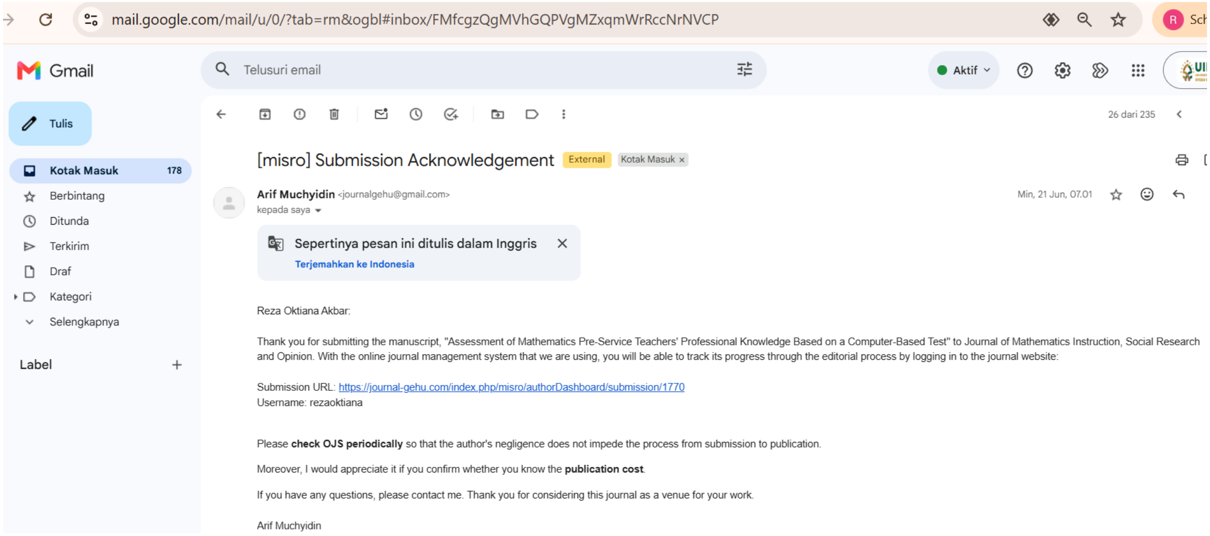
BUKTI KORESPONDENSI

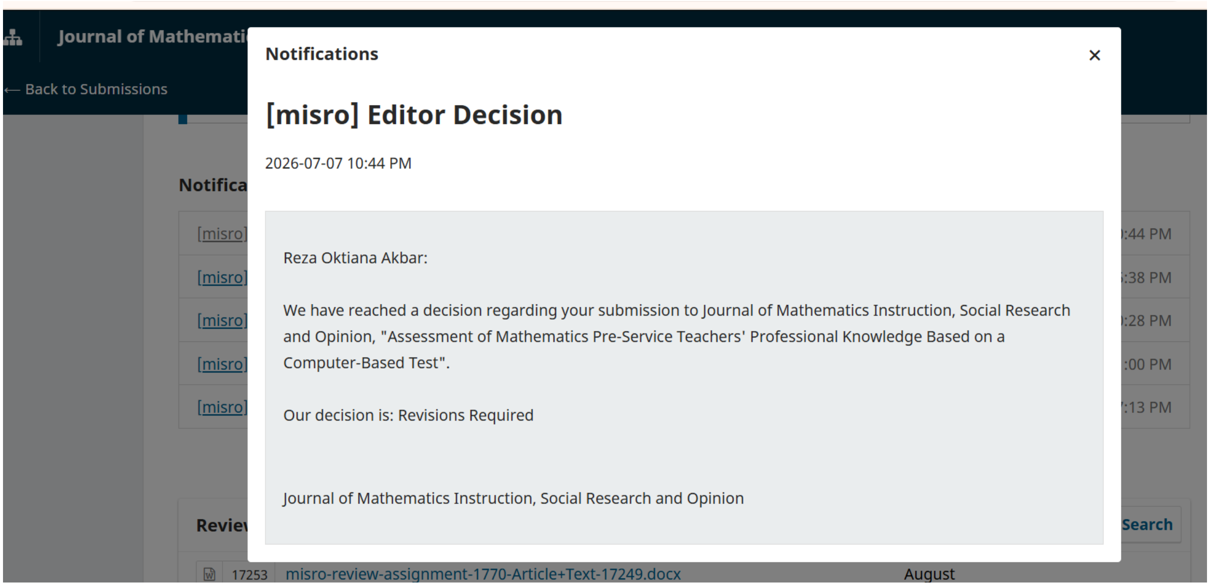
REZA OKTIANA AKBAR, M.Pd.

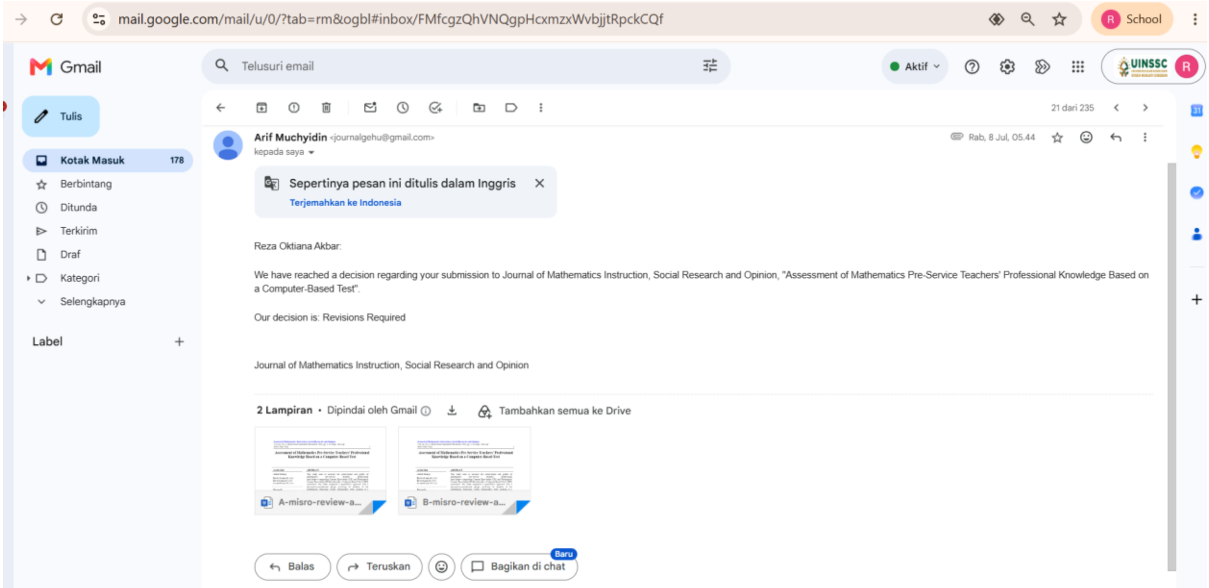
TITLE: Assessing Mathematics Pre-Service Teachers' Content Knowledge and Pedagogical Content Knowledge through a Diagnostic Computer-Based Test

Korespondensi 1 (Submit, 21 Juni 2026)



Korespondensi 2 (Reviewer 1 & Reviewer 2, perbaikan ke-1 dan 2, 8 Juli 2026)





mail.google.com/mail/u/0/?tab=rm&ogbl#inbox/FMfcgzQhVNQgpHcxmzxWvbjjtRpckCQf

Gmail

Telusuri email

Tulis

Kotak Masuk 178

Berbintang

Ditunda

Terkirim

Draf

Kategori

Selengkapnya

Label

21 dari 235

Arif Muchyidin <journalgehu@gmail.com>

kepada saya

Rab, 8 Jul, 05.44

Sepertinya pesan ini ditulis dalam Inggris

Terjemahkan ke Indonesia

Reza Oktiana Akbar:

We have reached a decision regarding your submission to Journal of Mathematics Instruction, Social Research and Opinion, "Assessment of Mathematics Pre-Service Teachers' Professional Knowledge Based on a Computer-Based Test".

Our decision is: Revisions Required

Journal of Mathematics Instruction, Social Research and Opinion

2 Lampiran • Dipindai oleh Gmail

Tambahkan semua ke Drive

A-misro-review-a...

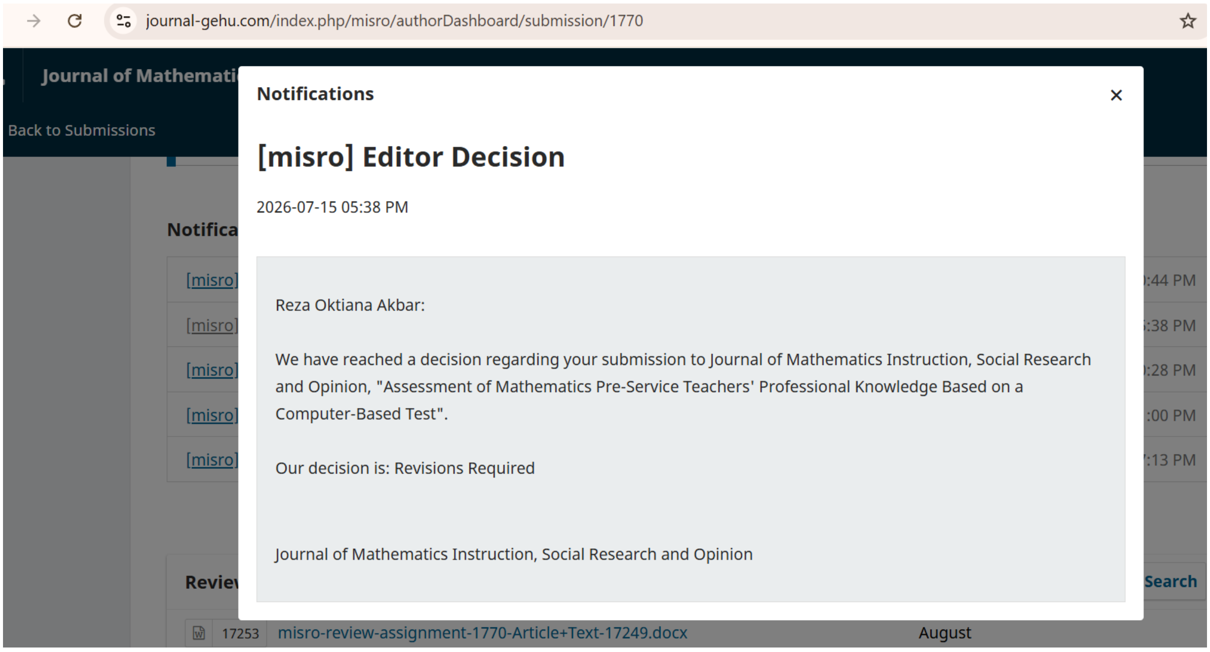
B-misro-review-a...

Balas

Teruskan

Bagikan di chat

Korespondensi 3 (Reviewer 3, perbaikan ke-3, 15 Juli 2026)



journal-gehu.com/index.php/misro/authorDashboard/submission/1770

Journal of Mathemati

Back to Submissions

Notifications

[misro] Editor Decision

2026-07-15 05:38 PM

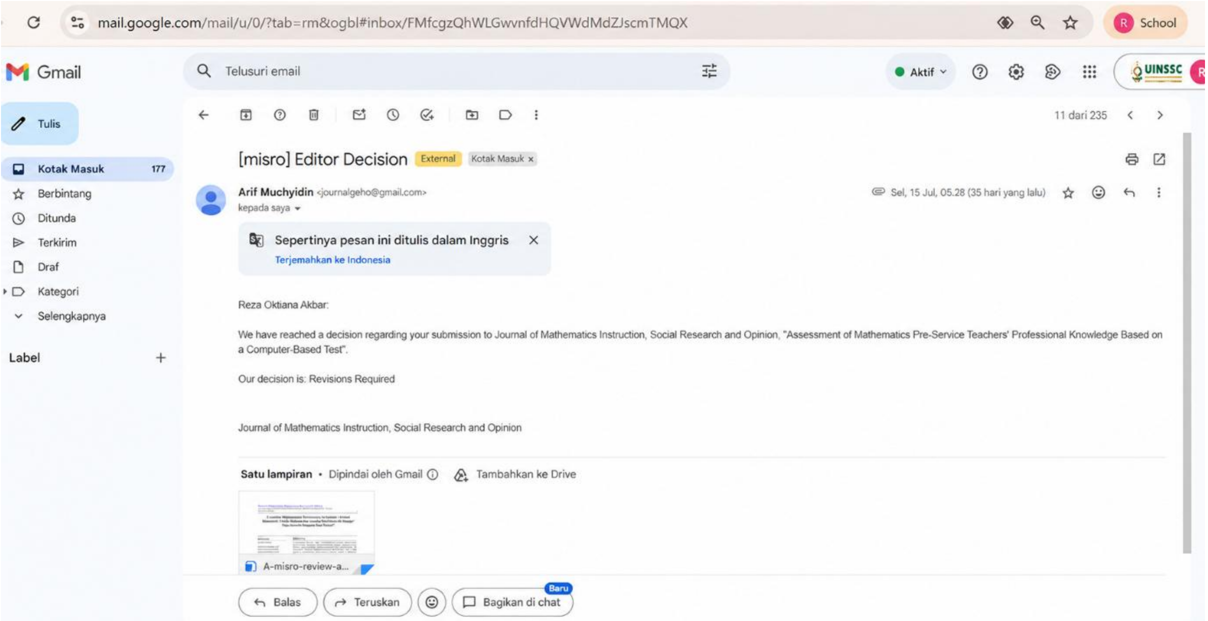
Reza Oktiana Akbar:

We have reached a decision regarding your submission to Journal of Mathematics Instruction, Social Research and Opinion, "Assessment of Mathematics Pre-Service Teachers' Professional Knowledge Based on a Computer-Based Test".

