JURNAL VARIDIKA
Vol. 38, No. 3, 2026, pp.302-332
p-ISSN 0852-0976 | e-ISSN 2460-3953
Website: https://journals2.ums.ac.id/index.php/varidika/index

Educational versus Non-Educational Backgrounds: A Nationwide Assessment of Physics Teacher Competencies in Indonesia

Dian Artha Kusumaningtyas[1]*, Fahrudin Nugroho[2], Muhammad Syahriandi Adhantoro[3], Efi Kurniasari[4]
[1]Department of Physics Education, Universitas Ahmad Dahlan, Indonesia
[2]Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada, Indonesia
[3]Department of Data Science, Universitas Muhammadiyah Surakarta, Indonesia
[4]Department of Elementary School Education, Universitas Muhammadiyah Sukabumi, Indonesia
*Corresponding author's email: dian.artha@pfis.uad.ac.id

Submission Track:
Received:
27 July 2026
Final Revision:
3 September 2026
Available online:
9 September 2026

ABSTRACT

The caliber of teacher training in Indonesia poses a significant obstacle, especially within the Teacher Professional Education Program (TPEP), which enrolls individuals from both educational and non-educational fields. Although common beliefs imply that graduates from educational disciplines exhibit enhanced pedagogical preparedness, comprehensive national comparative data is markedly limited. This ex post facto quantitative investigation seeks to fill this void by assessing the professional abilities of 728 physics TPEP graduates from 31 institutions. The study specifically contrasts performance based on academic backgrounds, explores institutional disparities, identifies determinants of Pedagogical Content Knowledge (PCK), and delineates graduate competencies. Statistical evaluations—including ANOVA, ANCOVA, and multiple regression—indicated that undergraduate background does not significantly influence professional competence following TPEP completion. Graduates from educational and non-educational tracks showed no statistically significant differences in PCK, UKIN (Portfolio), or UTUL (Written Test) scores. Instead, institutional context emerged as a notable source of variation. Regression modeling demonstrated that UTUL scores and recency of graduation are robust predictors of PCK, whereas undergraduate background is negligible. Interestingly, UKIN scores exhibited a negative association with PCK. Furthermore, cluster analysis segmented graduates into distinct low, moderate, and high-performance profiles, highlighting significant heterogeneity among the cohort. Ultimately, these findings indicate that a teacher's professional competence is driven by general academic preparedness, institutional environment, and study recency, rather than their initial undergraduate discipline. By providing a nationwide mapping of TPEP outcomes, this research challenges traditional assumptions regarding teacher preparation and provides a foundation for targeted, differentiated professional development strategies.

Keywords: Teacher Professional Education, Pedagogical Content Knowledge, Physics Teacher Education, Institutional Mapping, Cluster Analysis.

DOI: 10.23917/varidika.v38i3.19431

INTRODUCTION

Teacher quality is widely recognized as a key determinant of student achievement and overall educational improvement. In Indonesia, the Teacher Professional Education Program (TPEP) has been established as the primary pathway for preparing and certifying teachers across disciplines (C. F. Chang et al., 2024; Mardhiah et al., 2023). Mandated by national law, the program ensures that teachers acquire four essential competencies: pedagogical, professional, personal, and social (Basikin, 2024; Sari et al., 2023). While TPEP has significantly contributed to strengthening teacher quality, it continues to face challenges related to implementation, curriculum alignment, and the diverse academic backgrounds of its participants.

A distinctive feature of the program is its dual intake system, which accommodates both graduates from educational faculties and those from non-educational backgrounds (Agustiani, 2024). This policy responds to the urgent demand for qualified teachers, particularly in physics education where shortages remain acute (Simanjorang et al., 2020; Yuniawatika et al., 2022; Ishartono et al., 2026). Physics is a particularly instructive case for examining this policy because it combines a persistent national teacher shortage with content that is widely regarded as conceptually demanding, making the question of whether non-educational-background graduates can be brought to comparable pedagogical competence especially consequential for policy. However, it also raises critical questions regarding comparative readiness and competence. Non-educational graduates often encounter greater challenges in classroom practice, particularly in pedagogy and instructional strategies (Julia et al., 2020). Their teaching relies more on transferable skills such as communication, interpersonal abilities, and leadership, yet they frequently lack systematic training in pedagogy and curriculum design (Evens et al., 2018). Consequently, these graduates require supportive environments and structured mentorship to develop the competencies necessary for effective teaching. It should be noted, however, that non-educational-background graduates constitute a small minority of TPEP physics enrollees nationally, a numerical imbalance that is examined further in the Method and Discussion sections below.

In contrast, educational graduates typically demonstrate stronger initial readiness due to structured exposure to pedagogy, curriculum development, and classroom management during undergraduate studies. Their preparation not only provides a solid foundation in subject content but also equips them with teaching methodologies that enhance confidence and effectiveness (Julia et al., 2020; Simanjorang et al., 2020). Nonetheless, the assumption that educational graduates consistently outperform their non-educational counterparts has not been consistently confirmed by prior empirical comparisons, and the small number of directly comparative studies makes it difficult to draw firm conclusions either way. Existing studies on teacher professional education in Indonesia have largely focused on program design, implementation challenges, and certification outcomes (Abidin et al., 2023;

Loughran, 2020), leaving a gap in understanding how differences in academic background influence TPEP performance. Where prior work has examined related populations, it has typically addressed either pre-service instructional decision-making or discipline-specific content-knowledge development rather than background-based comparisons within a certification program (Estapa & Davis, 2023; Schiering et al., 2023), which is the specific gap the present study addresses. Furthermore, limited attention has been given to institutional variation and other predictors such as academic test performance or recency of graduation that may shape Pedagogical Content Knowledge (PCK), a construct that has itself been extensively theorized but less often examined empirically in relation to these structural and temporal factors (Abidin et al., 2023; Loughran, 2020).

Theoretically, this issue can be framed through Pedagogical Content Knowledge (PCK), a concept introduced by Shulman (1986). PCK highlights the integration of subject matter knowledge with pedagogical strategies as the foundation of effective teaching (Abidin et al., 2023; Hubbard, 2018; Ningsih et al., 2020). Research on teacher professional development further emphasizes that PCK is shaped not only by initial training but also by contextual factors, professional experiences, and institutional support. These insights underscore the importance of examining both background and institutional influences when evaluating teacher education outcomes (Krepf et al., 2018; Ningsih et al., 2020). This framework also motivates the specific variables examined in this study: undergraduate background is treated as a proxy for the type and depth of prior pedagogical training; UTUL, an academic-mastery measure, is treated as an indicator of the subject-matter-knowledge component of PCK; UKIN, a practicum-based teaching-performance measure, is treated as an indicator of the applied, classroom-enactment component of PCK; institutional affiliation is treated as a proxy for variation in program delivery and support; and year of graduation is treated as a proxy for the recency of formal training. Framing the variables this way clarifies why each was expected to relate to PCK, while also acknowledging that none of them measures the PCK construct directly.

Against this backdrop, the present study aims to map the performance of physics graduates with educational and non-educational backgrounds in the TPEP program across 31 Indonesian universities. Specifically, it investigates (1) differences in PCK, UKIN, and UTUL scores between the two groups, (2) institutional variations in outcomes, (3) the predictive power of background, year of graduation, and academic scores on PCK, and (4) graduate profiling through cluster analysis. By addressing these issues, the study contributes empirical, program-level evidence on a comparison that has rarely been examined directly within a single national TPEP cohort, offering a descriptive and associational basis that can inform, though not by itself determine, discussions of teacher professional education policy.

METHOD

This study employed a quantitative ex post facto design to analyze the performance of physics graduates in the Teacher Professional Education Program (TPEP). The focus was on mapping differences between graduates with educational and non-educational backgrounds, as well as identifying institutional variations and predictors of pedagogical content knowledge (PCK).

a. Participants

The study involved 728 physics graduates from 31 universities (LPTKs) across Indonesia who completed the TPEP program. Of these, 661 (90.8%) held an educational background and 67 (9.2%) came from non-educational backgrounds. The participants represented diverse graduation years, which allowed for analysis of temporal effects on performance. The 728 graduates represent the complete set of physics-track TPEP completers with fully recorded assessment data available in the national TPEP administrative dataset for the covered period across the 31 participating LPTKs; participants were not additionally sampled or selected from a larger eligible pool beyond the completeness criterion described in the Procedure below. This should be understood as a population-level administrative dataset for this specific program track and period rather than a probability sample, which affects how findings generalize. The resulting 10:1 imbalance between educational- and non-educational-background graduates reflects the actual enrollment composition of the program rather than a sampling artifact, but it is nonetheless an important methodological consideration: it reduces statistical power to detect differences involving the smaller group, and it was addressed directly by examining equality-of-variance assumptions before interpreting the group-comparison tests reported below (see Data Analysis).

b. Instruments and Data Sources

Secondary data were obtained from official TPEP assessment records. Three main assessment components were used: (1) Pedagogical Content Knowledge (PCK) test scores, (2) UKIN (performance-based teaching practicum scores), and (3) UTUL (academic test scores). These instruments are standardized components of the national TPEP evaluation system and are used for certification purposes. Data also included participant background (educational vs non-educational), university affiliation, and year of graduation. PCK is assessed through a standardized written examination targeting participants' integration of subject-matter and pedagogical knowledge, scored on a 0-100 scale. UKIN is a performance-based assessment of actual classroom teaching practice during the TPEP practicum, conducted by trained assessors using a structured rubric and scored on a 0-100 scale. UTUL is a standardized academic-mastery test of physics content knowledge, also scored on a 0-100 scale. All three instruments are administered nationally as mandatory components of TPEP certification under the Ministry of Education, Culture, Research, and Technology's assessment framework; as centrally administered certification instruments, their scoring procedures are standardized across institutions, though participant-level psychometric evidence (e.g., item-level reliability coefficients) was not available in the secondary dataset used for this study and is noted as a limitation below.

c. Procedure

Data collection was carried out by compiling official assessment results from the TPEP system. To ensure quality, only complete datasets with information on background, institution, and all three assessment scores were included in the analysis. Data were cleaned, coded, and categorized before statistical testing. A complete-case approach was used: records missing any of background, institutional affiliation, or one of the three assessment scores were excluded prior to analysis. The administrative dataset provided by the Directorate General of Higher Education did not retain a separable log of excluded incomplete records distinct from the final analytic file, so the exact number and characteristics of excluded cases cannot be reported retrospectively; this is acknowledged as a limitation, since a complete-case approach can introduce selection bias if missingness is not random with respect to background or institution.

d. Data Analysis

Analysis was done using SPSS and Python statistical software. The use of descriptive statistics was done to describe the participants and obtain score results. An independent t-test was performed on educational and non-educational graduates. Given the substantial imbalance between the two background groups ($n = 661$ vs. $n = 67$), Levene's test for equality of variances was examined prior to interpreting each t-test, and results were interpreted in light of this imbalance rather than assuming equal variances by default. One-way ANOVA was used to analyze institutional differences, addressing the research question on institutional variation, while two-way ANOVA was used to investigate the interaction effect of background and institutions; because several of the 31 universities contributed few non-educational-background graduates, cell sizes for some university-by-background combinations were small, and interaction results were interpreted with this sparse-cell limitation in mind. Analysis of covariance was done in order to control the year of graduation, addressing the predictive-power research question alongside the subsequently reported regression model, and finally, regression analysis was used to predict factors statistically associated with PCK. Cluster analysis was done to group the profiles of the graduates according to their PCK and UTUL score, the two variables that were entered into the clustering procedure and consistently reported in the corresponding Results subsection.

