between self-selection and targeted sampling. Research Methods in Applied Linguistics, 4(2). https://doi.org/10.1016/j.rmal.2025.100206

Hua, S., & Ren, Z. (2020). "Online + Offline" course teaching based on case teaching method: A case study of entrepreneurship education course. International Journal of Emerging Technologies in Learning, 15(10), 69–85. https://doi.org/10.3991/ijet.v15i10.13999

Jiménez Sierra, Á. A., Ortega Iglesias, J. M., Cabero-Almenara, J., & Palacios-Rodríguez, A. (2023). Development of the teacher's technological pedagogical content knowledge (TPACK) from the Lesson Study: A systematic review. Frontiers in Education, 8, 1078913. https://doi.org/10.3389/feduc.2023.1078913

Kislaya, I., Leite, A., Perelman, J., Machado, A., Torres, A. R., Tolonen, H., & Nunes, B. (2021). Combining self-reported and objectively measured survey data to improve hypertension prevalence estimates: Portuguese experience. Archives of Public Health, 79(1). https://doi.org/10.1186/s13690-021-00562-y

p-ISSN: 2597-7792 / e-ISSN: 2549-8525
DOI: https://doi.org/10.20961/ijpte.v10i1.110020

Koehler, M. J., Mishra, P., & Cain, W. (2013). What is technological pedagogical content knowledge (TPACK)? Journal of Education, 193(3). https://doi.org/10.1177/002205741319300303

König, J., Blömeke, S., Jentsch, A., Schlesinger, L., née Nehls, C. F., Musekamp, F., & Kaiser, G. (2021). The links between pedagogical competence, instructional quality, and mathematics achievement in the lower secondary classroom. Educational Studies in Mathematics, 107(1). https://doi.org/10.1007/s10649-020-10021-0

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. Journal of Educational Psychology, 105(3). https://doi.org/10.1037/a0032583

Le, M. T., Janssen, J., Wubbels, T., & Van Tartwijk, J. (2022). Teachers' interpersonal communication skills and its relationship with students' experiences of teacher support and student engagement. Learning and Instruction, 80, 101619. https://doi.org/10.1016/j.learninstruc.2021.101619

Li, Y., Zhou, X., Yin, H. B., & Chiu, T. K. F. (2025). Design language learning with artificial intelligence (AI) chatbots based on activity theory from a systematic review. Smart Learning Environments, 12(1). https://doi.org/10.1186/s40561-025-00379-0

Pearson, S. (2013). Professional capital: Transforming teaching in every school. Teacher Development, 17(1). https://doi.org/10.1080/13664530.2013.766491

Pérez-Álvarez, R., Breckler, J., Barger, M., et al. (2020). Constructive alignment in first-year biology and maths courses at a large-enrollment Hispanic-serving institution. Journal of Biological Education, 57(2), 326–337. https://doi.org/10.1080/00219266.2020.1791156

Praetorius, A. K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of three basic dimensions. ZDM Mathematics Education, 50(3). https://doi.org/10.1007/s11858-018-0918-4

Priestley, M., Philippou, S., Alvunger, D., & Soini, T. (2021). Curriculum making: A conceptual framing. In Curriculum making in Europe: Policy and practice within and across diverse contexts. https://doi.org/10.1108/978-1-83867-735-020211002

Puad, L. M. A. Z., & Ashton, K. (2023). A critical analysis of Indonesia's 2013 national curriculum: Tensions between global and local concerns. Curriculum Journal, 34(3). https://doi.org/10.1002/curj.194

Rothwell, A., Domagała-Zyśk, E., Mamcarz, P., Parra, A., Sánchez-Queija, M. I., De Rosa, A., & Lo Presti, A. (2025). Employability: An umbrella review. Australian Journal of Career Development, 34(3). https://doi.org/10.1177/10384162251366428

Sarstedt, M., Hair, J. F., Cheah, J. H., Becker, J. M., & Ringle, C. M. (2019). How to specify, estimate, and validate higher-order constructs in PLS-SEM. Australasian Marketing Journal, 27(3). https://doi.org/10.1016/j.ausmj.2019.05.003

Scherer, R., Siddiq, F., Howard, S. K., & Tondeur, J. (2023). The more experienced, the better prepared? New evidence on the relation between teachers' experience and their readiness for online teaching and learning. Computers in Human Behavior, 139. https://doi.org/10.1016/j.chb.2022.107530

Tondeur, J., van Braak, J., Ertmer, P. A., & Ottenbreit-Leftwich, A. (2017). Understanding the relationship between teachers' pedagogical beliefs and technology use in education: A systematic review of qualitative evidence. Educational Technology Research and Development, 65(3). https://doi.org/10.1007/s11423-016-9481-2

Valle, A., Regueiro, B., Piñeiro, I., Rodríguez, S., Ferradás, M. M., Freire, C., Núñez, J. C., & Rosário, P. (2024). Homework behavior and academic achievement throughout compulsory education. Learning and Individual Differences, 129, 102859. https://doi.org/10.1016/j.lindif.2024.102859

p-ISSN: 2597-7792 / e-ISSN: 2549-8525
DOI: https://doi.org/10.20961/ijpte.v10i1.110020

Voogt, J., Pieters, J., & Roblin, N. P. (2019). Collaborative curriculum design in teacher teams: Foundations. In Collaborative curriculum design for sustainable innovation and teacher learning. https://doi.org/10.1007/978-3-030-20062-6_1

Wang, X., Husu, J., & Toom, A. (2026). Teacher educators' mentoring competence—commonalities and distinctions in supporting learning to teach. Teaching and Teacher Education, 172. https://doi.org/10.1016/j.tate.2026.105408

p-ISSN: 2597-7792 / e-ISSN: 2549-8525
DOI: https://doi.org/10.20961/ijpte.v10i1.110020