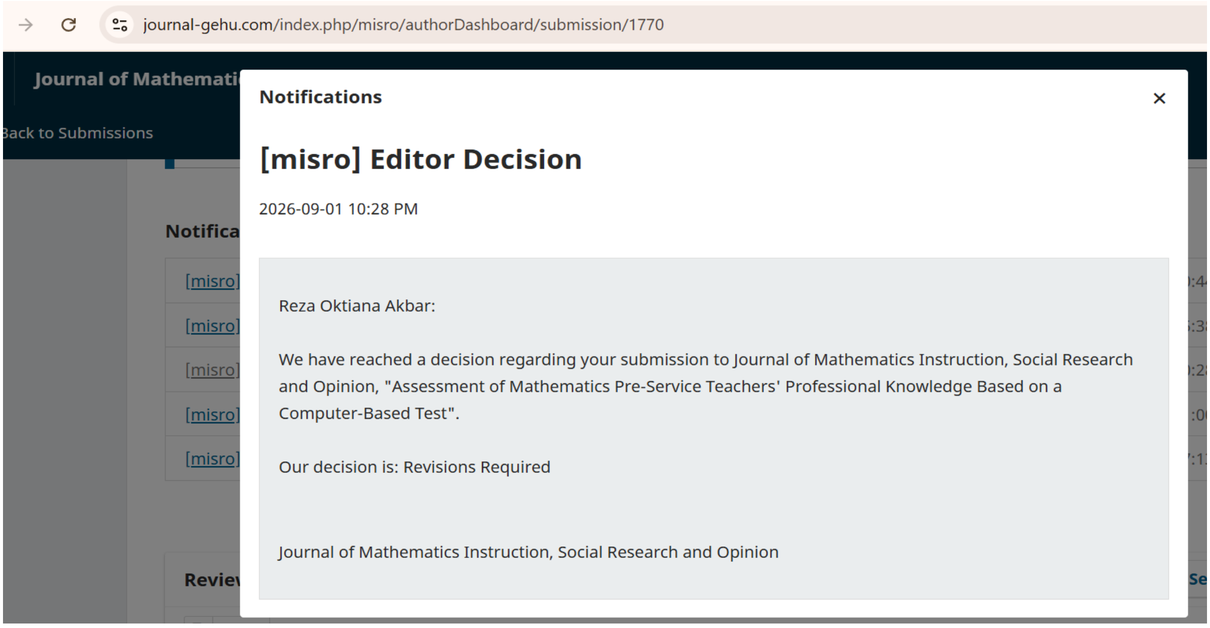
Our decision is: Revisions Required

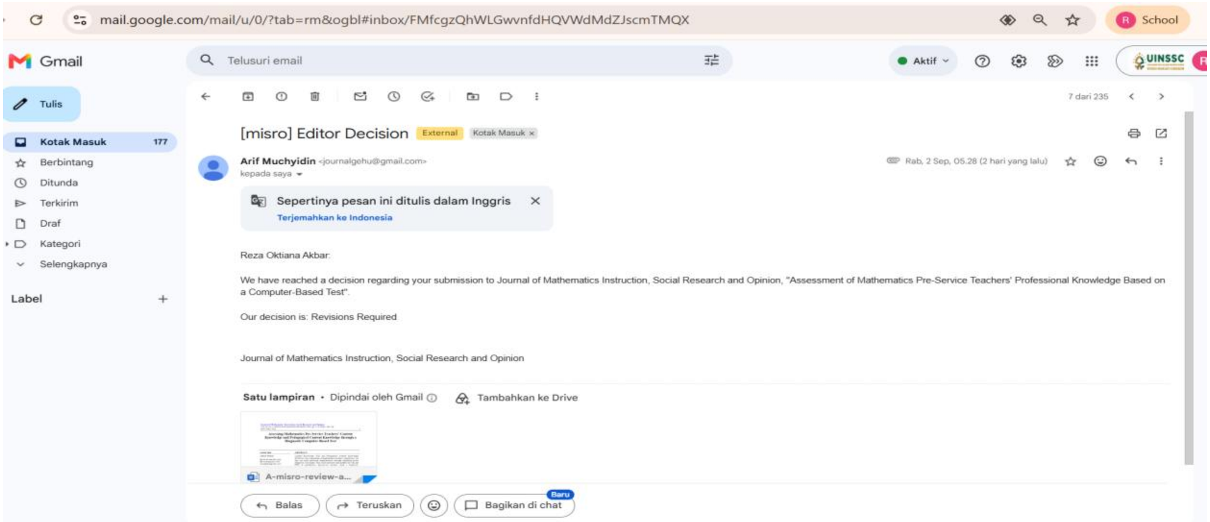
Journal of Mathematics Instruction, Social Research and Opinion

17253 misro-review-assignment-1770-Article+Text-17249.docx August

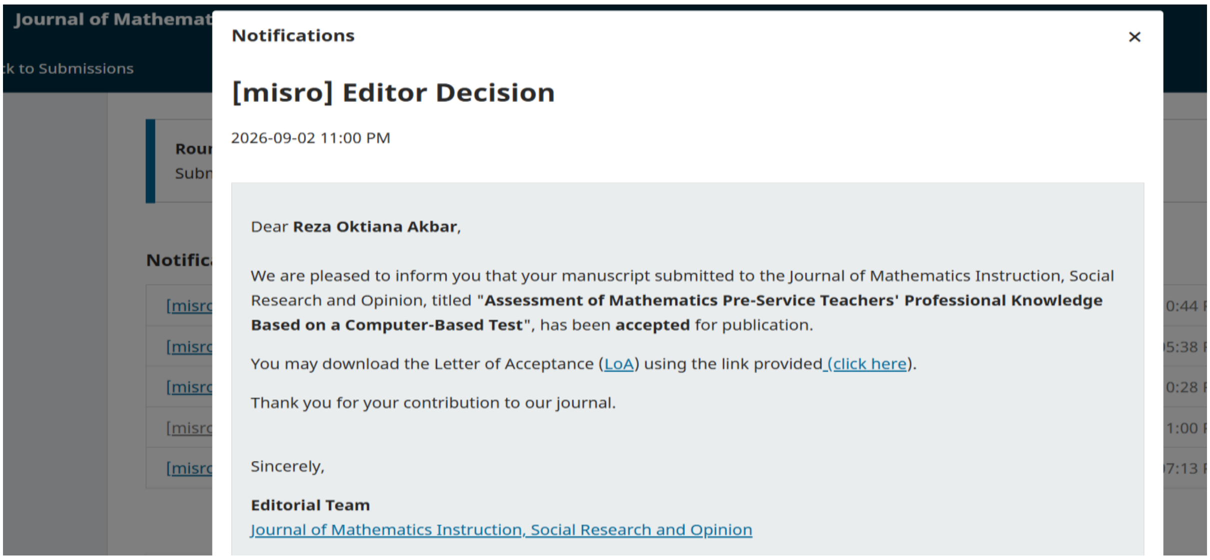


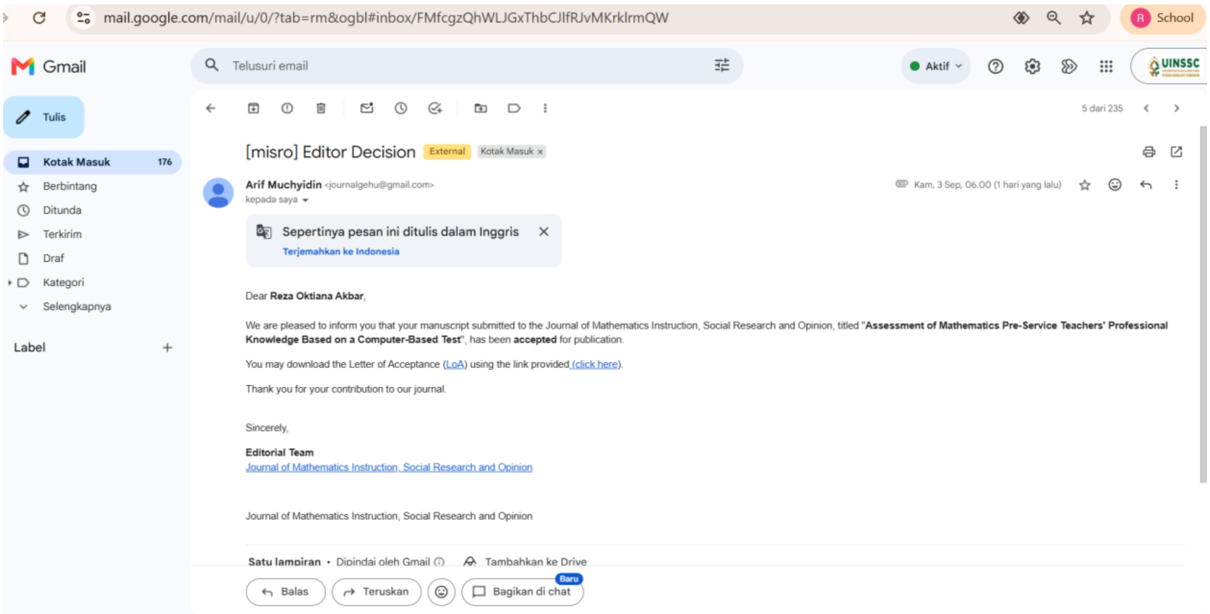
Korespondensi 4 (Reviewer 4, perbaikan ke-4, 2 September 2026)





Korespondensi 6 (Accepted)





ISSN 2962-7842

MISRO Journal of Mathematics Instruction, Social Research and Opinion

LETTER OF ACCEPTANCE

Dear Reza Oktiana Akbar, Muhamad Ali Misri

It is my pleasure to inform you that, after the peer review of your paper entitled:

"Assessing Mathematics Pre-Service Teachers' Content Knowledge and Pedagogical Content Knowledge through a Diagnostic Computer-Based Test"

It has been **ACCEPTED** to be published regularly in the Journal of Mathematics Instruction, Social Research and Opinion (MISRO) Vol. 5 No. 3, September 2026.

Thank you very much for submitting your article to MISRO. We look forward to submitting your other articles to our journal.

Best Regards,
Bandung, September 2, 2026
Editor-in-chief

Arif Muchyidin
| **Scopus** | **GS** | **SINTA** | **ORCID** |

Korespondensi 7 (copy editing discussion, 3 September 2026)

Arif Muchyidin <journalgehu@gmail.com> 3 Sep 2026, 06.00 (1 hari yang lalu)
kepada saya

Sepertinya pesan ini ditulis dalam Inggris
Terjemahkan ke Indonesia

You have a new notification from Journal of Mathematics Instruction, Social Research and Opinion:

You have been added to a discussion titled "Copyediting" regarding the submission "Assessment of Mathematics Pre-Service Teachers' Professional Knowledge Based on a Computer-Based Test".

Balas Teruskan Bagikan di chat Baru

Journal of Mathematics Instruction, Social Research and Opinion

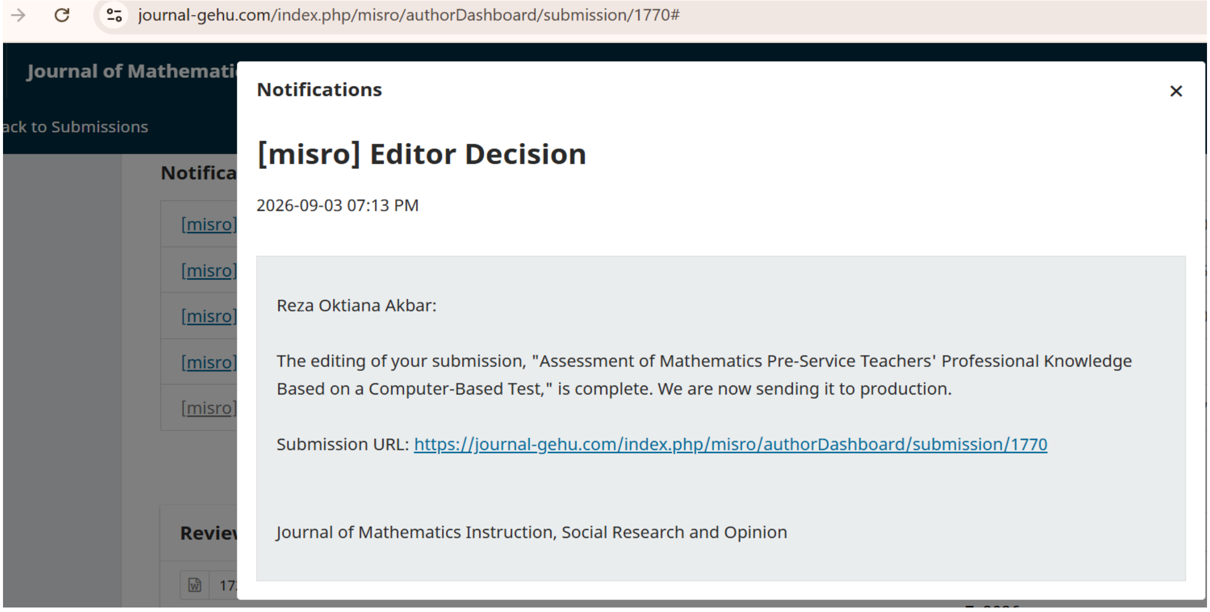
k to Submissions

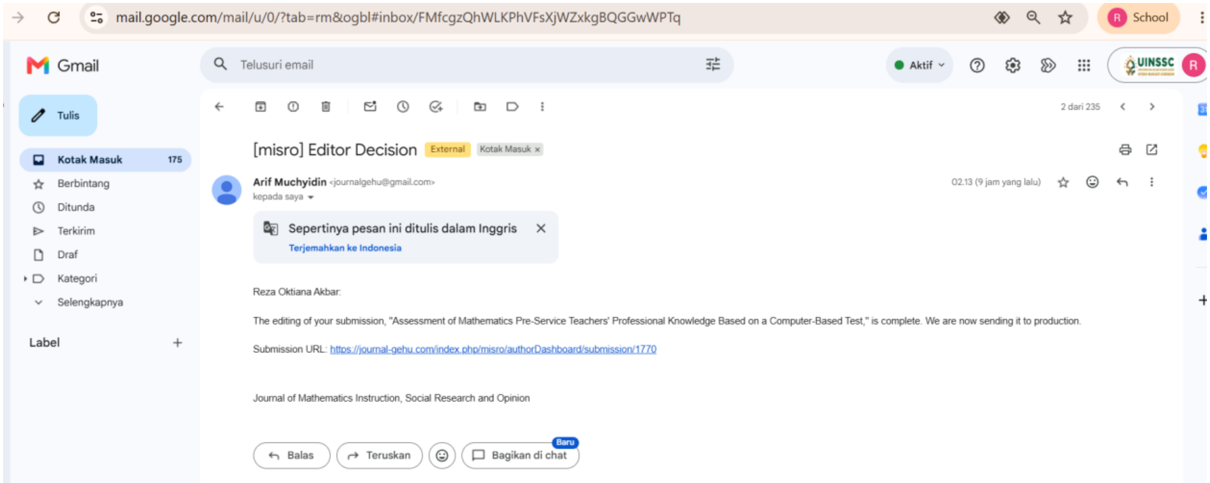
Submission Review Copyediting Production

Copyediting Discussions Add discussion

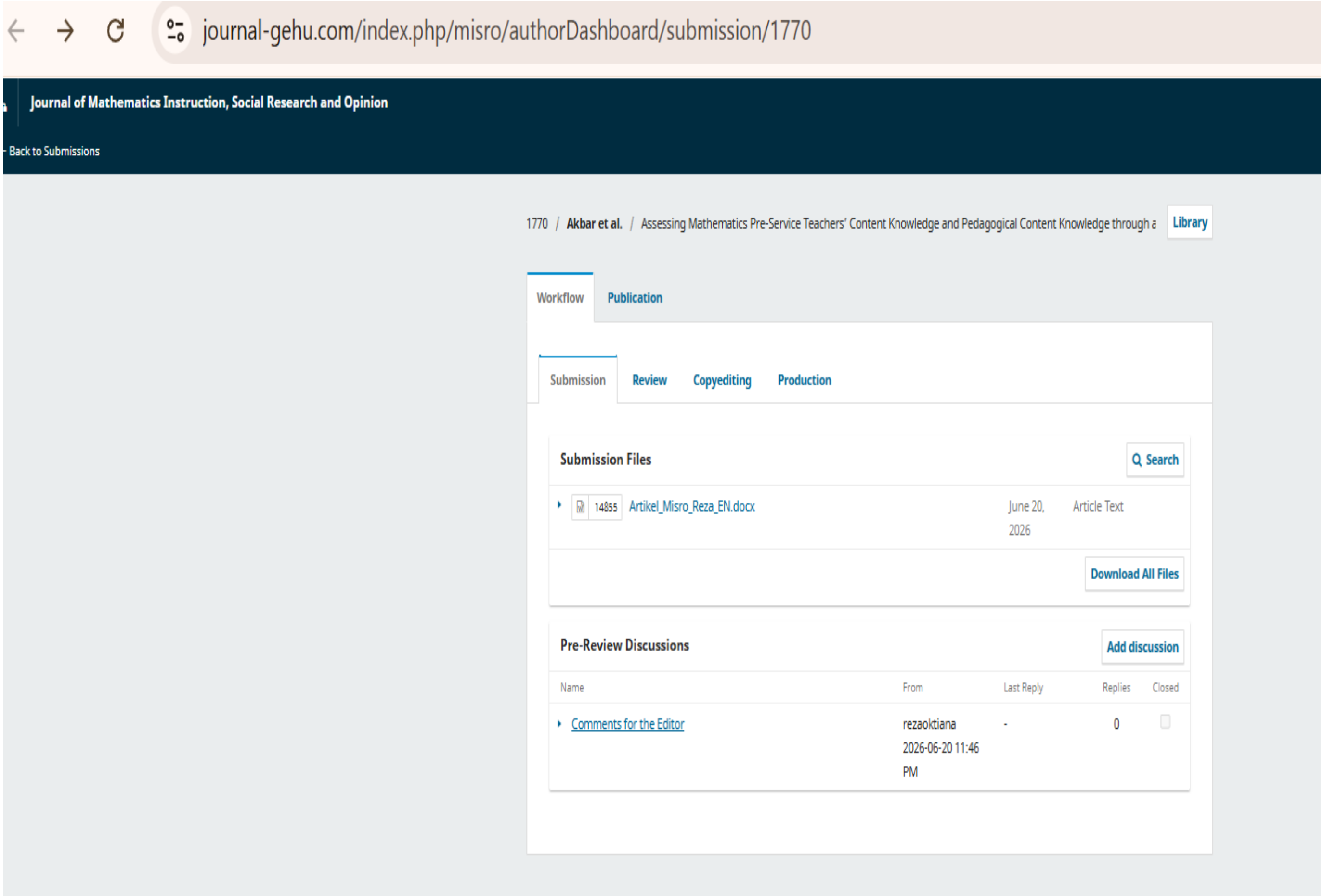
Name	From	Last Reply	Replies	Closed
Copyediting	arifmuchyidin 2026-09-02 11:00 PM	-	0	
fiksasi artikel tahap akhir	rezaoktiana 2026-09-03 01:11 PM	-	0	