RESULTS & DISCUSSION

This section presents the empirical findings of the study, beginning with participants' demographic profile, followed by descriptive statistics, comparative analyses, institutional mapping, regression modeling, and graduate profiling through cluster analysis. The results are subsequently interpreted in relation to previous studies and the theoretical framework of Pedagogical Content Knowledge (PCK).

Findings

a. Profile of Participants

Before moving on to the comparative and inferential analysis, it is important to first discuss the characteristics of the research participants in order to give a complete profile of the data set. This is important since the characteristic of the participants will determine the context in which the statistical analysis that follows will be considered. With regard to studies dealing with teacher professional education, participant characteristics such as educational background and institutional affiliation can influence how results will be interpreted as well as the generalizability of the results (Aydin-Gunbatar et al., 2020). For this reason, profiling the participants will help the readers to determine the representativeness of the participants used in the study and have an idea of the diversity of the participants in the analysis. In this case, participants came from different Teacher Education Institutions (LPTKs) from across Indonesia and they are from graduates of educational background and non-educational backgrounds. Looking at the characteristics of the participants gives us some clues regarding the balance of comparison groups and the institutional representation in the analysis. This is important in the interpretative stage of PCK, UKIN, UTUL, institutional differences, and graduate profile analysis.

Table 1. Distribution of Educational Backgrounds

Educational Background	Count	Percentage (%)
Educational	661	90.8
Non-Educational	67	9.2
Total	**728**	**100.0**

The participants' profile in Table 1 shows that the Teacher Professional Education (TPEP) for physics teachers program is composed mainly of graduates from an educational background, which is 90.8% of the whole sample, whereas those from a non-educational background only compose 9.2%. This composition is reflective of the current recruitment strategy for teachers in Indonesia wherein the graduates from teacher education institutions continue to become the main sources of applicants for professional education programs. The much smaller percentage of non-educational graduates also indicates that although there have been attempts to open up new channels to solve teacher shortages especially on science teaching, this has not yet materialized (Su, 2023). Nonetheless, the existence of these two groups of participants presents an important opportunity to see if there are differences in terms of their competence in their professional education in relation to their different backgrounds in undergraduate education. The substantial sample size, which consists of 728 participants from 31 universities, also provides a robust foundation of the dataset, making subsequent statistical analyses reliable. Additionally, the unbalance in group size also mirrors the real-world situation rather than sampling bias and therefore, enhances ecological validity of the research results (Ong & Annamalai, 2024). Thus, the participant composition becomes a suitable starting point in investigating if educational background has an impact on PCK, UKIN, and UTUL while at the same time providing an opportunity to analyze institutional differences in the later parts of this dissertation.

While the general composition of participants in terms of their educational backgrounds has provided some insight about the sample structure, it does not show the diversity of institutions participating in the study. As mentioned, the participants in the study were recruited from 31 Teacher Education Institutions (LPTKs) in Indonesia and therefore, it becomes necessary to first examine how they are distributed across universities before making any comparative and inferential analysis. Knowing this institutional distribution is critical since difference in participant composition can affect the interpretation of the analysis of the institutions in the next part of the discussion. Universities with more participants make a larger contribution to the overall dataset while those with smaller participants are more varied due to small sample sizes. Consequently, looking at the distribution of participants across universities can highlight the extent of the institutional coverage obtained in this study. The distribution is shown in Figure 1.

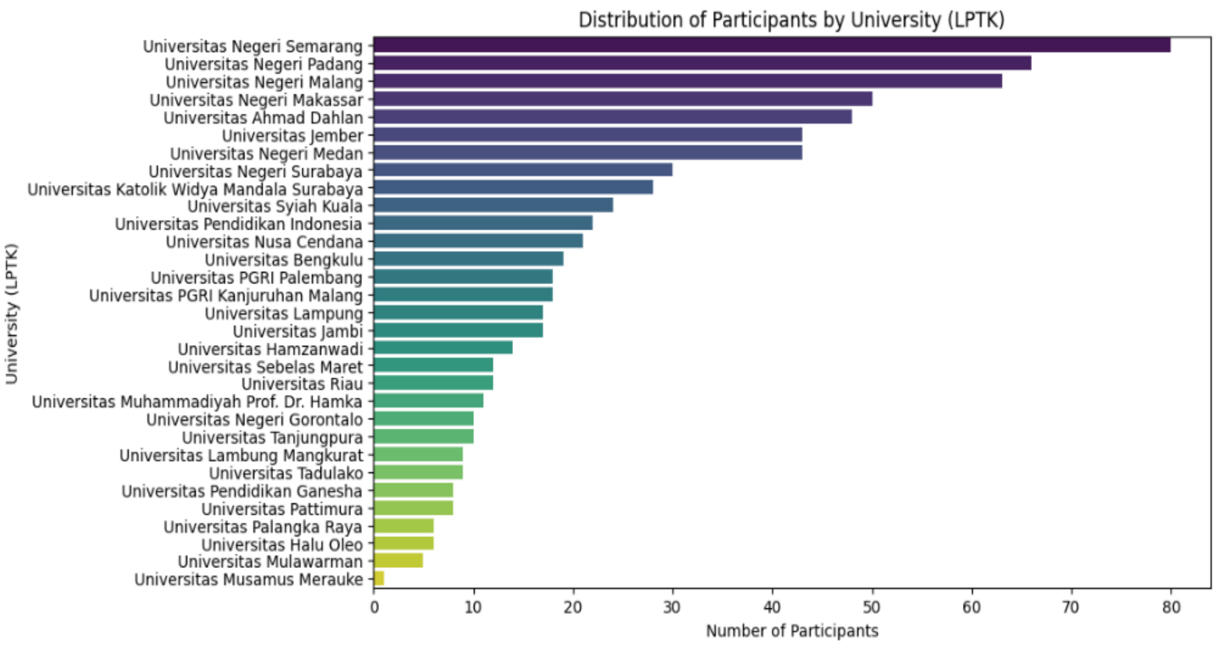

Figure 1. Distribution of Participants by University

Figure 1 below shows the representation of participants in the sample from among the 31 institutions of higher learning, i.e., LPTKs. It is evident that there were wide variations in the number of participants obtained from each institution, which means that there was an uneven distribution of participants in the research. Universitas Negeri Semarang had the highest representation of participants, followed by Universitas Negeri Padang and Universitas Negeri Malang, while other institutions like Universitas Palangka Raya, Universitas Halu Oleo, Universitas Mulawarman, and Universitas Musamus Merauke had very few participants. This type of representation is actually reflective of how many participants actually enrolled in the national TPEP program. This wide representation of institutions will make the external validity of the study better since it includes various types of institutions involved in teacher education in various geographical and educational settings in Indonesia. At the same time, this variation might imply that certain institutional factors could have caused some institutions to contribute more participants than others (Komarudin & Suherman, 2024). For example, some institutions might be more capable than others in providing participants due to certain institutional factors like the size of the program, recruitment strategy, availability of resources, and demand for physics teachers in a certain region.

b. Descriptive Analysis of Scores

After introducing participant characteristics, the next phase will be to analyze the descriptive statistics on the three components of assessment that were done as part of the Teacher Professional Education (TPEP) program, which include Pedagogical Content Knowledge (PCK), teaching practicum assessment (UKIN) and academic competency examination (UTUL). The descriptive statistics provide an early look at the performance of the participants by giving an indication of the score distributions and measures of central tendencies without making any inferential analysis (Nkundabakura et al., 2024). This is a very crucial phase of the study because it helps to provide an objective comparison between the graduates of educational programs and those of non-educational programs in terms of overall achievement. Apart from giving an idea about the difference in the means of two groups, the descriptive statistics also give an idea about the variability of scores and the range of performances of the two groups. Thus, the descriptive statistics in this section provide an initial picture of the performance of the graduates in the three competency components.

Descriptive statistics of the performance of participants in three assessment components in the Teacher Professional Education (TPEP) program are shown in Table 2 above. Generally, participants who graduated with both educational and non-educational degrees had very similar performance in terms of PCK, teaching practicum (UKIN) and academic competency (UTUL). This study's own descriptive comparison shows that although graduates of non-educational programs got a slightly better mean score for PCK, the difference between two groups is very insignificant relative to their standard deviations (cf. Mahler et al., 2024, who similarly report overlapping score distributions across comparably heterogeneous entry-background groups in a related certification context). This means that both distributions of scores should have considerable overlap. In addition, the differences in mean scores for UKIN and UTUL were insignificant, which suggests that graduates from both types of universities demonstrated very similar performance in teaching and academics. In other words, the score ranges of participants suggest that both groups include people who are either low-, moderate-, or high-achievers. Hence, at the descriptive level, the data show that the type of undergraduate degree can hardly make a difference in participants' performance in TPEP. On the contrary, these homogeneous score distributions can be explained by professional education. However, it does not mean that these differences are statistically significant. Descriptive statistics alone can tell very little about the significance of a difference (Li & Copur-Gencturk, 2024). In other words, to test whether these differences exist, it is necessary to conduct inferential analyses.