Korespondensi 8 (Produksi/Publish)





Bukti Riwayat Submission di OJS



Bukti Riwayat Review Round 1 di laman OJS

1770 / **Akbar et al.** / Assessing Mathematics Pre-Service Teachers' Content Knowledge and Pedagogical Content Knowledge through a Library

Workflow Publication

Submission Review Copyediting Production

Round 1 Round 2 Round 3

Round 1 Status

New reviews have been submitted and are being considered by the editor.

Notifications

[misro] Editor Decision	2026-07-07 10:44 PM
[misro] Editor Decision	2026-07-15 05:38 PM
[misro] Editor Decision	2026-09-01 10:28 PM
[misro] Editor Decision	2026-09-02 11:00 PM
[misro] Editor Decision	2026-09-03 07:13 PM

Reviewer's Attachments Search

15415	misro-review-assignment-1770-Article+Text-15407.docx	June 29, 2026
15881	misro-review-assignment-1770-Article+Text-15407.docx	July 7, 2026

Revisions Search Upload File

15899	Artikel_Misro_Revisi1.docx	July 8, 2026	Article Text

Review Discussions Add discussion

Name	From	Last Reply	Replies	Closed
LoA	arifmuchyidin 2026-09-02 10:56 PM	-	0	

Bukti Riwayat Review Round 2 di laman OJS

1770 / **Akbar et al.** / Assessing Mathematics Pre-Service Teachers' Content Knowledge and Pedagogical Content Knowledge through a | Library

Workflow | Publication

Submission | Review | Copyediting | Production

Round 1 | Round 2 | Round 3

Round 2 Status

All reviewers have responded and a decision is needed.

Notifications

[misro] Editor Decision	2026-07-07 10:44 PM
[misro] Editor Decision	2026-07-15 05:38 PM
[misro] Editor Decision	2026-09-01 10:28 PM
[misro] Editor Decision	2026-09-02 11:00 PM
[misro] Editor Decision	2026-09-03 07:13 PM

Reviewer's Attachments | Search

16152	misro-review-assignment-1770-Article+Text-16049.docx	July 11, 2026

Revisions | Search | Upload File

16386	Artikel_Misro_Revisi2.docx	July 17, 2026	Article Text

Review Discussions | Add discussion

Name	From	Last Reply	Replies	Closed
LoA	arifmuchyidin 2026-09-02 10:56 PM	-	0	

Bukti Riwayat Review Round 3 di laman OJS

1770 / **Akbar et al.** / Assessing Mathematics Pre-Service Teachers' Content Knowledge and Pedagogical Content Knowledge through a Library

Workflow | Publication

Submission | Review | Copyediting | Production

Round 1 | Round 2 | Round 3

Round 3 Status
Submission accepted.

Notifications

[misro] Editor Decision	2026-07-07 10:44 PM
[misro] Editor Decision	2026-07-15 05:38 PM
[misro] Editor Decision	2026-09-01 10:28 PM
[misro] Editor Decision	2026-09-02 11:00 PM
[misro] Editor Decision	2026-09-03 07:13 PM

Reviewer's Attachments Search

17253	misro-review-assignment-1770-Article+Text-17249.docx	August 7, 2026

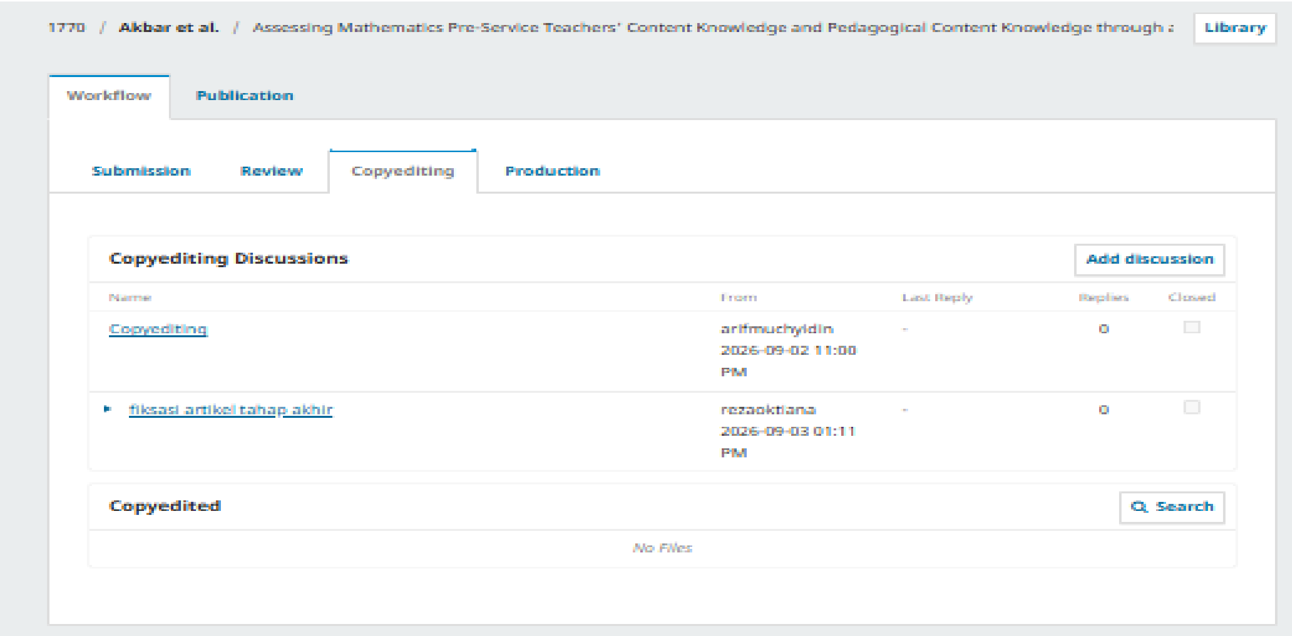
Revisions Search Upload File

18237	Artikel_Misro_Revisi 3.docx	September 1, 2026	Article Text

Review Discussions Add discussion

Name	From	Last Reply	Replies	Closed
LoA	arifmuchyidin 2026-09-02 10:56 PM	-	0	

Bukti Riwayat Copy Editing di OJS



Bukti Review 1 oleh Reviewer I

Journal of Mathematics Instruction, Social Research and Opinion
Vol. xx, No. x, March/June/September/December 202x, pp. 1-1x, https://doi.org/
ISSN 2962-7842

Assessment of Mathematics Pre-Service Teachers' Professional Knowledge Based on a Computer-Based Test

Commented [AM1]: Recommended revision: **"Assessing Mathematics Pre-Service Teachers' Content Knowledge and Pedagogical Content Knowledge through a Diagnostic Computer-Based Test"**

Article Info

Article history:

Received mm dd, yyyy
Revised mm dd, yyyy
Accepted mm dd, yyyy

Keywords:

First keyword
Second keyword
Third keyword
Fourth keyword
Fifth keyword

ABSTRACT

This study aims to measure the achievement and profile of mathematics pre-service teachers' professional knowledge—comprising Content Knowledge (CK) and Pedagogical Content Knowledge (PCK)—through a Computer-Based Test (CBT) instrument. The study employed a quantitative approach with a descriptive-correlational design involving 30 students of the Mathematics Education (Tadris Matematika) study program at a teacher education institution (LPTK) selected purposively. Data were analyzed using descriptive statistics, Pearson correlation, and simple linear regression. The results show mean scores of 50.00% for CK and 52.67% for PCK, both below the ideal competency threshold (75%), with 60.00% of the students in the low category and the Knowledge of Content and Curriculum (KCC) dimension as the weakest (46.25%). A strong positive correlation was found between CK and PCK ($r = 0.72$; $p < 0.001$), and CK predicted 52% of the variance in PCK ($PCK = 1.708 \times CK - 32.734$). Diagnostic analysis at the dimension, indicator, and individual-student levels demonstrates that the CBT is able to generate specific and actionable achievement profiles as a basis for instructional intervention. These findings underscore the need to strengthen LPTK curricula that integrate content mastery and pedagogical capacity grounded in diagnostic data.

Commented [AM2]: The abstract is informative and contains the research aim, method, sample, data analysis techniques, major findings, and implications.
The abstract clearly reports important quantitative findings, including CK mean score, PCK mean score, the percentage of students in the low category, the weakest dimension, correlation, and regression equation.
However, the abstract is quite dense and could be made more readable by reducing technical details or dividing long sentences.
The keywords are still placeholders: **"First keyword, Second keyword, Third keyword, Fourth keyword, Fifth keyword."**
These must be replaced.
Recommended keywords:
Computer-Based Test; Content Knowledge; Pedagogical Content Knowledge; Mathematics Pre-Service Teachers; Diagnostic Assessment; Mathematical Knowledge for Teaching.
The abstract states that the CBT is able to generate actionable diagnostic profiles, but it should briefly mention whether the CBT instrument had been validated or tested for reliability.
The phrase **"ideal competency threshold (75%)"** should be clarified. The abstract should state whether this threshold is based on institutional standards, previous studies, or researcher-defined criteria

Corresponding Author:

INTRODUCTION

The quality of a nation's mathematics education depends heavily on teachers' professional competence. Cross-national studies have shown that mathematics teachers' professional knowledge—comprising Content Knowledge (CK) and Pedagogical Content Knowledge (PCK)—is a primary predictor of instructional effectiveness and students' academic achievement [1], [2]. The 2022 PISA results placed Indonesian students' mathematics literacy 68th out of 81 countries, below the OECD average, indicating a fundamental problem in the quality of mathematics instruction [3]. This condition is inseparable from the quality of professional knowledge formed in teacher education institutions (LPTK), so ensuring an adequate professional-knowledge profile of pre-service teachers becomes a strategic imperative [4], [5].

Commented [AM3]: The introduction provides a strong rationale by linking mathematics education quality, teacher professional competence, CK, PCK, and Indonesian students' mathematics literacy.
The use of the Mathematical Knowledge for Teaching framework is appropriate and relevant to the research focus.
The article also successfully connects the study with the growing importance of digital assessment and Computer-Based Testing.
The research gap is generally clear: CBT has not been widely used to simultaneously measure CK and PCK while producing layered diagnostic feedback.
However, the introduction could be more focused. Some parts discuss teacher competence, CBT, digital transformation, and policy urgency at once, which may make the argument feel too broad.
The novelty is clearly stated in three aspects: theoretical, methodological, and contextual. This is a strength of the article.
However, the novelty claim should be supported more carefully. The author should avoid implying that Indonesian LPTK students are completely overlooked unless the literature review strongly supports this.

Theoretically, the Mathematical Knowledge for Teaching (MKT) framework distinguishes CK—comprising common, specialized, and horizon content knowledge—from PCK, which encompasses knowledge of content and students, content and teaching, and content and curriculum [6], [7]. Systematic reviews affirm that PCK is the most frequently studied yet the most difficult component to measure consistently because of its contextual and multidimensional nature [8], [9]. Studies in Indonesia show that the professional knowledge of mathematics pre-service teachers varies widely across institutions and curricula, with a notable weakness in the PCK aspect [10], [11]. Similar findings have been reported in cross-national contexts [12], [13].

Alongside the rapid digitalization of education, competency-evaluation systems are now shifting massively from traditional approaches toward technology-based assessment. Various studies indicate that CBT is often more valid and reliable than conventional examination methods [14], [15]; however, its use to objectively measure both the content knowledge and the pedagogical content knowledge of pre-service teachers simultaneously remains very limited because of the complexity of these constructs [16], [17]. The potential of CBT to generate diagnostic feedback based on learning analytics has also not been optimally exploited [18], even though interactive digital feedback has been shown to improve mathematical understanding [19], [20].

The urgency of this study is heightened in the context of Indonesia's uneven higher-education digital transformation [21], which is not yet supported by adequate standardization of assessment systems [22]. Measuring teacher competence in the digital era is required to produce data that are actionable for study-program managers and policymakers [23]. A number of studies affirm that CK and PCK are relatively independent constructs that require different intervention pathways [24], [25], and that teacher effects on student achievement are substantial [26].

The novelty of this study lies in three aspects. **First**, theoretically, the study integrates the MKT framework and technology-based assessment within a single coherent quantitative measurement design to identify CK and PCK profiles simultaneously [27], [28]. **Second**, methodologically, the CBT is used as an autonomous, automated assessment medium that generates layered diagnostic feedback—at the total, dimension, indicator, and individual-student levels—so that test results can be acted upon directly for instructional improvement [29]. **Third**, contextually, the study targets LPTK students in Indonesia who have so far been overlooked in in-depth empirical studies [30]. This study aims to: (1) measure students' CK and PCK achievement and profiles; (2) identify strengths and weaknesses in each dimension and indicator; and (3) demonstrate the use of diagnostic profiles—including at the individual-student level—as a basis for designing instructional interventions.

METHOD

This study employed a quantitative approach with a descriptive-survey and correlational design. The population comprised students of the Mathematics Education (Tadris Matematika) study program at an LPTK who were in their fifth and seventh semesters. A sample of 30 students was determined through purposive sampling with the inclusion criteria: (1) active students who had adequately completed mathematics-content and pedagogy courses; and (2) willing to take the entire series of CBT tests. Given the

Commented [AM4]: The method is appropriate for the study because the article uses a quantitative descriptive-correlational design.
The sample is clearly described: 30 fifth- and seventh-semester Mathematics Education students selected through purposive sampling.
The inclusion criteria are also clearly stated, which helps readers understand why the participants were selected.
The operationalization of CK and PCK is useful, especially because the article maps CK into CCK, SCK, and HCK, and PCK into KCS, KCT, and KCC.
Table 1 is helpful because it presents the number of CBT items for each MKT dimension.
However, the method needs stronger explanation of the CBT instrument development process.
The article should explain how the CBT items were constructed, reviewed, validated, and revised.
The article should report psychometric evidence, such as item validity, reliability, difficulty index, discrimination index, or expert validation results.
Since the study uses multiple-choice items to measure complex constructs such as PCK, the author should justify why multiple-choice CBT is sufficient for assessing pedagogical reasoning.
The sample size is quite small for correlation and regression analysis.
The article should acknowledge that n = 30 limits the generalizability and statistical power of the findings.
The use of Microsoft Excel is acceptable for basic analysis, but the article should explain how assumptions were checked, especially for Pearson correlation and linear regression.
The article should report whether normality, linearity, outliers, and homoscedasticity were examined.
The categorization into **High, Medium, and Low** should be explained clearly. The cut-off points should be stated in the method section.
Ethical considerations are briefly addressed through participant confidentiality, but the article should also mention informed consent and voluntary participation.

limited sample size, all data were analyzed as a single intact group (n = 30) and were not directed toward inter-semester comparison.

The operationalization of the variables comprised two main constructs. CK was defined as mastery of the mathematics content to be taught, measured through 25 multiple-choice items; PCK was defined as the ability to integrate content knowledge with pedagogical strategies, also measured through 25 multiple-choice items. Scores were converted into percentages of achievement relative to the ideal maximum score. The CBT instrument was designed to generate layered diagnostic reports, namely the total score, achievement per dimension, achievement per indicator, and an individual achievement profile for each student. Each item was mapped to a specific indicator and dimension (see the notes below Table 1) so that the CBT system could produce achievement per dimension and per indicator for all students at once, as well as their individual diagnostic profiles. The composition of the instrument is presented in Table 1.

Table 1. Composition of CBT Instrument Items by MKT Dimension

Domain	Dimension	Number of Items
CK	Common Content Knowledge (CCK)	7
CK	Specialized Content Knowledge (SCK)	14
CK	Horizon Content Knowledge (HCK)	4
PCK	Knowledge of Content and Students (KCS)	10
PCK	Knowledge of Content and Teaching (KCT)	9
PCK	Knowledge of Content and Curriculum (KCC)	6
Total		50

Notes on dimension codes and indicator coverage (based on the developed CK and PCK instruments). **CK – CCK** (Common Content Knowledge, items 1–7): understanding basic concepts (e.g., lines and angles), performing calculations and measurements accurately, and applying mathematical concepts in everyday life. **CK – SCK** (Specialized Content Knowledge, items 8–21): explaining concepts and principles, analyzing and connecting mathematical objects, proving properties, building models, and solving complex mathematical problems. **CK – HCK** (Horizon Content Knowledge, items 22–25): connecting school mathematics with advanced academic mathematics and with other disciplines. **PCK – KCS** (Knowledge of Content and Students, items 1–10): analyzing students' preconceptions and understanding as well as predicting and analyzing students' misconceptions, errors, and difficulties. **PCK – KCT** (Knowledge of Content and Teaching, items 11–19): selecting and organizing learning materials, selecting sources, strategies, and techniques, and selecting and analyzing representations for presenting material. **PCK – KCC** (Knowledge of Content and Curriculum, items 20–25): designing content and topic combinations, applying standards and curricula, and designing instruments to measure students' abilities.

The analysis techniques comprised descriptive and inferential analyses. Descriptive analysis produced the mean (M), standard deviation (SD), minimum, maximum, median, and percentage of achievement at the total, dimension, indicator, and individual levels, as well as a three-level categorization (High, Medium, Low). Inferential analysis was conducted through two techniques: (1) Pearson correlation to test the CK–PCK relationship and the corresponding inter-dimension relationships (CCK–KCS, SCK–KCT, HCK–KCC) at $\alpha = 0.05$; and (2) simple linear regression with CK as the predictor and PCK as the criterion.

All scoring, achievement computation, correlation, and linear regression were performed using Microsoft Excel through standardized formulas (including CORREL, SLOPE, INTERCEPT, and RSQ) so that every figure can be traced back to the raw data.

Participants' identities were kept confidential by using codes instead of names throughout the recording, processing, and reporting of data; in this article, the individual-profile example is presented under the pseudonym Student 1.

RESULTS AND DISCUSSION

3.1. Results

Overall, students' professional-knowledge achievement was in the medium category and had not reached the ideal competency threshold (≥ 75%). The mean CK score was 50.00% (SD = 10.16), PCK 52.67% (SD = 23.96), and the combined Professional Knowledge (PK) score 51.33% (SD = 16.04). Complete descriptive statistics are presented in Table 2.

Table 2. Descriptive Statistics of CK, PCK, and Professional Knowledge Scores (n = 30)

Variable	Mean (%)	SD	Min	Max	Median
Content Knowledge (CK)	50.00	10.16	24	68	52.00
Pedagogical Content Knowledge (PCK)	52.67	23.96	8	96	54.00
Professional Knowledge (PK)	51.33	16.04	26	78	50.00

Table 2 shows that the mean PCK is slightly higher than CK (a difference of 2.67 points), an interesting pattern indicating that pedagogical-transformation capacity slightly exceeds content mastery. The large standard deviation of PCK (SD = 23.96) and the much smaller one for CK (SD = 10.16) reflect substantial heterogeneity in knowledge profiles, with score ranges of 24–68% for CK and 8–96% for PCK. The distribution of Professional Knowledge categories is presented in Table 3.

Table 3. Distribution of Professional Knowledge Categories (n = 30)

Category	Frequency (f)	Percentage (%)
High	0	0.00
Medium	12	40.00
Low	18	60.00
Total	30	100.00

All students (100%) were in the Low or Medium category, with 60.00% in the Low category and 40.00% in the Medium category, and no student reaching the High category. This indicates that the majority of pre-service teachers have not attained an optimal level of professional-knowledge mastery. Achievement per dimension is presented in Table 4 and Figure 1.

Table 4. Average Achievement per CK and PCK Dimension

Domain	Dimension	Items	Mean (%)	SD
CK	CCK	7	73.33	20.47
CK	SCK	14	36.91	24.16
CK	HCK	4	55.00	35.60
PCK	KCS	10	54.07	33.18
PCK	KCT	9	57.50	26.14

Commented [AM5]: The results section is rich and well-organized. It presents descriptive statistics, category distribution, dimension-level achievement, indicator-level achievement, correlation, regression, and individual diagnostic profiles.
The article clearly reports that the mean CK score is **50.00%**, PCK is **52.67%**, and overall Professional Knowledge is **51.33%**.
The finding that **60% of students are in the low category** and none are in the high category is important and supports the urgency of instructional intervention.
Table 4 and Figure 1 are useful because they show the strengths and weaknesses across CK and PCK dimensions.
The finding that **CCK is the strongest dimension** while **SCK and KCC are the weakest** is meaningful for curriculum improvement.
Table 5 and Figure 2 provide a more detailed diagnostic picture at the indicator level. This is one of the strongest parts of the article.
The article's use of an individual example, **Student 1**, is also valuable because it demonstrates how CBT can generate personalized diagnostic profiles.
However, the results section is too long and sometimes overly interpretive. Some explanations would fit better in the discussion section.
Figures 5 and 6 are interesting as authentic CBT screenshots, but they are difficult to read. They should be redesigned or converted into clearer tables or higher-resolution figures.
Since Figure 5 and Figure 6 are screenshots from the CBT system, the article should ensure that no identifiable student information is visible. The pseudonym "Student 1" is appropriate.
The article should avoid saying that the CBT can "hardly" be provided by conventional tests. A more academic phrasing would be: **"The CBT provides diagnostic functions that are more difficult to implement efficiently in conventional paper-based tests."**
The regression result is useful, but the interpretation should be more cautious. The finding that CK explains around 52% of PCK variance does not prove causal influence.
The article should avoid causal wording such as **"CK predicted"** unless it clearly frames the regression as predictive association, not causal explanation.

Domain	Dimension	Items	Mean (%)	SD
PCK	KCC	6	46.25	32.30

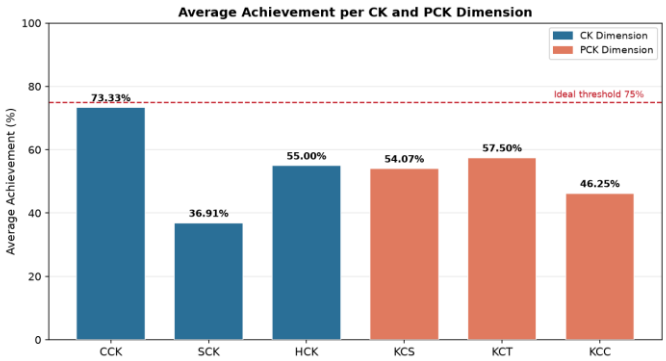


Figure 1. Average Achievement per CK and PCK Dimension

In the CK domain, the CCK dimension was the highest (73.33%) and SCK the lowest (36.91%); in the PCK domain, KCT was the highest (57.50%) while KCC was the lowest of all six dimensions (46.25%). Diagnostically, the low SCK and KCC signal a weak ability to relate content to the structure and organization of the curriculum—an area that needs to be prioritized in intervention. Achievement at the indicator level is presented in Table 5 and visualized in Figure 2.

Table 5. Average Achievement per CK and PCK Indicator

Domain	Indicator	Mean (%)
CK	CCK1	86.67
CK	CCK2	80.00
CK	CCK3	46.67
CK	SCK1	63.33
CK	SCK2	46.67
CK	SCK3	43.33
CK	SCK4	20.00
CK	SCK5	23.33
CK	SCK6	30.00
CK	SCK7	33.33
CK	HCK1	53.33
CK	HCK2	56.67
PCK	KCS1	58.33
PCK	KCS2	48.33
PCK	KCS3	58.33
PCK	KCS4	51.67
PCK	KCS5	53.33
PCK	KCT1	55.00
PCK	KCT2	66.67
PCK	KCT3	60.00
PCK	KCT4	36.67
PCK	KCC1	51.67
PCK	KCC2	43.33
PCK	KCC3	45.00

Notes on indicator codes (based on the CK and PCK instruments). CK indicators — **CCK1**: understanding basic school-mathematics concepts, such as lines, angles, sequences, and basic statistics (items 1–3); **CCK2**: performing calculations and measurements accurately (items 4–5); **CCK3**: applying mathematical concepts in everyday life (items 6–7); **SCK1**: explaining academic mathematical concepts and principles (items 8–9); **SCK2**: analyzing mathematical objects (items

10–11); **SCK3**: connecting mathematical properties and objects (items 12–13); **SCK4**: proving mathematical properties or theorems (items 14–16); **SCK5**: applying proof strategies (item 17); **SCK6**: building and constructing mathematical models (items 18–19); **SCK7**: solving complex mathematical problems (items 20–21); **HCK1**: connecting school mathematics with advanced academic mathematics (items 22–23); **HCK2**: connecting mathematics with other disciplines (items 24–25). PCK indicators — **KCS1**: analyzing students' preconceptions (items 1–2); **KCS2**: identifying students' level of understanding (items 3–4); **KCS3**: predicting students' misconceptions (items 5–6); **KCS4**: identifying sources of students' learning difficulties (items 7–8); **KCS5**: analyzing students' literacy issues and errors (items 9–10); **KCT1**: organizing and sequencing learning materials (items 11–12); **KCT2**: selecting learning sources, strategies, and techniques (items 13–14); **KCT3**: selecting representations and illustrations for presenting material (items 15–17); **KCT4**: analyzing the weaknesses and strengths of representations (items 18–19); **KCC1**: composing content and topic combinations (items 20–21); **KCC2**: applying standards and curricula (items 22–23); **KCC3**: designing instruments to measure students' abilities (items 24–25).

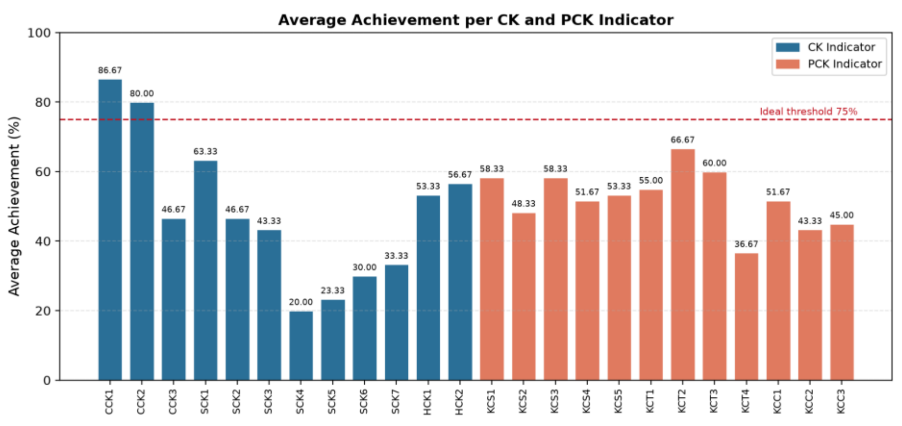


Figure 2. Average Achievement per CK and PCK Indicator

Figure 2 visualizes the average achievement of each indicator exactly as listed in Table 5. The range of achievement across indicators spans 20.00% to 86.67%. The highest achievement is on the CK indicator CCK1 (86.67%; understanding basic school-mathematics concepts), which is fundamental and therefore the most widely mastered by students. Conversely, the lowest overall achievement is on indicator SCK4 (20.00%; proving mathematical properties or theorems)—the skill with the highest cognitive demand in the CK domain—followed by SCK5 (23.33%; applying proof strategies) and SCK6 (30.00%; building and constructing mathematical models). The pattern in the SCK dimension declines gradually with complexity, from explaining concepts and principles (SCK1; 63.33%), analyzing and connecting mathematical objects (SCK2 46.67%; SCK3 43.33%), solving complex problems (SCK7; 33.33%), to the most difficult modeling and proving (SCK6 30.00%; SCK5 23.33%; SCK4 20.00%)—a pattern consistent with cognitive load theory. Because each dimension's achievement remains the mean of its constituent indicators—for example, KCS (54.07%) is the mean of KCS1–KCS5 and CCK (73.33%) the mean of CCK1–CCK3—this pattern maps precisely which items and indicators are the sources of weakness, so that lecturers can design remediation targeted at concrete rather than generic weaknesses. The results of the Pearson correlation test are presented in Table 6.

Table 6. Pearson Correlation between Domains and between Parallel Dimensions

Relationship	r	p	Interpretation

CK – PCK	0.72	< 0.001	Strong
CCK – KCS	0.42	< 0.05	Moderate
SCK – KCT	0.54	< 0.01	Moderate
HCK – KCC	0.86	< 0.001	Very strong

The strong positive correlation between CK and PCK ($r = 0.72$; $p < 0.001$) confirms that the two domains are constructively interdependent, reinforced by the correlations between parallel dimensions, especially HCK–KCC ($r = 0.86$) and SCK–KCT ($r = 0.54$). The predictive relationship of CK on PCK was tested through simple linear regression (Table 7 and Figure 3).