Table 2. Descriptive Statistics of Scores by Educational Background

Statistic	Non-Educational (n = 67)	Educational (n = 661)
PCK (Mean ± SD)	60.10 ± 7.06	59.36 ± 6.91
PCK (Min–Max)	40.7 – 76.26	36.1 – 79.67
UKIN (Mean ± SD)	81.76 ± 4.38	82.05 ± 4.21
UKIN (Min–Max)	71.7 – 91.2	66.7 – 85.0
UTUL (Mean ± SD)	68.90 ± 4.68	68.74 ± 4.41
UTUL (Min–Max)	56.55 – 78.70	54.76 – 80.99

Whereas descriptive statistics offer useful information on the average performance and score dispersion, they do not entirely show the relations between assessment components. The knowledge of how strongly PCK, UKIN and UTUL are related is necessary because each assessment evaluates the competences of the teachers in a particular area that can support one another. Correlation analysis of the three variables gives the opportunity to see whether good performance in one of the areas is accompanied by success in another one and thus gives a more detailed understanding of the structure of graduate competencies within the TPEP program (She et al., 2025). In addition, correlation analysis plays a role in the preparation stage for carrying out more complicated statistical analyses, like regression analysis. This analysis shows the strength and direction of the relations between the variables in question. The visualization of these relations through a correlation heatmap makes it easier to understand them than simply correlation coefficients. Therefore, the relations between PCK, UKIN and UTUL are presented in Figure 2.

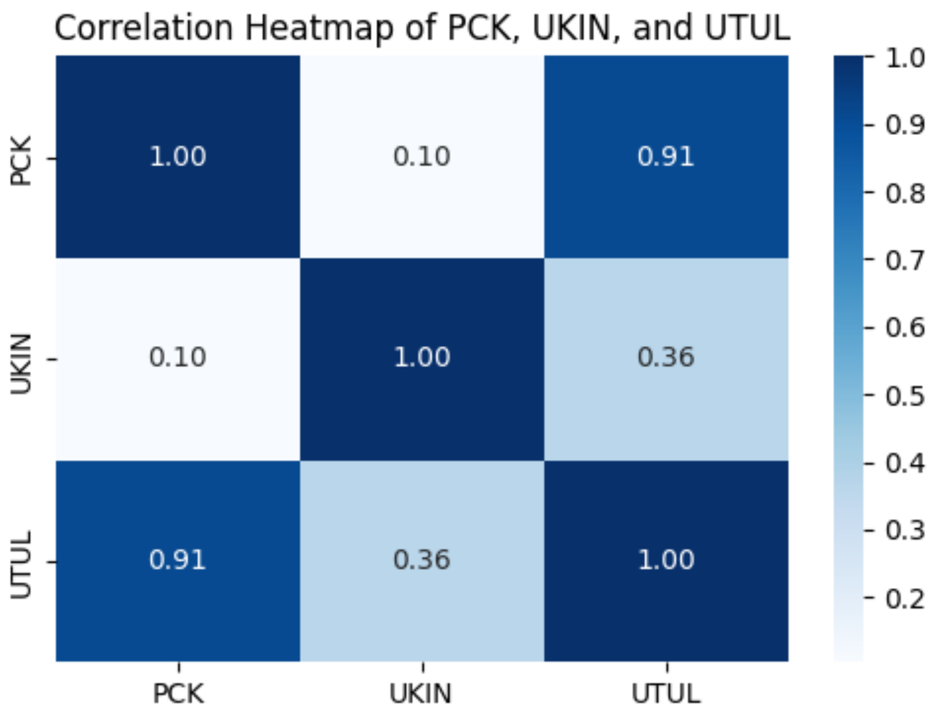

Figure 2. Correlation Heatmap of PCK, UKIN, and UTUL

Figure 2 shows the heatmap of correlation depicting the relationships between PCK, teaching practicum performance (UKIN), and academic competency (UTUL). In general, the presented heatmap demonstrates different patterns of relationships among three assessment measures, thus showing their complementarity. The highest level of correlation was found to be between PCK and UTUL, thus indicating that those students who had better academic mastery were more likely to have higher levels of pedagogical content knowledge (She et al., 2025). Such a result is consistent with the considered theoretical perspective on the fact that teaching implies not only the development of pedagogical skills but also the acquisition of the subject matter so that teachers can use this knowledge for teaching effectively. At the same time, the relationship between PCK and UKIN is relatively lower, thus suggesting that the theoretical knowledge of pedagogical skills and teaching practice are not the same concepts. Both make their contribution to the professional competence of teachers but might evaluate different aspects of their preparation. While the first one evaluates the knowledge of teaching, the latter one focuses on its implementation in practice (Favier et al., 2021). The heatmap also shows the relationship between UKIN and UTUL, which provides additional evidence that academic success of future teachers does not mean teaching practice per se. The described correlation patterns form a strong ground for conducting further regression analysis and examining the contribution of each of three assessment measures to PCK prediction.

c. Independent Samples t-test

Even though the descriptive statistics show that educational and non-educational graduates had comparable mean scores on all the three measures, the difference in their mean scores does not prove that the findings have any statistical significance. To examine if there is any connection between the participants' educational background and their performance, an inferential statistical test named the independent samples t-test was used, the standard approach for comparing means between two independent groups. It is suitable for this research since it compares the means of two different groups while considering the variation of observations in each group. Because the two background groups differ substantially in size ($n = 661$ vs. $n = 67$), Levene's test for equality of variances was examined for each comparison, and the degrees-of-freedom adjustment for unequal variances was applied where the equality assumption was not supported. The independent samples t-test was applied in this case to compare the performance of educational and non-educational graduates in Pedagogical Content Knowledge (PCK), teaching practicum (UKIN) and academic competency (UTUL). Effect sizes based on Cohen's d measure were used in addition to statistical significance of differences. Reporting both significance values and effect sizes will provide more detailed information by separating statistical and practical importance (Lakens, 2013). The outcomes of the independent samples t-test are shown in Table 3 below.

Table 3. Independent Samples t-test Results

Variable	t-value	p-value	Cohen's d
PCK	-0.820	0.415	-0.11
UKIN	0.520	0.604	0.07
UTUL	-0.265	0.792	-0.04

Table 3 summarizes the results of the independent samples t-test conducted to compare the performance of educational and non-educational graduates across the three assessment components. As can be seen, the differences between the two groups of participants turned out to be very insignificant in terms of magnitude and statistical significance. To begin with, in the case of Pedagogical Content Knowledge (PCK) the t-value equals to -0.820 while p-value is 0.415. Thus, despite the slightly higher average score (60.10 compared with 59.36) demonstrated by non-educational graduates, the obtained difference is not statistically significant, meaning that the observed score difference cannot be considered a reliable indicator of the true difference between the two groups of graduates. Besides, the effect size measured with Cohen's d in this case equals to -0.11 which also indicates that the difference is trivial and has no practical significance.

The same trend is observed for the assessment component called UKIN. Despite the slightly higher average score (82.05 compared with 81.76) achieved by the educational graduates, the difference equals to 0.520 for the t-value and to 0.604 for p-value, meaning that the observed score difference cannot be considered as statistically significant since it can be attributed to random sampling variability. Cohen's d of 0.07 in this case means that the difference is negligible and is unlikely to have any practical significance as well.

Finally, the assessment component UTUL revealed the smallest difference between the two groups. The average score of educational graduates (68.90) differs from the non-educational ones' (68.74) by only 0.16 point, resulting in t-value = -0.265 and p-value = 0.792. Extremely high p-value in this case clearly indicates the high similarity of academic competency between the two groups of graduates. The effect size (Cohen's d = -0.04) is almost zero as well, meaning that the educational background plays hardly any role in the differences between the two groups' scores.

Overall, the results obtained across the three variables indicate the consistent trend. Irrelevant of whether the teacher competence is assessed through pedagogical content knowledge, classroom teaching performance, or academic examinations, educational and non-educational graduates perform in the same way. Small effect sizes (from -0.11 to 0.07) provide the evidence that the differences are not only statistically insignificant but also negligibly small. No statistically significant difference was detected between the two background groups on any of the three measures; consistent with R1/R27-type cautions, this indicates an absence of detected difference rather than proof that background has no effect on competence. The obtained results are consistent with, though do not by themselves establish, the interpretation that TPEP program serves as an equalizing professional education path that enables people from different educational backgrounds to become equally competent after completion of the program; because this is a cross-sectional, post-completion comparison without a pre-TPEP baseline, comparable outcomes at completion could also reflect selection into the program or other unmeasured factors rather than the leveling effect of TPEP itself (see Limitations).

d. Institutional Mapping (University Differences)

The above analysis showed that educational and non-educational graduates performed equally well in all the three evaluation areas. But while educational background is one way of differentiating between the performance of the graduates, another factor that can have an impact on the learning of graduates is the university where the participants took the Teacher Professional Education (TPEP) program. According to Lachner et al. (2021) there could be many factors, like curriculum implementation, instructional practices, mentoring programs, academic supports, and assessment policies, that could lead to variations in the performance of the graduates based on their Pedagogical Content Knowledge (PCK). Thus, instead of assuming all the Teacher Education Institutions (LPTKs) produce graduates of equal level of competency, it is necessary to find out if there are any differences in the performance of graduates from different universities taking part in the national TPEP program. This will help in evaluating institutional differences in professional teacher education. A one-way ANOVA was done based on the university where the participant belongs. The result of this test is presented in Table 4.

Table 4. One-Way ANOVA Results for PCK by University (LPTK)

Source	Sum of Squares	df	F	p-value	Conclusion
LPTK	6703.27	30	5.53	<0.001	Significant difference
Residual	28145.00	697	—	—	—

The results of one-way ANOVA analyzing the differences in Pedagogical Content Knowledge (PCK) among LPTKs graduates are provided in Table 4. According to the results, there is a statistically significant effect of institutional affiliation, as it is evidenced by the high value of F-statistics equal to 5.53 and by the low value of the probability level (less than 0.001). This means that the differences in the mean values of PCK among universities are unlikely to occur just by accident and reflect real, systematic variation in mean PCK across institutions, but rather they are real differences caused by certain factors; the specific institutional factors responsible for this variation (e.g., curriculum implementation, mentoring, instructional quality) were not directly measured in this study and are discussed as candidate explanations, not established causes, in the Discussion section. The between-group sum of squares amounted to 6703.27, while the residual sum of squares was 28145.00. In spite of the higher value of the residual sum of squares compared to the one calculated based on the differences between universities, the ratio of these two variances is sufficient enough to ensure the significance of the F-statistic.