Table 7. Results of the Simple Linear Regression of CK on PCK

Parameter	Value
Slope (b)	1.708
Intercept (a)	−32.734
Coefficient of determination (R^2)	0.5245
Pearson r	0.72
Equation	PCK = 1.708 × CK − 32.734

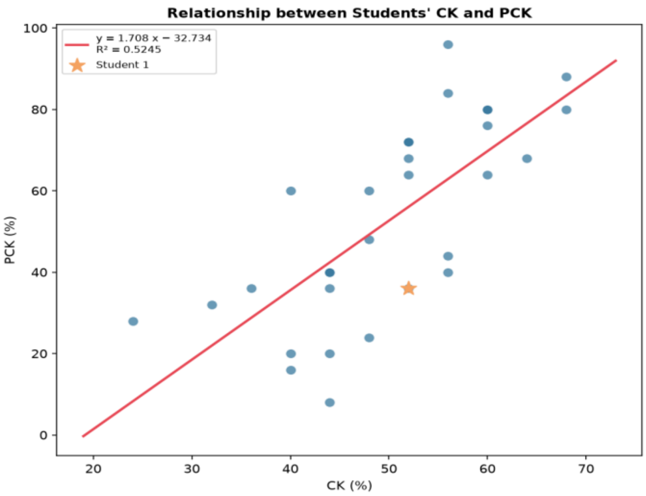


Figure 3. Scatter Plot and Regression Line of the CK–PCK Relationship

The equation PCK = 1.708 × CK − 32.734 with $R^2 = 0.52$ indicates that about 52% of the variance in PCK can be explained by CK. The remaining 48% of the variance signals that PCK does not form automatically from content mastery but requires an explicit pedagogical learning pathway.

The main advantage of the CBT instrument lies in its ability to break down aggregate achievement into individual diagnostic profiles automatically. As an illustration, Table 8 and Figure 4 summarize the achievement of one student disguised as Student 1, while Figure 5 and Figure 6 display authentic screenshots of the diagnostic scoring report generated directly by the CBT system for Student 1. Such individual reports are produced automatically by the system for every participant without any manual scoring.

Table 8. Individual Achievement Profile of Student 1 Compared with the Class Average

Domain	Student 1 Score (%)	Class Average (%)	Difference	Category
Content Knowledge (CK)	52.00	50.00	+2.00	Low
Pedagogical Content Knowledge (PCK)	36.00	52.67	−16.67	Low

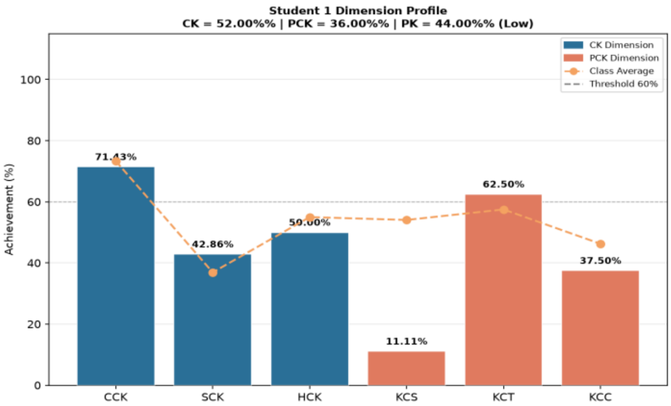


Figure 4. CK and PCK Achievement Profile of Student 1 Compared with the Class Average

At the individual level, the CBT system automatically scores the answers and generates a diagnostic scorecard for each participant. Student 1, for example, answered 13 of 25 CK items correctly (52.00%) and 9 of 25 PCK items correctly (36.00%); the CK achievement was slightly above the class average (a difference of +2.00 points) while the PCK achievement was far below it (a difference of −16.67 points), and neither reached the ideal competency threshold (75%), placing the student in the Low category for both domains. Authentic screenshots of the report are presented in Figure 5 (CK) and Figure 6 (PCK).

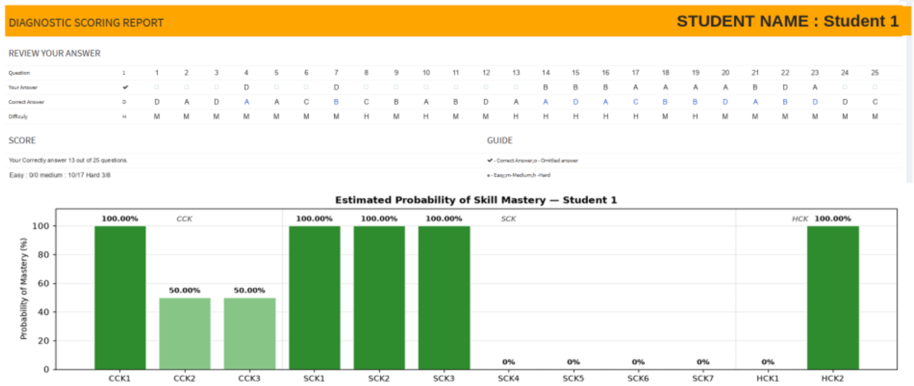


Figure 5. CBT Diagnostic Scorecard for the CK Domain for Student 1 (original system screenshot)

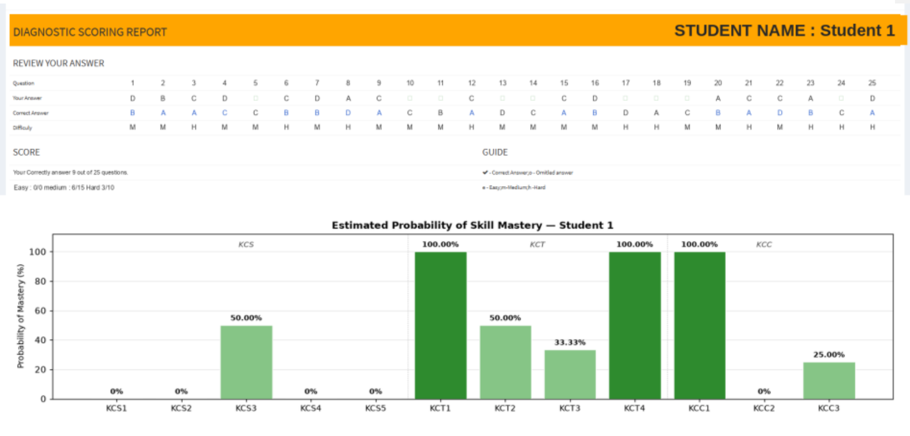


Figure 6. CBT Diagnostic Scorecard for the PCK Domain for Student 1 (system screenshot)

Figures 5 and 6 reveal a role of the CBT that conventional tests can hardly provide. In addition to recapping correct/incorrect answers per item along with their difficulty level—a combination of the Medium (M) and Hard (H) categories according to the item analysis of the instrument—the system presents an Estimated Probability of Skill Mastery graph that maps Student 1's mastery probability for each skill indicator (CCK1–HCK2 for CK and KCS1–KCC3 for PCK), complete with a mapping of sample items to each indicator. Thus, the CBT does not merely provide a final score but directly identifies which indicators have been mastered and which remain weak—in Student 1's case, several indicators show a mastery probability close to zero—so that lecturers can follow up with specific and immediate remediation. It is this automation of scoring, standardization of reporting, and diagnostic depth that affirms the added value of the CBT in assessing the professional knowledge of mathematics pre-service teachers.

Furthermore, the two scorecards allow Student 1's achievement to be traced down to the indicator level, not stopping at dimension averages. In the CK domain (Figure 5), Common Content Knowledge (CCK) was the strongest dimension (71.43%; 5 of 7 items): understanding basic school-mathematics concepts (CCK1) was fully mastered (100.00%), whereas performing calculations and measurements (CCK2) and applying mathematical concepts in everyday life (CCK3) were each only partially mastered (50.00%). Specialized Content Knowledge (SCK) stood at 42.86% (6 of 14 items) with a highly contrasting pattern: three basic indicators—explaining academic mathematical concepts and principles (SCK1), analyzing mathematical objects (SCK2), and connecting mathematical properties and objects (SCK3)—were all fully mastered (100.00%), while four indicators demanding higher-order reasoning—proving mathematical properties or theorems (SCK4), applying proof strategies (SCK5), building and constructing mathematical models (SCK6), and solving complex mathematical problems (SCK7)—were all at 0.00%, marking a total deficit in advanced mathematical reasoning. Horizon Content Knowledge (HCK) stood at 50.00% (2 of 4 items): connecting mathematics with other disciplines (HCK2) was fully mastered (100.00%), but connecting school mathematics with advanced academic mathematics (HCK1) remained at 0.00%. In the PCK domain (Figure 6), the indicator profile was also uneven within each dimension. Knowledge of Content and Teaching

(KCT; dimension mean 62.50%) was the strongest: organizing and sequencing learning materials (KCT1) and analyzing the weaknesses and strengths of representations (KCT4) were fully mastered (100.00%), followed by selecting learning sources, strategies, and techniques (KCT2; 50.00%) and selecting representations and illustrations for presenting material (KCT3; 33.33%). Knowledge of Content and Curriculum (KCC; dimension mean 37.50%) rested on composing content and topic combinations (KCC1), which was fully mastered (100.00%), while designing instruments to measure students' abilities (KCC3) was only 25.00% and applying standards and curricula (KCC2) was still 0.00%. Conversely, Knowledge of Content and Students (KCS) was the weakest PCK dimension and the weakest point of the entire profile (11.11%): only predicting students' misconceptions (KCS3) was partially mastered (50.00%), while analyzing students' preconceptions (KCS1), identifying students' level of understanding (KCS2), identifying sources of students' learning difficulties (KCS4), and analyzing students' literacy issues and errors (KCS5) were all at 0.00%. This pattern confirms that Student 1's weaknesses are concentrated in a number of indicators that are genuinely untouched—SCK4, SCK5, SCK6, SCK7, and HCK1 in the CK domain and KCS1, KCS2, KCS4, KCS5, and KCC2 in the PCK domain, all at 0.00%—so that these ten indicators are the most urgent remediation priorities, while the indicators already fully mastered (CCK1, SCK1, SCK2, SCK3, HCK2, KCT1, KCT4, and KCC1) can serve as a foundation for reinforcement. Thus, indicator-based diagnosis provides a far sharper and more specific direction for individual intervention than reading dimension averages alone.

3.2. Discussion

The main findings reveal three key facts. **First**, neither CK (50.00%) nor PCK (52.67%) reached the minimum competency threshold; interestingly, PCK was even slightly higher than CK, reflecting the knowledge-transformation challenge articulated by Shulman [7] and reinforced in the MKT framework [6]. **Second**, the SCK dimension was the weakest point of CK (36.91%) and KCC the weakest point of PCK (46.25%), indicating a double deficit in the mastery of advanced mathematical content and curriculum, which in the COACTIV model strongly determines the quality of instructional planning [2], [25]. **Third**, the lower standard deviation of CK (10.16) compared with PCK (23.96) shows that CK achievement is more uniform, whereas PCK is highly heterogeneous, indicating large variation in students' pedagogical capacity.

The significant positive correlations, especially HCK–KCC ($r = 0.86$) and SCK–KCT ($r = 0.54$), show that content mastery in particular dimensions goes hand in hand with the relevant pedagogical capacity. However, the regression equation ($R^2 = 0.52$) affirms that content mastery explains only part of the variance in PCK, consistent with findings that pre-service teachers in developing countries tend to have lower PCK scores than CK because LPTK curricula place greater emphasis on content mastery [1], [10]. The case of Student 1 reinforces this at the individual level: Student 1 (CK = 52%, PCK = 36%) falls into the Low category with a highly uneven profile, namely CK achievement slightly above the class average while PCK is far below it. In the PCK domain, KCT is relatively the highest (62.50%) while KCS is the lowest (11.11%) and KCC is low (37.50%), indicating an ability to select teaching strategies that is not yet accompanied by an adequate understanding of student characteristics or curriculum mastery. Intervention needs to target KCS and KCC simultaneously, especially the ability to analyze

Commented [AM6]: The discussion successfully identifies three key findings: low CK and PCK achievement, weaknesses in SCK and KCC, and the strong relationship between CK and PCK.
The discussion appropriately connects the findings to Shulman's PCK theory, the MKT framework, and the COACTIV model.
The article's claim that CK and PCK are interrelated but asymmetrically developing is conceptually strong.
The discussion of Student 1 is useful because it shows how individual diagnostic data can guide targeted remediation.
However, the discussion should include more comparison with previous empirical studies, especially studies on mathematics pre-service teachers' CK and PCK in Indonesia.
The article should explain why PCK is slightly higher than CK. This is an interesting finding, but it needs deeper interpretation.
The author should also discuss why SCK is particularly weak. For example, students may struggle with proof, mathematical modeling, and higher-order reasoning because teacher education curricula may emphasize procedural content over advanced mathematical reasoning.
The discussion should avoid overgeneralizing from one institution and 30 students.
The article should discuss the limitation of using only multiple-choice items. PCK often requires analysis of classroom scenarios, student work, teaching decisions, and pedagogical explanation, which may not be fully captured through multiple-choice CBT.
The limitations paragraph is useful but should be expanded.
Recommended limitations:
- Small sample size.
- Purposive sampling.
- Single-institution context.
- Limited use of bivariate analysis.
- Absence of qualitative data to explain why students performed poorly.
- Limited evidence of CBT instrument validity and reliability.
- Possible limitations of multiple-choice format for measuring complex professional knowledge.

preconceptions, identify levels of understanding, and identify sources of students' learning difficulties.

Theoretically, this study validates and extends the application of the MKT framework and the COACTIV model to the context of religiously affiliated LPTKs in Indonesia, which are almost unrepresented in the indexed literature [6], [25], [11]. Methodologically, the study demonstrates a CBT model that produces layered diagnostic profiles—from the total, dimension, and indicator levels to the individual student—thus addressing the previously identified gap in using CBT to measure pedagogical competence [22], [28]. Achievement profiles that can be decomposed to the individual level, as shown in the case of Student 1, strengthen the function of the CBT as an instrument that generates actionable diagnostic feedback [18]. In policy terms, the mean PK of 51.33% and the absence of any student in the High category provide urgent empirical evidence to drive data-based curriculum reform rather than normative assumptions [5], [23], particularly in the context of Indonesia's 2022 PISA results [3].

This study opens a new perspective by positioning the CBT not merely as an examination tool but as a technology-based diagnostic-assessment ecosystem that produces rich cognitive data, decomposed per dimension and per individual, and actionable for program managers [23], [30]. Several limitations should be noted, namely the relatively small sample size ($n = 30$) and the use of purposive sampling, which limit generalizability, as well as the still-bivariate scope of the analysis, which has not yet modeled other determinants simultaneously.

CONCLUSION AND RECOMMENDATIONS

Commented [AM7]: The conclusion clearly summarizes the main findings, including CK and PCK scores, category distribution, weakest dimensions, correlation, regression, and the CBT's diagnostic potential.
The recommendations are highly specific and practical. This is a major strength of the article.
The recommendation to strengthen mathematical proof, proof strategies, modeling, and complex problem solving is well-aligned with the weakest CK indicators.
The recommendation to strengthen curriculum understanding, representation analysis, and assessment instrument design is also relevant to the weakest PCK indicators.
However, the conclusion is too long and contains many detailed recommendations that could be moved into a separate **Implications** or **Recommendations** section.
The conclusion should be more concise and should directly answer the research objectives.
The recommendations should be organized into categories, for example:
- **For lecturers**
- **For LPTK curriculum managers**
- **For future researchers**
- **For CBT system development**

The article should add a practical recommendation for improving the CBT itself, such as adding automatic feedback, item-level explanations, or links to remedial learning materials.

This study succeeded in identifying the professional-knowledge profile of mathematics pre-service teachers at an LPTK through a CBT instrument. The mean CK (50.00%) and PCK (52.67%) scores were both below the ideal competency threshold (75%), with students distributed across the Low (60.00%) and Medium (40.00%) categories and none reaching the High category, with CCK as the highest CK dimension (73.33%), SCK the weakest (36.91%), and KCC as the weakest PCK dimension (46.25%). There was a strong positive correlation between CK and PCK ($r = 0.72$; $p < 0.001$), and CK predicted 52% of the variance in PCK ($PCK = 1.708 \times CK - 32.734$). Layered diagnostic analysis—down to the individual level as exemplified in the case of Student 1 (CK 52.00%; PCK 36.00%)—shows that the CBT is able to generate specific and actionable achievement profiles. Overall, these findings reinforce the view that CK and PCK are interrelated yet asymmetrically developing constructs, and affirm the feasibility of the CBT as an efficient and standardized diagnostic-assessment instrument in LPTKs.

Recommendations. In practical terms, the Mathematics Education (Tadris Matematika) study program is recommended to design remedial interventions targeted not only at the weakest dimensions (SCK in CK and KCC in PCK) but also down to the weakest indicators as read in Table 5. In the CK domain, the top priority is strengthening mathematical proof ability—namely the indicators proving properties or theorems (SCK4; 20.00%) and applying proof strategies (SCK5; 23.33%)—through gradually structured proof practice (scaffolded proof), analysis of proof errors, and habituation to writing deductive arguments; followed by strengthening mathematical modeling (SCK6; 30.00%) and complex problem solving (SCK7; 33.33%) through problem-based learning and mathematization projects. In the PCK domain, the focus is directed at the indicators analyzing the weaknesses and strengths of representations (KCT4; 36.67%), applying standards and curricula (KCC2; 43.33%), and designing measurement instruments (KCC3; 45.00%)—through case-based learning, analysis of teaching videos, examination of curriculum documents, and assignments to develop test blueprints and assessment tools. This indicator-targeting strategy can also be applied at the individual level: the profile of Student 1, for example, who falls into the Low category, calls for prioritizing the untouched indicators—the four advanced-reasoning SCK indicators (proving, proof strategies, modeling, and complex problem solving; SCK4–SCK7) and connecting school mathematics with advanced academic mathematics (HCK1) in the CK domain, as well as analyzing students' preconceptions (KCS1), identifying students' level of understanding (KCS2), identifying sources of students' learning difficulties (KCS4), analyzing students' literacy issues and errors (KCS5), and applying standards and curricula (KCC2) in the PCK domain—all of which are still at 0.00%. For LPTK managers, these results encourage the integration of a diagnostic CBT system—capable of mapping achievement down to the indicator and individual levels—into the quality-assurance cycle so that specific weaknesses can be monitored and followed up continuously. Theoretically and methodologically, further research is recommended to use a larger and more representative sample, a longitudinal design, cross-institution comparison, and multivariate models involving other determinants (academic background, learning motivation, and the quality of lecturers' teaching) to strengthen external validity and deepen understanding of the formation of mathematics pre-service teachers' professional knowledge.

ACKNOWLEDGMENTS

The authors thank the Mathematics Education (Tadris Matematika) study program of the LPTK for facilitating the implementation of the research, as well as all the students who participated as respondents.

REFERENCES

[1] X. Yang and G. Kaiser, “The impact of mathematics teachers’ professional competence on instructional quality and students’ mathematical learning outcomes,” Teach. Teach. Educ., vol. 110, art. 103582, 2022, doi: 10.1016/j.tate.2021.103582

[2] S. Krauss et al., “Pedagogical content knowledge and content knowledge of secondary mathematics teachers,” Journal of Educational Psychology, 100(3), 716–725, 2008, doi: 10.1037/0022-0663.100.3.716.

[3] A. Schleicher, PISA 2022 Results: The State of Learning and Equity in Education. Paris: OECD Publishing, 2023.

[4] M. T. Tatto, J. Schwille, S. L. Senk, L. Ingvarson, R. Peck, G. Rowley, K. Bankov, M. Rodriguez, and M. Reckase, “Policy, Practice, and Readiness to Teach Primary and Secondary Mathematics in 17 Countries: Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)”. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement, 2012.

[5] H. Qian and P. Youngs, “The effect of teacher education programs on future elementary mathematics teachers’ knowledge: A five-country analysis using TEDS-M data”. J. Math. Teacher Educ., vol. 19, no. 4, pp. 371–396, 2016, doi: 10.1007/s10857-014-9297-9.

[6] D. L. Ball, M. H. Thames, and G. Phelps, “Content knowledge for teaching: What makes it special?,” J. Teacher Educ., vol. 59, no. 5, pp. 389–407, 2008, doi: 10.1177/0022487108324554.

[7] L. S. Shulman, “Those who understand: Knowledge growth in teaching,” Educ. Res., vol. 15, no. 2, pp. 4–14, 1986, doi: 10.3102/0013189X015002004.

[8] M. Grigaliūnienė, E. Lehtinen, L. Verschaffel, and F. Depaepe, “Systematic review of research on pedagogical content knowledge in mathematics: Insights from a topic-specific approach,” *ZDM Mathematics Education*, vol. 57, pp. 777–794, 2025, doi: 10.1007/s11858-025-01684-1.

[9] F. Depaepe, L. Verschaffel, and G. Kelchtermans, “Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research,” Teach. Teach. Educ., vol. 34, pp. 12–25, 2013, doi: 10.1016/j.tate.2013.03.001.

[10] Ningsih, S. Y., Turmudi, & Juandi, D. (2020). *Pedagogical content knowledge (PCK) profile of prospective teachers in mathematics learning*. Journal of Physics: Conference Series, 1521(3), 032057, doi: 10.1088/1742-6596/1521/3/032057

[11] S. Fiangga, S. Khabibah, S. M. Amin, dan R. Ekawati, "A Learning Design Analysis of the Pre-service Teachers’ Mathematics Pedagogical Content Knowledge," *J. Phys.: Conf. Ser.*, vol. 1899, no. 1, p. 012153, 2021, doi: 10.1088/1742-6596/1899/1/012153.

[12] Yang, X., Schwarz, B. & Leung, I.K.C. Pre-service mathematics teachers’ professional modeling competencies: a comparative study between Germany, Mainland China, and Hong Kong. *Educ Stud Math* **109**, 409–429 (2022). doi: 10.1007/s10649-021-10064-x.

[13] Njiku, J. (2025). Pre-service teachers manifested mathematics pedagogical content knowledge: The role of the teaching practicum. Pedagogical Research, 10(1), em0229. doi: 10.29333/pr/15647.

[14] Boevé AJ, Meijer RR, Albers CJ, Beetsma Y, Bosker RJ (2015) Introducing Computer-Based Testing in High-Stakes Exams in Higher Education: Results of a Field Experiment. PLOS ONE 10(12): e0143616. https://doi.org/10.1371/journal.pone.0143616

[15] C. Y. Piaw, "Replacing paper-based testing with computer-based testing in assessment: Are we doing wrong?." *Procedia-Social and Behavioral Sciences* 64 (2012): 655-664, doi: 10.1016/j.sbspro.2012.11.077

[16] Julie, Gess-Newsome. "A model of teacher professional knowledge and skill including PCK: Results of the thinking from the PCK Summit." *Re-examining pedagogical content knowledge in science education*. Routledge, 2015. 28-42.

[17] Schmid, Mirjam, Eliana Brianza, and Dominik Petko. "Self-reported technological pedagogical content knowledge (TPACK) of pre-service teachers in relation to digital technology use in lesson plans." *Computers in human behavior* 115 (2021): 106586, doi: 10.1016/j.chb.2020.106586

[18] Maas, L.; Brinkhuis, M.J.S.; Kester, L.; Wijngaards-de Meij, L. Cognitive Diagnostic Assessment in University Statistics Education: Valid and Reliable Skill Measurement for Actionable Feedback Using Learning Dashboards. *Appl. Sci.* **2022**, *12*, 4809. doi: 10.3390/app12104809.

[19] A. Barana, M. Marchisio, and M. Sacchet, “Interactive feedback for learning mathematics in a digital learning environment,” Educ. Sci., vol. 11, no. 6, art. 279, 2021, doi: 10.3390/educsci11060279.

[20] Rezat, S. How automated feedback from a digital mathematics textbook affects primary students’ conceptual development: two case studies. *ZDM Mathematics Education* **53**, 1433–1445 (2021). doi: 10.1007/s11858-021-01263-0.

[21] L. N. Yaqin, L. D. Prasojo, N. A. Haji-Othman, N. Yusof, and A. Habibi, "Addressing the Digital Divide in Indonesian Higher Education: Insights, Implications, and Potential Solutions," in *From Digital Divide to Digital Inclusion*, Singapore: Springer Nature Singapore, 2024, pp. 291–307. doi: 10.1007/978-981-99-7645-4_13

[22] N. Rajiha and S. Rini, "Computer-Based Standardized Testing (CBT) Summative Assessment: One Institution Experience," *JISAE (Journal of Indonesian Student Assessment and Evaluation)*, vol. 9, no. 1, 2023. doi: 10.21009/jisae.v9i1.33233.

[23] G. Kaiser and J. König, “Competence measurement in (mathematics) teacher education and beyond: Implications for policy,” High. Educ. Policy, vol. 32, no. 4, pp. 597–615, 2019, doi: 10.1057/s41307-019-00139-z.

[24] T. Kleickmann et al., “Teachers’ content knowledge and pedagogical content knowledge: The role of structural differences in teacher education,” J. Teacher Educ., vol. 64, no. 1, pp. 90–106, 2013, doi: 10.1177/0022487112460398.

[25] J. Baumert and M. Kunter, “The COACTIV model of teachers’ professional competence,” in Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers, M. Kunter et al., Eds. New York: Springer, 2013, pp. 25–48, doi: 10.1007/978-1-4614-5149-5_2.

[26] B. Rowan, R. Correnti, and R. J. Miller, “What large-scale, survey research tells us about teacher effects on student achievement: Insights from the Prospects Study of Elementary Schools,” J. Teacher Educ., vol. 104, no. 8, pp. 1525–1567, 2002, doi: 10.1177/00224871211001323.

[27] S. Blömeke, J. E. Gustafsson, and R. J. Shavelson, “Beyond dichotomies: Competence viewed as a continuum,” Z. Psychol., vol. 223, no. 1, pp. 3–13, 2015, doi: 10.1027/2151-2604/a000398.

[28] E. Mohammadpour *et al.*, "A performance-based test to measure teachers' mathematics and science content knowledge (CK) and pedagogical knowledge (PK) of teachers," *Heliyon*, vol. 9, no. 3, 2023, doi: 10.1007/s10763-021-10198-6.

[29] T. Clark and H. Endres, "Computer-based diagnostic assessment of high school students’ grammar skills with automated feedback – an international trial," *Assessment in Education: Principles, Policy & Practice*, vol. 28, no. 5-6, pp. 602–632, 2021, doi: 10.1080/0969594X.2021.1970513

[30] G. A. Nortvedt and N. Buchholtz, “Assessment in mathematics education: Responding to issues regarding methodology, policy, and equity,” ZDM Math. Educ., vol. 50, no. 4, pp. 555–570, 2018, doi: 10.1007/s11858-018-0932-3.

Bukti Review 1 oleh Reviewer II

Assessment of Mathematics Pre-Service Teachers' Professional Knowledge Based on a Computer-Based Test

Article Info

Article history:

Received mm dd, yyyy
Revised mm dd, yyyy
Accepted mm dd, yyyy

Keywords:

First keyword
Second keyword
Third keyword
Fourth keyword
Fifth keyword

ABSTRACT

This study aims to measure the achievement and profile of mathematics pre-service teachers' professional knowledge—comprising Content Knowledge (CK) and Pedagogical Content Knowledge (PCK)—through a Computer-Based Test (CBT) instrument. The study employed a quantitative approach with a descriptive-correlational design involving 30 students of the Mathematics Education (Tadris Matematika) study program at a teacher education institution (LPTK) selected purposively. Data were analyzed using descriptive statistics, Pearson correlation, and simple linear regression. The results show mean scores of 50.00% for CK and 52.67% for PCK, both below the ideal competency threshold (75%), with 60.00% of the students in the low category and the Knowledge of Content and Curriculum (KCC) dimension as the weakest (46.25%). A strong positive correlation was found between CK and PCK ($r = 0.72$; $p < 0.001$), and CK predicted 52% of the variance in PCK ($PCK = 1.708 \times CK - 32.734$). Diagnostic analysis at the dimension, indicator, and individual-student levels demonstrates that the CBT is able to generate specific and actionable achievement profiles as a basis for instructional intervention. These findings underscore the need to strengthen LPTK curricula that integrate content mastery and pedagogical capacity grounded in diagnostic data.

Corresponding Author:

Commented [AM1]: **The abstract must have the following structure:**
 Background / **Context**
Briefly introduce the research topic and the main issue being studied.
 Research Gap / Problem Statement
Explain what is still unclear, underexplored, or problematic in previous studies.
 Objective / Purpose of the Study
State the main aim of the research clearly and directly.
 Methodology
Briefly describe the research design, data source, sample/respondents, instruments, and analysis technique.
 Key Findings / Results
Present the most important research findings. Avoid too much detail.
 Conclusion
Summarize the main meaning of the findings.
 Implications / Contribution
Explain the theoretical, practical, or policy contribution of the study.
 Keywords
Provide 4–6 relevant keywords that represent the main topic, method, theory, and research context.

INTRODUCTION

The quality of a nation's mathematics education depends heavily on teachers' professional competence. Cross-national studies have shown that mathematics teachers' professional knowledge—comprising Content Knowledge (CK) and Pedagogical Content Knowledge (PCK)—is a primary predictor of instructional effectiveness and students' academic achievement [1], [2]. The 2022 PISA results placed Indonesian students' mathematics literacy 68th out of 81 countries, below the OECD average, indicating a fundamental problem in the quality of mathematics instruction [3]. This condition is inseparable from the quality of professional knowledge formed in teacher education institutions (LPTK), so ensuring an adequate professional-knowledge profile of pre-service teachers becomes a strategic imperative [4], [5].

Commented [AM2]: Here is the **structure of a good Introduction section** in an international scientific article:
 General Background / Research Context
Introduce the broader topic of the study. Explain why the issue is important, relevant, and worth investigating. This part should provide the general context before moving to a more specific problem.
 Specific Research Problem
Narrow the discussion to the specific problem addressed in the study. Explain what issue, phenomenon, or challenge needs to be examined.
 Review of Previous Studies
Summarize relevant previous research. Do not simply list studies one by one. Instead, show the pattern of previous findings, such as what has been studied, what has been found, and what remains unclear.
 Research Gap
Identify the gap in the existing literature. This may include limited studies in a certain context, inconsistent findings, underexplored variables, methodological limitations, or a lack of empirical evidence.
 Novelty / Originality of the Study
Explain what makes the current study different from previous research. The novelty may come from the object, context, method, variables, theoretical approach, or data used in the study.
 Research Objective
Clearly state the aim of the study. This section usually begins with phrases such as:
 Research Contribution / Significance
Explain the contribution of the study. This may include theoretical contribution, practical implications, policy relevance, or benefits for stakeholders.

Theoretically, the Mathematical Knowledge for Teaching (MKT) framework distinguishes CK—comprising common, specialized, and horizon content knowledge—from PCK, which encompasses knowledge of content and students, content and teaching, and content and curriculum [6], [7]. Systematic reviews affirm that PCK is the most frequently studied yet the most difficult component to measure consistently because of its contextual and multidimensional nature [8], [9]. Studies in Indonesia show that the professional knowledge of mathematics pre-service teachers varies widely across institutions and curricula, with a notable weakness in the PCK aspect [10], [11]. Similar findings have been reported in cross-national contexts [12], [13].