The results of ANOVA were calculated based on 31 universities using 30 degrees of freedom for the institutional factor and 697 degrees of freedom for the residual error. High residual degrees of freedom suggest that there was a sufficiently high number of observations used in order to conduct the analysis (Tseng et al., 2022). Thus, a significant F-value means that at least one institution differs significantly from other institutions in its average performance measured in terms of PCK. However, since the ANOVA does not reveal the particular institution(s), further descriptive visualization will be needed in order to analyze the performance of the particular universities. Formal post-hoc multiple-comparison testing (e.g., Tukey's HSD) with institution-pairwise significance and an accompanying effect-size measure was not conducted for this analysis; the visual comparison presented below (Figure 3) is therefore descriptive rather than inferential, and claims about which specific institutions differ from which others should be read with this limitation in mind.

Thus, the results are consistent with the role of the institutional context in the competence of graduates being non-trivial. There may be many aspects of the instructional process, which may affect the pedagogical development of graduates in the TPEP program, including differences in instructional quality, learning environment, mentoring system, assessment procedures, etc. These remain plausible explanations rather than tested mechanisms, since none of these process-level variables was directly measured in this study. Thus, it means that the quality of teacher education cannot be assessed only based on individual characteristics but also, potentially, on the performance of institutions.

While it is true that one-way ANOVA has established the existence of statistically significant institutional differences, the statistical numbers themselves fail to provide any information regarding the variations in performance of graduates from individual universities. It would be helpful, therefore, to analyze the average PCK scores obtained by graduates from the respective participating university. The graphical representation makes it easier for the reader to notice such patterns as high, moderate, or low average PCK scores for universities (Angeli & Valanides, 2009). At the same time, it makes it easier to visualize the degree of difference, which is reflected in statistically significant ANOVA results. In addition, the information about the distribution of graduate performance across the entire national TPEP program will help in determining universities that can act as an example of best practices as well as universities requiring additional academic assistance. The average PCK score distribution across the participating universities is represented in Figure 5.

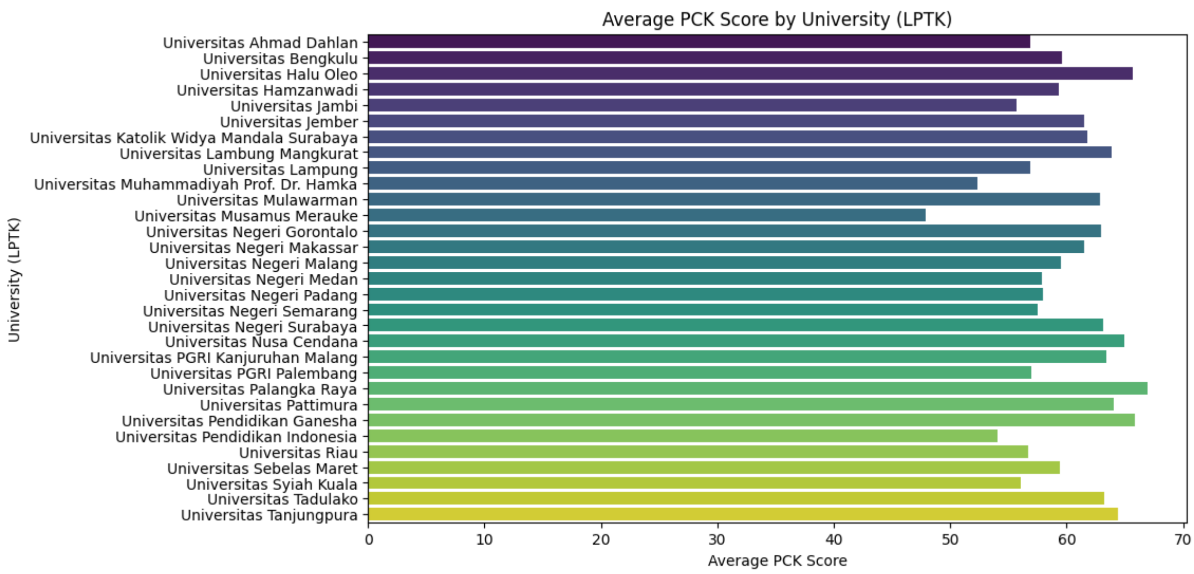

Figure 3. PCK Scores by University

Figure 3 shows the distribution of average PCK scores between institutions involved in this research. From this figure, it can be seen that there is not only one level of graduate performance in pedagogical content knowledge among all universities. Some of them have managed to get significantly high mean scores of PCK while other institutions had quite low results. The institution having the highest mean PCK score is Universitas Palangka Raya (66.98). Then comes Universitas Pendidikan Ganesha (65.89) and Universitas Halu Oleo (65.66). It can be suggested that graduates of these universities have higher achievements in PCK because they have more effective teaching methods or positive study conditions in their TPEP programs.

On the contrary, some institutions showed quite low results. The institution that had the lowest mean score of PCK was Universitas Musamus Merauke (47.85). In fact, this score is more than 19 points lower than the highest average score at Universitas Palangka Raya. Universitas Muhammadiyah Prof. Dr. Hamka also got a quite low average score (52.30). These figures confirm great diversity in graduates' results in different universities. The wide distribution of average scores shown in Figure 3 gives additional proof for the significance of the ANOVA analysis conducted before.

Such significant differences in performance among universities may mean that curriculum implementation, supervision effectiveness, learning facilities, academic mentoring, and evaluation techniques have an impact on PCK skills of graduates. However, these factors may not have an equal impact on graduates from educational and non-educational backgrounds. Therefore, it becomes necessary to check whether educational background influences graduate performance depending on institutions. It can be done with the help of two-way ANOVA presented in Table 5.

Table 5. Two-Way ANOVA Results for PCK by University and Background

Source	Sum of Squares	df	F	p-value	Conclusion
University (LPTK)	6674.50	30	5.43	< 0.001	Significant
Background (Educational vs Non-Educational)	13.38	1	0.33	0.568	Not significant
University × Background	3220.97	30	2.62	< 0.01	Significant

The results of two-way ANOVA with university affiliation, educational background, and interaction of both variables as independent predictors and PCK as dependent variable are displayed in Table 5. There were three crucial observations made during the interpretation of the two-way ANOVA results. First, university affiliation was a highly significant predictor of PCK, as it yielded F-value of 5.43 with a p-value under 0.001. This result corroborated with the previous one-way ANOVA and showed that pedagogical competence of graduates was significantly associated with participating universities. Institutional sum of squares was equal to 6674.50, thus showing that universities play an important role in explaining the difference in PCK scores.

Second, the educational background per se did not make a big difference in the performance of graduates. The educational background factor resulted in a sum of squares of 13.38 with 1 degree of freedom, which led to F-value of 0.33 and p-value of 0.568. As compared to a much bigger institutional sum of squares, the variation explained by the educational background is insignificant. It confirms that the graduates from educational and non-educational backgrounds show similar performance after undergoing the TPEP program, thus confirming the independent samples t-test results (Schmidt et al., 2009).

Third, the most interesting result relates to the interaction of university affiliation and educational background. The interaction effect resulted in a sum of squares of 3220.97 with 30 degrees of freedom and yielded F-value of 2.62 with p-value under 0.01. Even though the variation explained by the interaction was smaller than institutional effect, the interaction was still statistically significant. It means that the influence of the educational background varies depending on the university rather than being constant for all participating universities (Mishra & Koehler, 2006). While at some universities the performance of graduates from both backgrounds is almost identical, at other universities there is a noticeable difference between one group and another. Because non-educational-background graduates (n = 67) are spread thinly across 31 universities, several university-by-background cells contain few observations; this sparse-cell pattern should be kept in mind when interpreting the interaction, as estimates for institutions with very few non-educational graduates are less stable than those with larger cell sizes.

Thus, the two-way ANOVA showed that the most important factor associated with graduate performance was the institutional context, while the educational background per se had no statistically significant average (main-effect) association with graduate performance once university affiliation and their interaction were taken into account. Still, a statistically significant interaction showed that there is a moderation of the development of pedagogical competencies among graduates with different educational backgrounds by the institutional practices; because formal simple-effects follow-up tests were not conducted, the non-significant main effect should not be read as background being uniformly irrelevant, since the interaction indicates that its association with PCK does vary meaningfully by institution. In order to illustrate the pattern of interactions, Figure 4 presents the distribution of PCK scores by university and educational background.

The interaction effect revealed by the two-way ANOVA implies that there is no linear dependency between the educational background of respondents and their Pedagogical Content Knowledge (PCK). Although the interaction effect has been verified statistically, numeric values are not enough to understand in what way the interaction effect is observed at different universities. It is necessary to visualize the difference between performance of respondents who received an education and those who did not receive an education in order to be able to reveal the level of this difference at different universities (Taşar & Yılmaz Ergül, 2023). Thus, it will become possible to reveal universities at which the performance of the two groups of respondents does not vary and universities at which performance gaps appear. The interaction effect between university and educational background of respondents is shown in Figure 4 below.