Alongside the rapid digitalization of education, competency-evaluation systems are now shifting massively from traditional approaches toward technology-based assessment. Various studies indicate that CBT is often more valid and reliable than conventional examination methods [14], [15]; however, its use to objectively measure both the content knowledge and the pedagogical content knowledge of pre-service teachers simultaneously remains very limited because of the complexity of these constructs [16], [17]. The potential of CBT to generate diagnostic feedback based on learning analytics has also not been optimally exploited [18], even though interactive digital feedback has been shown to improve mathematical understanding [19], [20].

The urgency of this study is heightened in the context of Indonesia's uneven higher-education digital transformation [21], which is not yet supported by adequate standardization of assessment systems [22]. Measuring teacher competence in the digital era is required to produce data that are actionable for study-program managers and policymakers [23]. A number of studies affirm that CK and PCK are relatively independent constructs that require different intervention pathways [24], [25], and that teacher effects on student achievement are substantial [26].

The novelty of this study lies in three aspects. **First**, theoretically, the study integrates the MKT framework and technology-based assessment within a single coherent quantitative measurement design to identify CK and PCK profiles simultaneously [27], [28]. **Second**, methodologically, the CBT is used as an autonomous, automated assessment medium that generates layered diagnostic feedback—at the total, dimension, indicator, and individual-student levels—so that test results can be acted upon directly for instructional improvement [29]. **Third**, contextually, the study targets LPTK students in Indonesia who have so far been overlooked in in-depth empirical studies [30]. This study aims to: (1) measure students' CK and PCK achievement and profiles; (2) identify strengths and weaknesses in each dimension and indicator; and (3) demonstrate the use of diagnostic profiles—including at the individual-student level—as a basis for designing instructional interventions.

METHOD

This study employed a quantitative approach with a descriptive-survey and correlational design. The population comprised students of the Mathematics Education (Tadris Matematika) study program at an LPTK who were in their fifth and seventh semesters. A sample of 30 students was determined through purposive sampling with the inclusion criteria: (1) active students who had adequately completed mathematics-content and pedagogy courses; and (2) willing to take the entire series of CBT tests. Given the

Commented [AM3]: Here is the **structure of a good Methods section** in an international scientific article:

1.**Research Design**
Explain the type of research used in the study.
For example: qualitative, quantitative, mixed-methods, experimental, survey, descriptive, case study, or literature review.

1.**Research Location and Time**
State where and when the research was conducted. This is important, especially for field research.

2.**Population and Sample / Participants**
Explain who or what became the subject of the study.
This section may include population, sample size, characteristics of respondents, and inclusion or exclusion criteria.

3.**Sampling Technique**
Describe how the sample or participants were selected.
Common techniques include random sampling, purposive sampling, stratified sampling, snowball sampling, or census.

4.**Data Sources**
Explain the sources of data used in the study.
Data may be primary data, secondary data, or both.

5.**Data Collection Techniques**
Describe how the data were collected.
This may include surveys, interviews, observation, documentation, laboratory tests, experiments, or focus group discussions.

6.**Research Instruments**
Explain the tools used to collect data, such as questionnaires, interview guidelines, observation sheets, laboratory equipment, or measurement scales.

7.**Variables and Measurement**
For quantitative studies, clearly define the independent, dependent, mediator, moderator, or control variables.
Also explain how each variable was measured.

8.**Validity and Reliability / Trustworthiness**
For quantitative research, explain validity and reliability testing.
For qualitative research, explain credibility, transferability, dependability, and confirmability.

1.**Data Analysis Technique**
Explain how the data were analyzed.
For quantitative studies, this may include descriptive statistics, regression, SEM, ANOVA, DEA, or correlation analysis.
For qualitative studies, this may include thematic analysis, content analysis, coding, or interactive model analysis.

2.**Ethical Considerations**
Explain how ethical issues were handled, especially when involving human participants.
This may include informed consent, confidentiality, anonymity, and voluntary participation.

limited sample size, all data were analyzed as a single intact group (n = 30) and were not directed toward inter-semester comparison.

The operationalization of the variables comprised two main constructs. CK was defined as mastery of the mathematics content to be taught, measured through 25 multiple-choice items; PCK was defined as the ability to integrate content knowledge with pedagogical strategies, also measured through 25 multiple-choice items. Scores were converted into percentages of achievement relative to the ideal maximum score. The CBT instrument was designed to generate layered diagnostic reports, namely the total score, achievement per dimension, achievement per indicator, and an individual achievement profile for each student. Each item was mapped to a specific indicator and dimension (see the notes below Table 1) so that the CBT system could produce achievement per dimension and per indicator for all students at once, as well as their individual diagnostic profiles. The composition of the instrument is presented in Table 1.

Table 1. Composition of CBT Instrument Items by MKT Dimension

Domain	Dimension	Number of Items
CK	Common Content Knowledge (CCK)	7
CK	Specialized Content Knowledge (SCK)	14
CK	Horizon Content Knowledge (HCK)	4
PCK	Knowledge of Content and Students (KCS)	10
PCK	Knowledge of Content and Teaching (KCT)	9
PCK	Knowledge of Content and Curriculum (KCC)	6
Total		50

Notes on dimension codes and indicator coverage (based on the developed CK and PCK instruments). **CK – CCK** (Common Content Knowledge, items 1–7): understanding basic concepts (e.g., lines and angles), performing calculations and measurements accurately, and applying mathematical concepts in everyday life. **CK – SCK** (Specialized Content Knowledge, items 8–21): explaining concepts and principles, analyzing and connecting mathematical objects, proving properties, building models, and solving complex mathematical problems. **CK – HCK** (Horizon Content Knowledge, items 22–25): connecting school mathematics with advanced academic mathematics and with other disciplines. **PCK – KCS** (Knowledge of Content and Students, items 1–10): analyzing students' preconceptions and understanding as well as predicting and analyzing students' misconceptions, errors, and difficulties. **PCK – KCT** (Knowledge of Content and Teaching, items 11–19): selecting and organizing learning materials, selecting sources, strategies, and techniques, and selecting and analyzing representations for presenting material. **PCK – KCC** (Knowledge of Content and Curriculum, items 20–25): designing content and topic combinations, applying standards and curricula, and designing instruments to measure students' abilities.

The analysis techniques comprised descriptive and inferential analyses. Descriptive analysis produced the mean (M), standard deviation (SD), minimum, maximum, median, and percentage of achievement at the total, dimension, indicator, and individual levels, as well as a three-level categorization (High, Medium, Low). Inferential analysis was conducted through two techniques: (1) Pearson correlation to test the CK–PCK relationship and the corresponding inter-dimension relationships (CCK–KCS, SCK–KCT, HCK–KCC) at $\alpha = 0.05$; and (2) simple linear regression with CK as the predictor and PCK as the criterion.

All scoring, achievement computation, correlation, and linear regression were performed using Microsoft Excel through standardized formulas (including CORREL, SLOPE, INTERCEPT, and RSQ) so that every figure can be traced back to the raw data.

Participants' identities were kept confidential by using codes instead of names throughout the recording, processing, and reporting of data; in this article, the individual-profile example is presented under the pseudonym Student 1.

RESULTS AND DISCUSSION

3.1. Results

Overall, students' professional-knowledge achievement was in the medium category and had not reached the ideal competency threshold (≥ 75%). The mean CK score was 50.00% (SD = 10.16), PCK 52.67% (SD = 23.96), and the combined Professional Knowledge (PK) score 51.33% (SD = 16.04). Complete descriptive statistics are presented in Table 2.

Table 2. Descriptive Statistics of CK, PCK, and Professional Knowledge Scores (n = 30)

Variable	Mean (%)	SD	Min	Max	Median
Content Knowledge (CK)	50.00	10.16	24	68	52.00
Pedagogical Content Knowledge (PCK)	52.67	23.96	8	96	54.00
Professional Knowledge (PK)	51.33	16.04	26	78	50.00

Table 2 shows that the mean PCK is slightly higher than CK (a difference of 2.67 points), an interesting pattern indicating that pedagogical-transformation capacity slightly exceeds content mastery. The large standard deviation of PCK (SD = 23.96) and the much smaller one for CK (SD = 10.16) reflect substantial heterogeneity in knowledge profiles, with score ranges of 24–68% for CK and 8–96% for PCK. The distribution of Professional Knowledge categories is presented in Table 3.

Table 3. Distribution of Professional Knowledge Categories (n = 30)

Category	Frequency (f)	Percentage (%)
High	0	0.00
Medium	12	40.00
Low	18	60.00
Total	30	100.00

All students (100%) were in the Low or Medium category, with 60.00% in the Low category and 40.00% in the Medium category, and no student reaching the High category. This indicates that the majority of pre-service teachers have not attained an optimal level of professional-knowledge mastery. Achievement per dimension is presented in Table 4 and Figure 1.

Table 4. Average Achievement per CK and PCK Dimension

Domain	Dimension	Items	Mean (%)	SD
CK	CCK	7	73.33	20.47
CK	SCK	14	36.91	24.16
CK	HCK	4	55.00	35.60
PCK	KCS	10	54.07	33.18
PCK	KCT	9	57.50	26.14

Commented [AM4]: Here is the **structure of a good Results section** in an international scientific article:

1. **Brief Opening Statement**
Start with a short statement explaining what the Results section presents.
Example:
This section presents the findings obtained from the data analysis, including respondent characteristics, descriptive statistics, and hypothesis testing results.
2. **Description of Respondents / Sample Characteristics**
Present the profile of respondents, samples, objects, or research units.
This may include age, gender, education, location, experience, commodity type, firm size, or other relevant characteristics.
3. **Descriptive Statistics / General Findings**
Show the general pattern of the data before presenting deeper analysis.
This may include mean, percentage, frequency, standard deviation, minimum and maximum values.
4. **Main Research Findings**
Present the core findings that directly answer the research objectives or research questions.
The order of presentation should follow the order of the objectives, hypotheses, or analytical framework.
5. **Statistical Test Results / Analytical Results**
If the study is quantitative, present the results of statistical analysis, such as regression, correlation, ANOVA, SEM, DEA, or other tests.
Include important values such as coefficient, p-value, t-value, F-value, R-square, or confidence interval.
6. **Tables and Figures**
Use tables and figures to present important findings clearly.
Do not repeat all numbers in the text. Instead, highlight the most important patterns.
7. **Qualitative Findings / Themes**
If the study is qualitative, present the findings based on themes, categories, or patterns from interviews, observations, or documents.
Use selected quotations only when they strongly support the finding.
8. **Hypothesis Testing Results**
If the article includes hypotheses, clearly state whether each hypothesis is accepted or rejected.
9. **Summary of Key Findings**
End the Results section with a brief summary of the most important findings before moving to the Discussion.

Domain	Dimension	Items	Mean (%)	SD
PCK	KCC	6	46.25	32.30

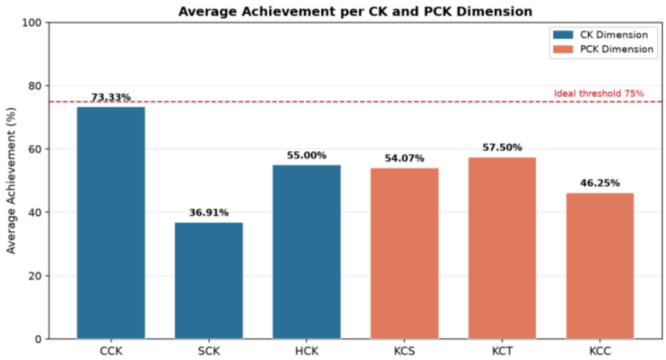


Figure 1. Average Achievement per CK and PCK Dimension

In the CK domain, the CCK dimension was the highest (73.33%) and SCK the lowest (36.91%); in the PCK domain, KCT was the highest (57.50%) while KCC was the lowest of all six dimensions (46.25%). Diagnostically, the low SCK and KCC signal a weak ability to relate content to the structure and organization of the curriculum—an area that needs to be prioritized in intervention. Achievement at the indicator level is presented in Table 5 and visualized in Figure 2.

Table 5. Average Achievement per CK and PCK Indicator

Domain	Indicator	Mean (%)
CK	CCK1	86.67
CK	CCK2	80.00
CK	CCK3	46.67
CK	SCK1	63.33
CK	SCK2	46.67
CK	SCK3	43.33
CK	SCK4	20.00
CK	SCK5	23.33
CK	SCK6	30.00
CK	SCK7	33.33
CK	HCK1	53.33
CK	HCK2	56.67
PCK	KCS1	58.33
PCK	KCS2	48.33
PCK	KCS3	58.33
PCK	KCS4	51.67
PCK	KCS5	53.33
PCK	KCT1	55.00
PCK	KCT2	66.67
PCK	KCT3	60.00
PCK	KCT4	36.67
PCK	KCC1	51.67
PCK	KCC2	43.33
PCK	KCC3	45.00

Notes on indicator codes (based on the CK and PCK instruments). CK indicators — **CCK1**: understanding basic school-mathematics concepts, such as lines, angles, sequences, and basic statistics (items 1–3); **CCK2**: performing calculations and measurements accurately (items 4–5); **CCK3**: applying mathematical concepts in everyday life (items 6–7); **SCK1**: explaining academic mathematical concepts and principles (items 8–9); **SCK2**: analyzing mathematical objects (items

10–11); **SCK3**: connecting mathematical properties and objects (items 12–13); **SCK4**: proving mathematical properties or theorems (items 14–16); **SCK5**: applying proof strategies (item 17); **SCK6**: building and constructing mathematical models (items 18–19); **SCK7**: solving complex mathematical problems (items 20–21); **HCK1**: connecting school mathematics with advanced academic mathematics (items 22–23); **HCK2**: connecting mathematics with other disciplines (items 24–25). PCK indicators — **KCS1**: analyzing students' preconceptions (items 1–2); **KCS2**: identifying students' level of understanding (items 3–4); **KCS3**: predicting students' misconceptions (items 5–6); **KCS4**: identifying sources of students' learning difficulties (items 7–8); **KCS5**: analyzing students' literacy issues and errors (items 9–10); **KCT1**: organizing and sequencing learning materials (items 11–12); **KCT2**: selecting learning sources, strategies, and techniques (items 13–14); **KCT3**: selecting representations and illustrations for presenting material (items 15–17); **KCT4**: analyzing the weaknesses and strengths of representations (items 18–19); **KCC1**: composing content and topic combinations (items 20–21); **KCC2**: applying standards and curricula (items 22–23); **KCC3**: designing instruments to measure students' abilities (items 24–25).