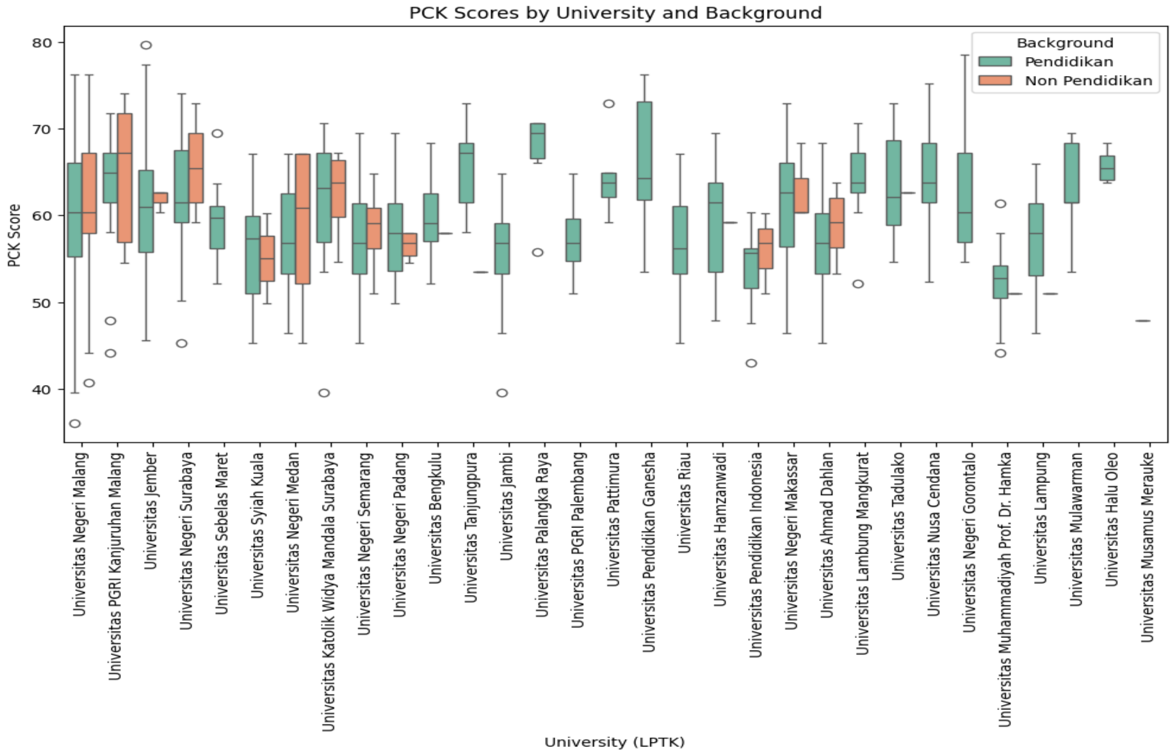

Figure 4. PCK Scores by University and Background

Figure 4 illustrates the interaction between university affiliation and educational background in explaining graduates' Pedagogical Content Knowledge (PCK). The figure demonstrates that the performance differences between educational and non-educational graduates vary considerably across the participating universities. In several institutions, the average PCK scores of the two groups are almost identical, indicating that educational background has little influence on graduate performance within those universities — a pattern broadly consistent with the non-significant main effect reported above. This pattern suggests that institutional learning environments and professional preparation effectively support participants regardless of their undergraduate discipline.

In contrast, several universities exhibit noticeably larger performance gaps between the two groups. In these institutions, either educational graduates outperform non-educational graduates or vice versa, indicating that institutional practices may interact differently with participants' prior academic experiences. For example, some universities appear to provide learning environments that enable non-educational graduates to perform at levels comparable to educational graduates, whereas other institutions display greater disparities between the two groups. These variations visually explain why the interaction effect observed in the two-way ANOVA reached statistical significance despite the absence of a significant main effect for educational background (Celik, 2023).

Overall, Figure 4 reinforces the conclusion that educational background should not be interpreted independently of institutional context. Rather than functioning as a universal determinant of graduate competence, its statistical association appears to depend on how individual universities implement the TPEP curriculum, organize mentoring systems, provide academic support, and facilitate professional learning experiences. These findings further are consistent with the interpretation that institutional quality plays a more influential role than initial academic background in shaping graduates' pedagogical content knowledge, though, as noted above, the specific institutional mechanisms were not directly measured and this remains an association rather than a confirmed causal pathway within Indonesia's Teacher Professional Education Program.

e. ANCOVA and Regression Analysis

It was shown in the previous analysis that there were no significant differences in PCK of the graduates depending on their educational background, whereas the institution played an important role in the differences. However, the analyses conducted previously have not taken into account the other factors that could influence the performance of graduates at the same time. In particular, the graduation year may play an important role in pedagogical competence of the participants since the most recently graduated ones have more up-to-date academic knowledge and skills. Thus, a more complex analysis should be conducted in order to find out whether the educational background continues to have a certain impact on PCK after the effect of time has been controlled for. For this purpose, ANCOVA was conducted in which the year of graduation served as a covariate and educational background was kept as a factor. Such an analysis allows evaluating the independent effect of each variable while eliminating the possible confounding factors (Iskandar et al., 2021). The results of ANCOVA also serve as the basis for regression analysis in which the contribution of several predictors in explaining the variance in PCK is assessed. The results of ANCOVA are presented in Table 6.

Table 6. ANCOVA and *Regression Analysis*

Source	Sum of Squares	df	F	p-value	Conclusion
Background	33.39	1	0.70	0.404	Not significant
Year of Graduation	242038.33	1	5047.26	< 0.001	Significant

The results of ANCOVA are provided in Table 6, where the effects of the explanatory variables, such as educational background and year of graduation, on pedagogical content knowledge were considered. The analysis shows evident differences between the two explanatory variables. The educational background explains a sum of squares of 33.39 with 1 degree of freedom and thus has the F-value of 0.70 and the p-value of 0.404. These statistics demonstrate that after accounting for the year of graduation, the educational background does not explain the variability in graduates' PCK scores. In addition, the rather low value of the sum of squares proves that a negligible amount of the total variability in PCK is caused by the type of graduates' undergraduate education.

On the contrary, the year of graduation proved to be a very important variable statistically associated with the graduates' PCK. The year of graduation explains a sum of squares of 242,038.33, which is thousands of times higher than the effect of educational background. The F-value of 5047.26 and the p-value less than 0.001 prove that the year of graduation is strongly associated with graduates' PCK in this ex post facto design; given the unusually large F-value, this result is also flagged for cautious interpretation, since it may partly reflect cohort-, curriculum-, or dataset-structure effects that a purely statistical test cannot distinguish from a direct effect of time itself (see Limitations). The high value of F-statistics shows that the systematic variation caused by the year of graduation largely surpasses the unexplained residual variation.

Such a difference between the explanatory powers of the two variables is consistent with the interpretation that the recency of graduates' academic preparation has a greater statistical association with PCK compared to their educational background. Students who have finished their studies recently are expected to possess better knowledge in the discipline and be familiar with recent trends in teaching methods and curriculum, which can affect their PCK (Mayer & Girwidz, 2019). These findings are consistent with the results of the previous analysis according to which the educational background alone has no explanatory power but simultaneously identifies the year of graduation as an important statistical determinant of graduates' achievements. However, ANCOVA analyzes the adjusted mean difference associated with a categorical factor (educational background) while statistically controlling for a covariate (year of graduation), providing a controlled group comparison; it does not consider the effects of the variables simultaneously as a joint predictive model. Therefore, the multiple regression analysis was performed next to estimate the influence of the educational background, year of graduation, UKIN, and UTUL on pedagogical content knowledge simultaneously, which serves the complementary purpose of quantifying each predictor's independent statistical association with PCK while holding the others constant. The results of regression analysis are provided in Table 7.

Table 7. Regression Results

	coef	std er	t	P>|t|	[0.025	0.975]
Background_bin	-0.368	0.295	-1.244	0.214	-0.947	0.212
Year_Graduation	-0.006	0.001	-7.091	0.000	-0.008	-0.005
UKIN	-0.429	0.022	-19.765	0.000	-0.472	-0.387
UTUL	1.572	0.021	75.945	0.000	1.532	1.613

Table 7 summarizes the findings of multiple regression that was conducted to examine the simultaneous statistical association of the explanatory variables (educational background, year of graduation, UKIN, and UTUL) with the level of Pedagogical Content Knowledge (PCK). The obtained regression shows that all the explanatory variables have different statistical associations with the level of graduates' pedagogical competence. First of all, the magnitude and statistical significance of the regression coefficients vary greatly.

In this study, educational background produced a regression coefficient β=-0.368 (t=-1.244; p=0.214). Given that the probability value is higher than the commonly accepted criterion of significance, which equals 0.05, it is possible to conclude that graduates' educational background is insignificant for the explanation of graduates' PCK after accounting for the impact of the other variables. In addition, the 95% confidence interval (-0.947 to 0.212) contains zero and supports the conclusion that the contribution of educational background to the prediction of graduates' PCK can be neglected. This finding complements the results of the independent t-test and ANCOVA, according to which graduates with educational and non-educational backgrounds demonstrated equal levels of pedagogical competence.

Another variable, year of graduation, shows a completely different picture. Despite the relatively low value of the regression coefficient β=-0.006, this variable produced a t-value of -7.091 with a p-value below 0.001. Therefore, the contribution of the year of graduation into the regression is highly significant. The very narrow confidence interval (-0.008 to -0.005) confirms the reliability of this estimation. According to the negative coefficient, graduates of earlier years had slightly lower PCK scores in comparison to more recent graduates. This association is compatible with, but does not by itself confirm, the interpretation that knowledge recency plays an important role in professional teacher education; because graduation year may also capture cohort, curriculum-revision, or institutional-policy differences that were not directly measured, these alternative explanations cannot be ruled out and are noted here as a caution against a purely 'knowledge recency' reading.

The strongest predictor in the regression model is UTUL. This variable produced a large regression coefficient (β =1.572) that is associated with an exceptionally large t-value (75.945) and p-value below 0.001. The confidence interval (1.532 to 1.613) is narrow and contains only positive numbers. Therefore, it is possible to conclude that UTUL score is a strong and reliable positive predictor of PCK score. Numerically, it means that every one-point increase in UTUL score is associated with approximately 1.57-point increase in PCK. Thus, UTUL shows that academic achievement is the key determinant of graduates' pedagogical competence among the predictors considered.

It is interesting that UKIN has a statistically significant and negative association with graduates' PCK. The regression coefficient $\beta = -0.429$ produced a t-value of -19.765 with a p-value below 0.001 and an entirely negative confidence interval (-0.472 to -0.387). Despite the fact that the negative association of this predictor seems contradictory, one plausible explanation is that the negative relationship between UKIN and PCK can be explained as reflecting a difference in the dimensions of professionals' competences, measured by these variables. UTUL and PCK reflect the aspect of cognitive and conceptual mastery, while UKIN measures only practical classroom teaching performance of the graduate teachers. Therefore, high classroom teaching performance does not guarantee high scores in PCK. Alternative explanations should also be considered, including differences in how the two instruments are scored and scaled, measurement error in either assessment, or a suppression effect arising from UKIN's correlation with the other predictors in the same multivariable model; the present cross-sectional design cannot adjudicate between these explanations, and this finding is best treated as an empirical association warranting further investigation rather than settled evidence that UKIN and PCK are simply 'different dimensions' of competence.

Thus, regression analysis demonstrates that academic competency (UTUL) and year of graduation are the key determinants of pedagogical content knowledge, while graduates' educational background contributes little to prediction of their performance. These results confirm the conclusion about the insignificance of educational background for the effectiveness of the TPEP program.

f. Cluster Analysis (Graduate Profiles)

The preceding analyses primarily focused on identifying differences among participant groups and determining the factors that significantly predict Pedagogical Content Knowledge (PCK). While these approaches provide valuable information regarding statistical relationships and the relative contribution of individual variables, they do not reveal whether participants can be grouped into distinct competency profiles based on their overall performance. To obtain a more comprehensive understanding of graduate characteristics, a cluster analysis was performed using PCK and UTUL scores. Unlike comparative statistical techniques, cluster analysis classifies participants according to similarities in their performance patterns without relying on predefined categories (Aartun et al., 2022). The clustering was performed using the k-means algorithm on standardized PCK and UTUL scores, with the number of clusters ($k = 3$) selected to distinguish low-, moderate-, and high-performing profiles. This data-driven approach enables the identification of naturally occurring groups of graduates who share similar levels of academic and pedagogical competence. Such profiling provides practical insights into the heterogeneity of participant achievement and supports the development of differentiated interventions tailored to specific learner needs. The resulting clusters are illustrated in Figure 5. A formal internal validation index (e.g., a silhouette coefficient) for the three-cluster solution is not reported here and is noted as a direction for methodological strengthening in future work (see Limitations); cluster centroids and the number of participants per cluster are summarized descriptively below.

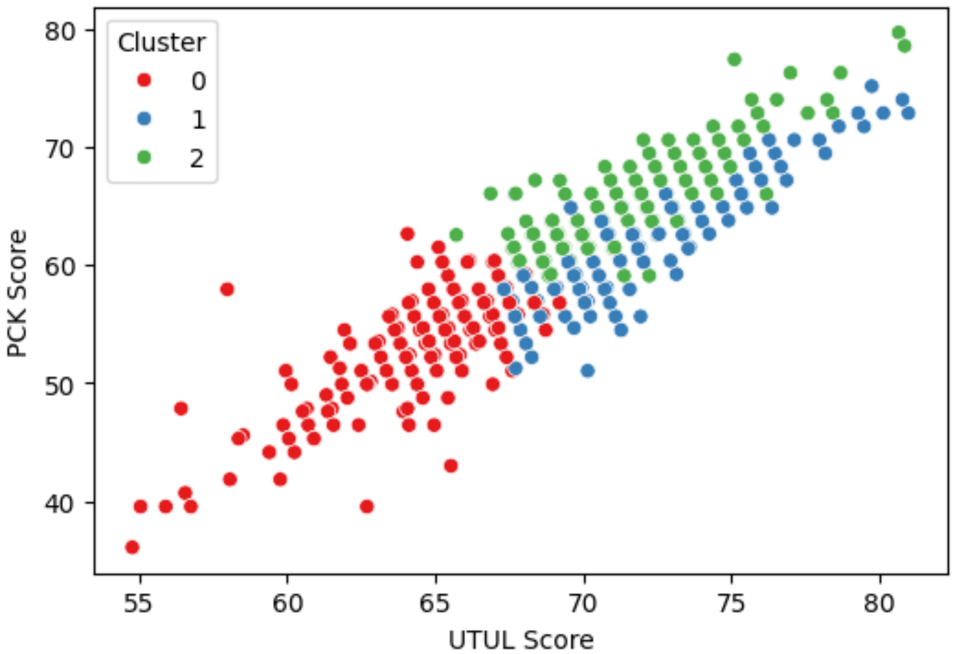

Figure 5. Cluster of graduates based on PCK and UTUL

Figure 5 provides the results of clustering analysis using the PCK and UTUL scores of the participants to detect performance patterns. From the visual representation of the results of this type of analysis, it becomes evident that participants are distributed between three clusters demonstrating three different levels of performance. The first cluster (Cluster 0), represented in red, is located at the lower-left corner of the graph and comprises participants with relatively low PCK and UTUL scores. The participants in this cluster are those with lower academic mastery of content knowledge and pedagogical skills, which means that these participants might need more academic and mentoring support from their mentors during the early years of their work experience.

The second cluster, represented in blue, is located in the center of the graph. This cluster comprises participants with moderately high levels of PCK and UTUL. The participants in this cluster demonstrate satisfactory levels of competency; however, there is still a potential for improvement. The placement of this cluster in the middle, between the lower and the higher clusters, suggests that through professional development, reflective practices, and pedagogical training, the participants can move to the upper level of competence (Leonet et al., 2020). Finally, the third cluster, located in the upper right part of the figure and colored in green, includes participants with high PCK and UTUL scores. The participants in this cluster have high academic mastery and well-developed pedagogical skills, which shows high professional readiness.

The relatively clear distinction of the three clusters suggests that participants have performance levels that differ in meaningful ways rather than randomly across the population. This result corresponds to the results obtained from regression analysis and suggests that even though the educational background does not significantly distinguish graduate competencies, participants tend to form groups of distinct performance levels. Thus, this result supports the importance of using the differentiated strategy for professional development of the TPEP participants based on their performance levels.

Discussion

This study set out to provide an empirical mapping of physics graduates with educational and non-educational backgrounds in Indonesia's Teacher Professional Education Program (TPEP). The findings revealed no significant differences in PCK, UKIN, and UTUL scores between the two groups. This challenges the common assumption that educational graduates are inherently better prepared than non-educational graduates, suggesting instead that the TPEP program may plausibly function as a leveling mechanism that enables graduates with diverse academic trajectories to achieve comparable outcomes (Kamenarac, 2024; Shing et al., 2015) — an interpretation these two conceptual sources help frame theoretically, though neither provides direct empirical evidence for it, and the present cross-sectional, post-completion design cannot rule out alternative explanations such as selection into the program (see Limitations).

The analysis further identified the key factors statistically associated with graduates' performance in TPEP. Regression and ANCOVA results consistently demonstrated that educational background was not a significant predictor of PCK. Instead, UTUL scores and year of graduation emerged as the strongest statistical predictors. These findings underscore the importance of academic mastery and knowledge recency, indicating that graduates who have more recent academic exposure and strong content knowledge are statistically associated with stronger integration of subject matter into effective pedagogy. This is broadly consistent with Shulman's (1986) originating framework of Pedagogical Content Knowledge (PCK), which emphasizes the dynamic interplay between content knowledge and pedagogy, and with subsequent applications of the framework in other teaching contexts such as in-service teacher practice (Bostock, 2019) and social-science research-methods teaching (Nind, 2020). Interestingly, the regression model revealed a statistically significant negative association between UKIN and PCK, suggesting that performance-based assessments and theoretical pedagogical knowledge may capture different dimensions of teacher competence—a stronger and more specific claim than the merely 'relatively lower' PCK-UKIN correlation reported earlier in the correlation analysis, and one that should be read as this study's regression-based finding rather than as something already established by the descriptive correlation results (Sarkar et al., 2024; Schiering et al., 2023; Wang et al., 2024; Amalia & Hodiyanto, 2026).

Cluster analysis added another dimension to the mapping by classifying graduates into three distinct profiles based on PCK and UTUL scores: low, moderate, and high performers. This heterogeneity demonstrates that TPEP graduates are not a homogeneous group but vary considerably in these two dimensions of competence. Low performers require intensive pedagogical reinforcement, moderate performers benefit from capacity-building strategies, and high performers can be supported through advanced professional development opportunities (Bloomfield et al., 2024). Such differentiated pathways resonate with teacher professional development literature, which highlights the need for tailored interventions to address varying levels of readiness and expertise (Ruifeng & Ping, 2021; Solihin et al., 2021).

From a practical perspective, these findings suggest that entry background should not be overemphasized in recruitment policies for TPEP, though this implication should be drawn cautiously: because the analytic sample consists only of graduates who completed TPEP, the data cannot directly speak to how background relates to admission, retention, or completion, and any recruitment-policy implication is an extrapolation beyond what a post-completion outcomes comparison can establish. Instead, greater attention should be given to strengthening academic preparation, designing refresher opportunities for older graduates, and ensuring differentiated support according to graduate profiles. These tiered-support recommendations describe existing performance patterns identified through cluster analysis and are offered as plausible, evidence-informed implications rather than as interventions whose effectiveness has been tested in this study. By focusing on academic knowledge, knowledge recency, and targeted professional development, TPEP can better fulfill its role as a potentially equalizing platform for diverse graduates (Julia et al., 2020).

In the broader context of teacher education, these findings contribute to ongoing debates about the relative importance of initial academic background versus professional preparation. International research has shown that strong teacher education programs can compensate for differences in prior training, provided that they emphasize integration of content and pedagogy, reflective practice, and ongoing professional support (Kertesz & Brett, 2020; Werler & Tahirsylaj, 2020). The present study strengthens this perspective by demonstrating that within the Indonesian TPEP framework, non-educational graduates can achieve outcomes comparable to educational graduates, provided that academic knowledge and knowledge recency are ensured (Darling-Hammond, 2023; Palincsar et al., 2020).

The contribution of this study also lies in the use of cluster analysis to profile graduate competencies. While most previous research on background-based comparisons in teacher certification programs has compared groups based on single dimensions such as test performance or background using group-mean comparisons, the identification of low, moderate, and high performers through a data-driven clustering approach provides a more nuanced mapping of graduate diversity than a simple two-group comparison could reveal (Roca-Campos et al., 2021; Nsingo & Chirume, 2026), though whether this three-cluster solution is sufficiently robust to support practical use should be confirmed through internal validation (e.g., silhouette analysis) and, ideally, replication in future work. This typology can serve as a practical tool for designing differentiated interventions within teacher education programs, offering tailored support that addresses the unique needs of each group.

Finally, while the implications extend beyond the Indonesian context only tentatively, since this study examined a single program (TPEP) in a single country and a single subject area (physics), many countries facing teacher shortages adopt flexible entry pathways that bring together candidates with diverse academic backgrounds (Pawan et al., 2023). The present findings suggest that the success of such policies may plausibly depend less on prior background and more on the capacity of teacher education programs to provide rigorous, targeted, and adaptive training (Nganga & Kambutu, 2024; Parkhouse et al., 2019), though this cross-national extension has not itself been tested here and would require comparative, multi-country data to confirm. As such, this study underscores the importance of designing teacher professional education systems that are responsive to graduate diversity and that build competence through structured academic and pedagogical reinforcement.

Despite its contributions, this study has limitations. The data were restricted to physics graduates, which may limit generalizability to other subject areas, and, correspondingly, the policy recommendations offered above should be understood as most directly applicable to physics TPEP rather than to teacher certification generally. Moreover, while the study employed robust quantitative methods, complementary qualitative insights into graduate experiences would enrich the understanding of how TPEP shapes professional competence (Musa et al., 2024). Longitudinal studies tracking graduates into their teaching careers are also needed to evaluate the long-term impact of TPEP training (C.-F. Chang et al., 2024; Warahmah et al., 2021). Several additional limitations should be noted. First, this is a secondary, cross-sectional, ex post facto design: it can establish statistical associations and group differences but cannot establish causal relationships among educational background, graduation year, institutional context, and PCK, and all causal-sounding language in this manuscript should be read in that light. Second, the dataset includes no pre-TPEP baseline measures, so comparable outcomes between background groups at completion cannot distinguish a genuine equalizing effect of TPEP from pre-existing similarities or selection into the program. Third, the sampling frame reflects a population-level administrative extract for physics-track TPEP completers with complete records rather than a probability sample, and the exact number and characteristics of records excluded for incompleteness could not be retrospectively reconstructed from the dataset provided, which leaves some uncertainty about whether excluded cases differed systematically from those analyzed. Fourth, the substantial imbalance between educational-background (n = 661) and non-educational-background (n = 67) graduates reduces statistical power for detecting true differences involving the smaller group and produces sparse cells in some university-by-background comparisons; equality-of-variance checks were applied, but this imbalance remains a structural constraint on the precision of group comparisons. Fifth, formal post-hoc multiple-comparison tests for the one-way ANOVA and formal simple-effects tests for the significant interaction were not conducted, so institution-specific and background-by-institution claims remain descriptive rather than confirmatory. Sixth, the three-cluster solution was not accompanied by a formal internal-validation statistic, and its robustness should be confirmed in future replications. Finally, this study analyzed a single national program in a single country; extensions of the findings to other teacher-education systems or country contexts are offered as tentative implications rather than tested conclusions.

CONCLUSION

This study mapped the performance of physics graduates with educational and non-educational backgrounds in Indonesia's Teacher Professional Education Program (TPEP). The findings revealed no significant differences between the two groups in PCK, UKIN, and UTUL scores, indicating that this study did not detect a difference between the groups rather than proving that academic background alone does not determine graduates' success in professional teacher training; this null result should be interpreted with the substantial group-size imbalance (661 vs. 67) in mind, as it limits statistical power to detect true differences involving the smaller group. However, institutional mapping indicated significant variation in outcomes across universities, and the interaction analysis showed that the performance gap between backgrounds was not uniform but differed by institution. Regression and ANCOVA results further demonstrated that year of graduation and UTUL scores were the strongest statistical predictors of PCK, while background was not a significant factor; UKIN was also a statistically significant predictor, with a negative coefficient that warrants further investigation rather than a settled explanation (see Discussion). Cluster analysis identified three distinct graduate profiles based on PCK and UTUL scores low, moderate, and high performers underscoring the heterogeneity of participants' competencies. Overall, these results are consistent with the interpretation that institutional context, academic knowledge, and recency of study play a more prominent statistical role than academic background in association with PCK outcomes, within the limits of this cross-sectional, ex post facto design, which does not permit causal claims. The findings highlight the need for differentiated support in TPEP, with targeted interventions for low performers, capacity-building for moderate performers, and advanced professional development for high performers, offered here as evidence-informed implications rather than tested interventions. At the policy level, strengthening institutional quality and aligning academic preparation with professional competencies are crucial steps toward improving teacher education and certification in physics TPEP specifically, with extension to teacher certification more broadly requiring further, subject-diverse research in Indonesia.

ACKNOWLEDGMENT

The authors would like to express their gratitude to the Directorate General of Higher Education, Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia, for providing access to the Teacher Professional Education (TPEP) data. Special thanks are also extended to the participating universities (LPTKs) for their collaboration and support in facilitating this study.

DECLARATION

Author Contribution	:	DAK: Conceptualization, Methodology, Data curation, Formal analysis, Investigation, Visualization, Writing original draft, and Writing review & editing. FN: Methodology, Validation, Formal analysis, Writing review & editing, and Supervision. MSA: Data curation, Software, Validation, Visualization, Writing review & editing, and Project administration. EK: Conceptualization, Supervision, Resources, Project administration, Writing review & editing, and Final approval of the manuscript.
Funding Statement	:	This research was funded by the Director General of Strengthening Research and Development with the Ministry of Research, Technology, and Higher Education of the Republic of Indonesia, for supporting and funding this research.
Conflict of Interest	:	The authors declare no conflict of interest.
Additional Information	:	Additional information is available for this paper.
AI Declaration Statement	:	The authors used AI solely to assist with language editing, grammar refinement, and manuscript readability. All scientific content, data analysis, interpretation of results, and conclusions were developed, verified, and approved by the authors.

REFERENCES

Aartun, I., Walseth, K., Standal, Ø. F., & Kirk, D. (2022). Pedagogies of embodiment in physical education – a literature review. *Sport, Education and Society*, *27*(1), 1–13. https://doi.org/10.1080/13573322.2020.1821182

Abidin, Z., Hindriana, A., & Arip, A. (2023). Pedagogic content knowledge (PCK) in teacher competence development. *Community Empowerment*. https://doi.org/10.31603/ce.9058

Agustiani, M. (2024). Teacher professional education program in indonesia: benefit and challenge. *Enrich: Jurnal Pendidikan, Bahasa, Sastra Dan Linguistik*, *5*(2), 28–37. https://doi.org/10.36546/enrich.v5i2.1314

Amalia, T., & Hodiyanto, H. (2026). Deep Learning Research Trends in Mathematics Learning: A Scoping Review And Bibliometric Analysis (ScoRBA). International Journal of Review in Mathematics Education, 1(3), 218–239. https://doi.org/10.23917/ijrime.18504

Angeli, C., & Valanides, N. (2009). Epistemological and methodological issues for the conceptualization, development, and assessment of ICT–TPCK: Advances in technological pedagogical content knowledge (TPCK). *Computers & Education*, *52*(1), 154–168. https://doi.org/10.1016/j.compedu.2008.07.006

Aydin-Gunbatar, S., Ekiz-Kiran, B., & Oztay, E. S. (2020). Pre-service chemistry teachers' pedagogical content knowledge for integrated STEM development with LESMeR model. *Chemistry Education Research and Practice*, *21*(4), 1063–1082. https://doi.org/10.1039/D0RP00074D

Basikin, B. (2024). The Contribution of Inservice Teacher Education Program (TPEP) on Teachers' Professionalism. *English Language Teaching Educational Journal*. https://doi.org/10.12928/eltej.v6i2.10076

Bloomfield, B. S., Fox, R. A., & Leif, E. S. (2024). Multi-Tiered Systems of Educator Professional Development: A Systematic Literature Review of Responsive, Tiered Professional Development Models. Journal of Positive Behavior Interventions, 26(3), 168–188. https://doi.org/10.1177/10983007231224028

Bostock, J. (2019). Exploring in-service trainee teacher expertise and practice: Developing pedagogical content knowledge. *Innovations in Education and Teaching International*, *56*(5), 605–616. https://doi.org/10.1080/14703297.2018.1562358

Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, *138*, 107468. https://doi.org/10.1016/j.chb.2022.107468

Chang, C. F., Annisa, N., & Chen, K. Z. (2024). Pre-service teacher professional education program (TPEP) and Indonesian science teachers' TPACK development: A career-path comparative study. *Education and Information Technologies*, *30*(7), 8689–8711. https://doi.org/10.1007/S10639-024-13160-6/TABLES/7

Darling-Hammond, L. (2023). Reprint: How Teacher Education Matters. *Journal of Teacher Education*, *74*, 151–156. https://doi.org/10.1177/00224871231161863

Estapa, A., & Davis, J. (2023). Prospective Teachers' Instructional Decisions and Pedagogical Moves when Responding to Student Thinking in Elementary Mathematics and Science Lessons. *International Journal of Science and Mathematics Education*, *21*(5), 1703–1724. https://doi.org/10.1007/S10763-022-10304-3/TABLES/6

Evens, M., Elen, J., Larmuseau, C., & Depaepe, F. (2018). Promoting the development of teacher professional knowledge: Integrating content and pedagogy in teacher education. *Teaching and Teacher Education*, *75*, 244–258. https://doi.org/10.1016/j.tate.2018.07.001

Favier, T., Van Gorp, B., Cyvin, J. B., & Cyvin, J. (2021). Learning to teach climate change: students in teacher training and their progression in pedagogical content knowledge. *Journal of Geography in Higher Education*, *45*(4), 594–620. https://doi.org/10.1080/03098265.2021.1900080

Hubbard, A. (2018). Pedagogical content knowledge in computing education: a review of the research literature. *Computer Science Education*, *28*(2), 117–135. https://doi.org/10.1080/08993408.2018.1509580

Ishartono, N., Putri, R. O. E., Fenyvesi, K., & Malonisio, M. O. (2026). A Systematic Critical Review: Flipped Learning and Its Impact on Self-Efficacy in Mathematics Education. International Journal of Review in Mathematics Education, 1(3), 205–217. https://doi.org/10.23917/ijrime.17530

Iskandar, I., Jumadi, J., Sastradika, D., & Defrianti, D. (2021). Development of TPACK and EQ-based 21st century learning through the teacher certification programme in Indonesia. *South African Journal of Education*, *41*(Supplement 2), S1–S9. https://doi.org/10.15700/saje.v41ns2a1952

Julia, J., Subarjah, H., Maulana, M., Sujana, A., Isrok'atun, I., Nugraha, D., & Rachmatin, D. (2020). Readiness and Competence of New Teachers for Career as Professional Teachers in Primary Schools. *European Journal of Educational Research*, *9*, 655–673. https://doi.org/10.12973/eu-jer.9.2.655

Kamenarac, O. (2024). Why shouldn't teacher education produce 'ready-to-teach' graduates? Mapping a pathway to alternative teacher education and teacher becomings. *Higher Education Research & Development*, *44*, 431–445. https://doi.org/10.1080/07294360.2024.2399068
Kertesz, J., & Brett, P. (2020). Defining and designing impact consciousness in teacher education. *Teaching Education*, *31*, 363–380. https://doi.org/10.1080/10476210.2019.1583728
Komarudin, K., & Suherman, S. (2024). Evaluación del conocimiento tecnológico pedagógico del contenido (TPACK) entre los profesores en formación: modelo de medición Rasch. *Pixel-Bit, Revista de Medios y Educación*, (71), 59–82. https://doi.org/10.12795/pixelbit.107599
Krepf, M., Plöger, W., Scholl, D., & Seifert, A. (2018). Pedagogical content knowledge of experts and novices—what knowledge do they activate when analyzing science lessons? *Journal of Research in Science Teaching*, *55*, 44–67. https://doi.org/10.1002/TEA.21410
Lachner, A., Fabian, A., Franke, U., Preiß, J., Jacob, L., Führer, C., Küchler, U., Paravicini, W., Randler, C., & Thomas, P. (2021). Fostering pre-service teachers' technological pedagogical content knowledge (TPACK): A quasi-experimental field study. *Computers & Education*, *174*, 104304. https://doi.org/10.1016/j.compedu.2021.104304
Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Frontiers in Psychology, 4, 863. https://doi.org/10.3389/fpsyg.2013.00863
Leonet, O., Cenoz, J., & Gorter, D. (2020). Developing morphological awareness across languages: translanguaging pedagogies in third language acquisition. *Language Awareness*, *29*(1), 41–59. https://doi.org/10.1080/09658416.2019.1688338
Li, J., & Copur-Gencturk, Y. (2024). Learning through teaching: the development of pedagogical content knowledge among novice mathematics teachers. *Journal of Education for Teaching*, *50*(4), 582–597. https://doi.org/10.1080/02607476.2024.2358041
Loughran, J. (2020). Pedagogical Content Knowledge. *Science Education for Australian Students*. https://doi.org/10.4324/9781003117223-11
Mahler, D., Bock, D., Schauber, S., & Harms, U. (2024). Using longitudinal models to describe preservice science teachers' development of content knowledge and pedagogical content knowledge. *Teaching and Teacher Education*, *144*, 104583. https://doi.org/10.1016/j.tate.2024.104583
Mardhiah, M., Musgamy, A., & Lubis, M. (2023). Teacher Professional Development through the Teacher Education Program (TPEP) at Islamic Education Institutions. *International Journal of Learning, Teaching and Educational Research*, *22*(11), 80–95. https://doi.org/10.26803/ijlter.22.11.5
Mayer, P., & Girwidz, R. (2019). Physics Teachers' Acceptance of Multimedia Applications—Adaptation of the Technology Acceptance Model to Investigate the Influence of TPACK on Physics Teachers' Acceptance Behavior of Multimedia Applications. *Frontiers in Education*, *4*. https://doi.org/10.3389/feduc.2019.00073
Mishra, P., & Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. *Teachers College Record: The Voice of Scholarship in Education*, *108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x
Musa, H., Rusli, R., & Jufri, H. (2024). Perceptions of Mathematics Education Professional Teacher Education Program (TPEP) Students towards the TPEP Program. *ARRUS Journal of Mathematics and Applied Science*. https://doi.org/10.35877/mathscience2763
Nganga, L., & Kambutu, J. (2024). Culturally Responsive Professional Development Programs for Teacher Educators Using Community-Based Collaborative Learning: Lessons Learned from a Native American Community. *Education Sciences*. https://doi.org/10.3390/educsci14070787
Nind, M. (2020). A new application for the concept of pedagogical content knowledge: teaching advanced social science research methods. *Oxford Review of Education*, *46*(2), 185–201. https://doi.org/10.1080/03054985.2019.1644996
Ningsih, S., Turmudi, & Juandi, D. (2020). Pedagogical content knowledge (PCK) profile of prospective teachers in mathematics learning. *Journal of Physics: Conference Series*, *1521*. https://doi.org/10.1088/1742-6596/1521/3/032057

Nkundabakura, P., Nsengimana, T., Uwamariya, E., Nyirahabimana, P., Nkurunziza, J. B., Mukamwambali, C., Dushimimana, J. C., Batamuliza, J., Byukusenge, C., & Iyamuremye, A. (2024). Effectiveness of the continuous professional development training on upper primary mathematics and science and elementary technology teachers' Pedagogical Content Knowledge in Rwanda. *Discover Education*, *3*(1), 12. https://doi.org/10.1007/s44217-024-00091-0

Nsingo, E., & Chirume, S. (2026). ARIMA Forecasting of Ordinary Level Mathematics Pass Rates: A 13-Year Critical Review from a Zimbabwean High School. International Journal of Review in Mathematics Education, 1(3), 194–204. https://doi.org/10.23917/ijrime.15830

Ong, Q. K. L., & Annamalai, N. (2024). Technological pedagogical content knowledge for twenty-first century learning skills: the game changer for teachers of industrial revolution 5.0. *Education and Information Technologies*, *29*(2), 1939–1980. https://doi.org/10.1007/s10639-023-11852-z

Palincsar, A., DellaVecchia, G., & Easley, K. M. (2020). Teacher Education and its Effects on Teaching and Learning. In *Oxford Research Encyclopedia of Education*. Oxford University Press. https://doi.org/10.1093/acrefore/9780190264093.013.905

Parkhouse, H., Lu, C., & Massaro, V. (2019). Multicultural Education Professional Development: A Review of the Literature. *Review of Educational Research*, *89*, 416–458. https://doi.org/10.3102/0034654319840359

Pawan, F., Li, S., Nijiati, S., Dopwell, M., Harris, A., & Iruoje, T. (2023). Culturally- and Linguistically-Responsive Online Teacher Learning Professional Development. *Online Learning*. https://doi.org/10.24059/olj.v27i4.4003

Roca-Campos, E., Renta-Davids, A.-I., Marhuenda-Fluixá, F., & Flecha, R. (2021). Educational Impact Evaluation of Professional Development of In-Service Teachers: The Case of the Dialogic Pedagogical Gatherings at Valencia "On Giants' Shoulders". *Sustainability*, *13*, 4275. https://doi.org/10.3390/SU13084275

Ruifeng, L., & Ping, D. (2021). A Thorough Review of Pedagogical Content Knowledge (PCK) in TOBACCO framework. *Tobacco Regulatory Science*. https://doi.org/10.18001/trs.7.5.16

Sari, W. R., Komarudin, K., Badrujaman, A., & Tola, B. (2023). Systematic Review of The Impact Evaluation of Teacher Profession Education Programmes in Indonesia. *International Journal of Business, Law, and Education*. https://doi.org/10.56442/ijble.v4i2.186

Sarkar, M., Gutierrez-Bucheli, L., Yip, S. Y., Lazarus, M., Wright, C., White, P. J., Ilic, D., Hiscox, T. J., & Berry, A. (2024). Pedagogical content knowledge (PCK) in higher education: A systematic scoping review. *Teaching and Teacher Education*, *144*, 104608. https://doi.org/10.1016/J.TATE.2024.104608

Schiering, D., Sorge, S., Tröbst, S., & Neumann, K. (2023). Course quality in higher education teacher training: What matters for pre-service physics teachers' content knowledge development? *Studies in Educational Evaluation*, *78*, 101275. https://doi.org/10.1016/J.STUEDUC.2023.101275

Schmidt, D. A., Baran, E., Thompson, A. D., Mishra, P., Koehler, M. J., & Shin, T. S. (2009). Technological Pedagogical Content Knowledge (TPACK). *Journal of Research on Technology in Education*, *42*(2), 123–149. https://doi.org/10.1080/15391523.2009.10782544

She, J., Chan, K. K. H., Wang, J., Hu, X., & Liu, E. (2025). Effect of Science Teachers' Pedagogical Content Knowledge on Student Achievement: Evidence From Both Text- and Video-Based Pedagogical Content Knowledge Tests. *American Educational Research Journal*, *62*(1), 92–135. https://doi.org/10.3102/00028312241278627

Shing, C. L., Saat, R., & Loke, S. H. (2015). The Knowledge of Teaching--Pedagogical Content Knowledge (PCK). *Malaysian Online Journal of Educational Sciences*, *3*, 40–55.

Shulman, L. S. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher*, *15*(2), 4–14. https://doi.org/10.3102/0013189X015002004

Simanjorang, M., Mansyur, A., & Gultom, S. (2020). Reconstruction of Professional Teacher Education Program. *Journal of Education and Practice*. https://doi.org/10.7176/jep/11-16-12

Solihin, R., Iqbal, M., & Muin, M. T. (2021). Konstruksi Kompetensi Pedagogik Guru dalam Pembelajaran. *SCAFFOLDING: Jurnal Pendidikan Islam Dan Multikulturalisme*. https://doi.org/10.37680/scaffolding.v3i2.1085

Su, Y. (2023). Delving into EFL teachers' digital literacy and professional identity in the pandemic era: Technological Pedagogical Content Knowledge (TPACK) framework. *Heliyon*, *9*(6), e16361. https://doi.org/10.1016/j.heliyon.2023.e16361

Taşar, M. F., & Yılmaz Ergül, D. (2023). Technological Pedagogical Content Knowledge (TPACK) in Physics Education. In *The International Handbook of Physics Education Research: Teaching Physics* (pp. 1-1-1–30). AIP Publishing LLCMelville, New York. https://doi.org/10.1063/9780735425712_001

Tseng, J.-J., Chai, C. S., Tan, L., & Park, M. (2022). A critical review of research on technological pedagogical and content knowledge (TPACK) in language teaching. *Computer Assisted Language Learning*, *35*(4), 948–971. https://doi.org/10.1080/09588221.2020.1868531

Wang, X., Hamat, A. Bin, & Shi, N. L. (2024). Designing a pedagogical framework for mobile-assisted language learning. *Heliyon*, *10*(7), e28102. https://doi.org/10.1016/J.HELIYON.2024.E28102/ASSET/05196F2C-2ED7-4956-B11F-64BA1C2F16F2/MAIN.ASSETS/GR5.JPG

Warahmah, M., Daud, A., & Novitri, N. (2021). OBSTACLES FACED BY PRE-SERVICE ENGLISH TEACHERS DURING TEACHERS PROFESSIONAL EDUCATION PROGRAM (TPEP). *International Journal of Educational Best Practices*. https://doi.org/10.31258/ijebp.v5n2.p153-167

Werler, T., & Tahirsylaj, A. (2020). Differences in teacher education programmes and their outcomes across Didaktik and curriculum traditions. *European Journal of Teacher Education*, *45*, 154–172. https://doi.org/10.1080/02619768.2020.1827388

Yuniawatika, Muksar, M., Angga Rini, T., & Alfan, M. (2022). Online Teacher Professional Education Program (TPEP) Practices Through Moodle-Based E-learning. *2nd International Conference on Information Technology and Education (ICIT&E)*, 60–64. https://doi.org/10.1109/ICITE54466.2022.9759894