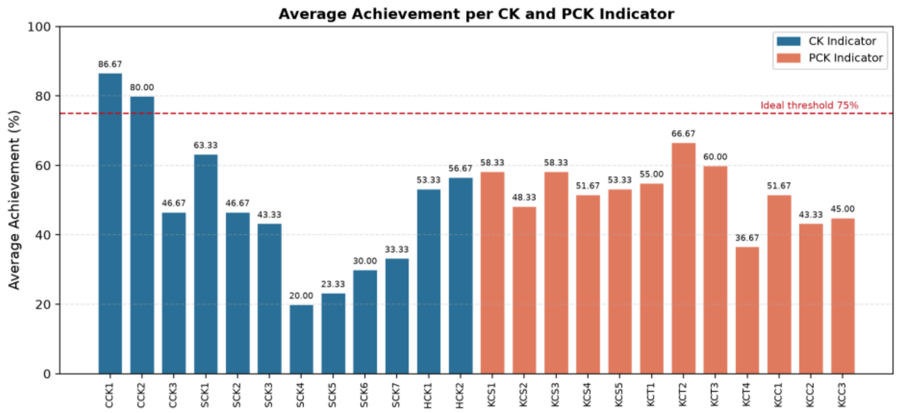


Figure 2. Average Achievement per CK and PCK Indicator

Figure 2 visualizes the average achievement of each indicator exactly as listed in Table 5. The range of achievement across indicators spans 20.00% to 86.67%. The highest achievement is on the CK indicator CCK1 (86.67%; understanding basic school-mathematics concepts), which is fundamental and therefore the most widely mastered by students. Conversely, the lowest overall achievement is on indicator SCK4 (20.00%; proving mathematical properties or theorems)—the skill with the highest cognitive demand in the CK domain—followed by SCK5 (23.33%; applying proof strategies) and SCK6 (30.00%; building and constructing mathematical models). The pattern in the SCK dimension declines gradually with complexity, from explaining concepts and principles (SCK1; 63.33%), analyzing and connecting mathematical objects (SCK2 46.67%; SCK3 43.33%), solving complex problems (SCK7; 33.33%), to the most difficult modeling and proving (SCK6 30.00%; SCK5 23.33%; SCK4 20.00%)—a pattern consistent with cognitive load theory. Because each dimension's achievement remains the mean of its constituent indicators—for example, KCS (54.07%) is the mean of KCS1–KCS5 and CCK (73.33%) the mean of CCK1–CCK3—this pattern maps precisely which items and indicators are the sources of weakness, so that lecturers can design remediation targeted at concrete rather than generic weaknesses. The results of the Pearson correlation test are presented in Table 6.

Table 6. Pearson Correlation between Domains and between Parallel Dimensions

Relationship	**r**	**p**	**Interpretation**

CK – PCK	0.72	< 0.001	Strong
CCK – KCS	0.42	< 0.05	Moderate
SCK – KCT	0.54	< 0.01	Moderate
HCK – KCC	0.86	< 0.001	Very strong

The strong positive correlation between CK and PCK ($r = 0.72$; $p < 0.001$) confirms that the two domains are constructively interdependent, reinforced by the correlations between parallel dimensions, especially HCK–KCC ($r = 0.86$) and SCK–KCT ($r = 0.54$). The predictive relationship of CK on PCK was tested through simple linear regression (Table 7 and Figure 3).

Table 7. Results of the Simple Linear Regression of CK on PCK

Parameter	Value
Slope (b)	1.708
Intercept (a)	−32.734
Coefficient of determination (R^2)	0.5245
Pearson r	0.72
Equation	PCK = 1.708 × CK − 32.734

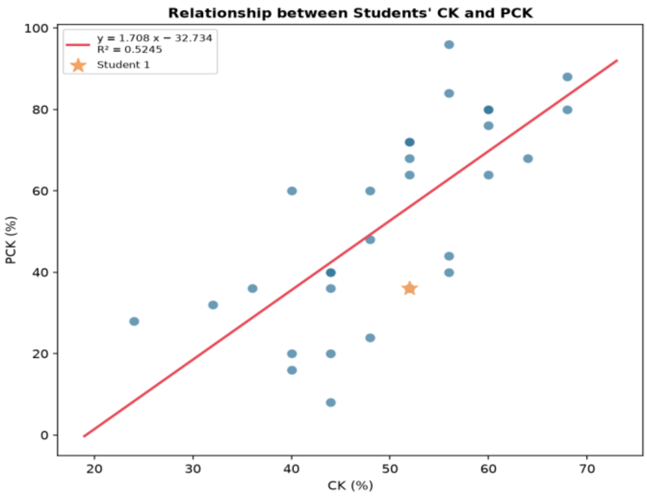


Figure 3. Scatter Plot and Regression Line of the CK–PCK Relationship

The equation PCK = 1.708 × CK − 32.734 with $R^2 = 0.52$ indicates that about 52% of the variance in PCK can be explained by CK. The remaining 48% of the variance signals that PCK does not form automatically from content mastery but requires an explicit pedagogical learning pathway.

The main advantage of the CBT instrument lies in its ability to break down aggregate achievement into individual diagnostic profiles automatically. As an illustration, Table 8 and Figure 4 summarize the achievement of one student disguised as Student 1, while Figure 5 and Figure 6 display authentic screenshots of the diagnostic scoring report generated directly by the CBT system for Student 1. Such individual reports are produced automatically by the system for every participant without any manual scoring.

Table 8. Individual Achievement Profile of Student 1 Compared with the Class Average

Domain	Student 1 Score (%)	Class Average (%)	Difference	Category
Content Knowledge (CK)	52.00	50.00	+2.00	Low
Pedagogical Content Knowledge (PCK)	36.00	52.67	−16.67	Low

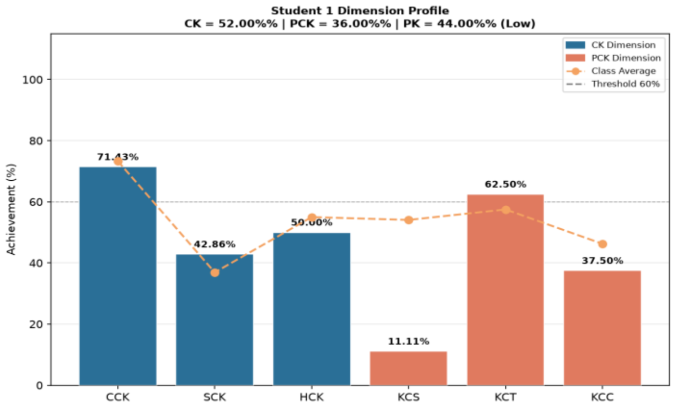


Figure 4. CK and PCK Achievement Profile of Student 1 Compared with the Class Average

At the individual level, the CBT system automatically scores the answers and generates a diagnostic scorecard for each participant. Student 1, for example, answered 13 of 25 CK items correctly (52.00%) and 9 of 25 PCK items correctly (36.00%); the CK achievement was slightly above the class average (a difference of +2.00 points) while the PCK achievement was far below it (a difference of −16.67 points), and neither reached the ideal competency threshold (75%), placing the student in the Low category for both domains. Authentic screenshots of the report are presented in Figure 5 (CK) and Figure 6 (PCK).

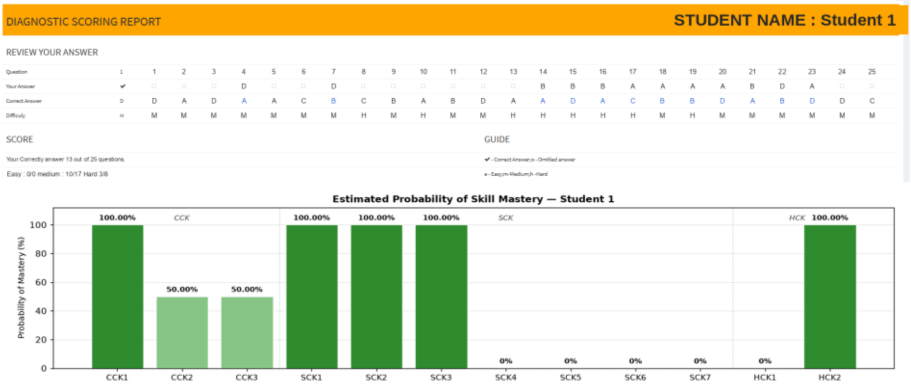


Figure 5. CBT Diagnostic Scorecard for the CK Domain for Student 1 (original system screenshot)

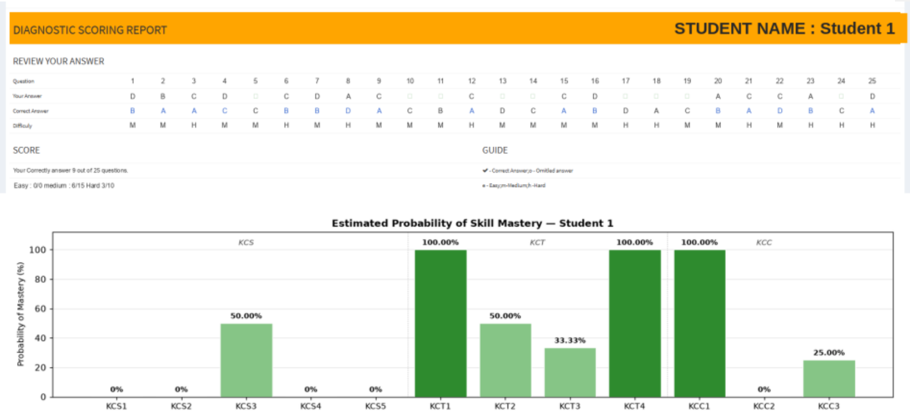


Figure 6. CBT Diagnostic Scorecard for the PCK Domain for Student 1 (system screenshot)

Figures 5 and 6 reveal a role of the CBT that conventional tests can hardly provide. In addition to recapping correct/incorrect answers per item along with their difficulty level—a combination of the Medium (M) and Hard (H) categories according to the item analysis of the instrument—the system presents an Estimated Probability of Skill Mastery graph that maps Student 1's mastery probability for each skill indicator (CCK1–HCK2 for CK and KCS1–KCC3 for PCK), complete with a mapping of sample items to each indicator. Thus, the CBT does not merely provide a final score but directly identifies which indicators have been mastered and which remain weak—in Student 1's case, several indicators show a mastery probability close to zero—so that lecturers can follow up with specific and immediate remediation. It is this automation of scoring, standardization of reporting, and diagnostic depth that affirms the added value of the CBT in assessing the professional knowledge of mathematics pre-service teachers.

Furthermore, the two scorecards allow Student 1's achievement to be traced down to the indicator level, not stopping at dimension averages. In the CK domain (Figure 5), Common Content Knowledge (CCK) was the strongest dimension (71.43%; 5 of 7 items): understanding basic school-mathematics concepts (CCK1) was fully mastered (100.00%), whereas performing calculations and measurements (CCK2) and applying mathematical concepts in everyday life (CCK3) were each only partially mastered (50.00%). Specialized Content Knowledge (SCK) stood at 42.86% (6 of 14 items) with a highly contrasting pattern: three basic indicators—explaining academic mathematical concepts and principles (SCK1), analyzing mathematical objects (SCK2), and connecting mathematical properties and objects (SCK3)—were all fully mastered (100.00%), while four indicators demanding higher-order reasoning—proving mathematical properties or theorems (SCK4), applying proof strategies (SCK5), building and constructing mathematical models (SCK6), and solving complex mathematical problems (SCK7)—were all at 0.00%, marking a total deficit in advanced mathematical reasoning. Horizon Content Knowledge (HCK) stood at 50.00% (2 of 4 items): connecting mathematics with other disciplines (HCK2) was fully mastered (100.00%), but connecting school mathematics with advanced academic mathematics (HCK1) remained at 0.00%. In the PCK domain (Figure 6), the indicator profile was also uneven within each dimension. Knowledge of Content and Teaching

(KCT; dimension mean 62.50%) was the strongest: organizing and sequencing learning materials (KCT1) and analyzing the weaknesses and strengths of representations (KCT4) were fully mastered (100.00%), followed by selecting learning sources, strategies, and techniques (KCT2; 50.00%) and selecting representations and illustrations for presenting material (KCT3; 33.33%). Knowledge of Content and Curriculum (KCC; dimension mean 37.50%) rested on composing content and topic combinations (KCC1), which was fully mastered (100.00%), while designing instruments to measure students' abilities (KCC3) was only 25.00% and applying standards and curricula (KCC2) was still 0.00%. Conversely, Knowledge of Content and Students (KCS) was the weakest PCK dimension and the weakest point of the entire profile (11.11%): only predicting students' misconceptions (KCS3) was partially mastered (50.00%), while analyzing students' preconceptions (KCS1), identifying students' level of understanding (KCS2), identifying sources of students' learning difficulties (KCS4), and analyzing students' literacy issues and errors (KCS5) were all at 0.00%. This pattern confirms that Student 1's weaknesses are concentrated in a number of indicators that are genuinely untouched—SCK4, SCK5, SCK6, SCK7, and HCK1 in the CK domain and KCS1, KCS2, KCS4, KCS5, and KCC2 in the PCK domain, all at 0.00%—so that these ten indicators are the most urgent remediation priorities, while the indicators already fully mastered (CCK1, SCK1, SCK2, SCK3, HCK2, KCT1, KCT4, and KCC1) can serve as a foundation for reinforcement. Thus, indicator-based diagnosis provides a far sharper and more specific direction for individual intervention than reading dimension averages alone.

3.2. Discussion

The main findings reveal three key facts. **First**, neither CK (50.00%) nor PCK (52.67%) reached the minimum competency threshold; interestingly, PCK was even slightly higher than CK, reflecting the knowledge-transformation challenge articulated by Shulman [7] and reinforced in the MKT framework [6]. **Second**, the SCK dimension was the weakest point of CK (36.91%) and KCC the weakest point of PCK (46.25%), indicating a double deficit in the mastery of advanced mathematical content and curriculum, which in the COACTIV model strongly determines the quality of instructional planning [2], [25]. **Third**, the lower standard deviation of CK (10.16) compared with PCK (23.96) shows that CK achievement is more uniform, whereas PCK is highly heterogeneous, indicating large variation in students' pedagogical capacity.

The significant positive correlations, especially HCK–KCC ($r = 0.86$) and SCK–KCT ($r = 0.54$), show that content mastery in particular dimensions goes hand in hand with the relevant pedagogical capacity. However, the regression equation ($R^2 = 0.52$) affirms that content mastery explains only part of the variance in PCK, consistent with findings that pre-service teachers in developing countries tend to have lower PCK scores than CK because LPTK curricula place greater emphasis on content mastery [1], [10]. The case of Student 1 reinforces this at the individual level: Student 1 (CK = 52%, PCK = 36%) falls into the Low category with a highly uneven profile, namely CK achievement slightly above the class average while PCK is far below it. In the PCK domain, KCT is relatively the highest (62.50%) while KCS is the lowest (11.11%) and KCC is low (37.50%), indicating an ability to select teaching strategies that is not yet accompanied by an adequate understanding of student characteristics or curriculum mastery. Intervention needs to target KCS and KCC simultaneously, especially the ability to analyze

Commented [AM5]: Here is the **structure of a good Discussion section** in an international scientific article:

1. **Brief Restatement of the Main Findings**
Begin by summarizing the key findings of the study. Do not repeat all results in detail. Focus only on the most important findings.
2. **Interpretation of the Findings**
Explain what the findings mean. This is the core of the Discussion section.
The author should not only state the result, but also explain why the result occurred.
3. **Comparison with Previous Studies**
Compare your findings with previous research. Explain whether your results support, extend, or contradict earlier studies.
4. **Theoretical Explanation**
Connect the findings with relevant theories, concepts, or frameworks.
This part shows the academic strength of the article.
5. **Explanation of Unexpected Findings**
If there are findings that are not significant, inconsistent, or different from previous studies, explain possible reasons.
6. **Implications of the Study**
Explain the meaning of the findings for theory, practice, policy, or stakeholders.
The implications may include:
- **Theoretical implications**: contribution to academic knowledge.
- **Practical implications**: recommendations for practitioners, farmers, managers, or institutions.
- **Policy implications**: suggestions for government or regulatory bodies.

1. **Limitations of the Study**
State the limitations honestly. This does not weaken the article; instead, it shows academic awareness.
2. **Suggestions for Future Research**
Recommend what future researchers can examine based on the limitations or unanswered questions.

preconceptions, identify levels of understanding, and identify sources of students' learning difficulties.

Theoretically, this study validates and extends the application of the MKT framework and the COACTIV model to the context of religiously affiliated LPTKs in Indonesia, which are almost unrepresented in the indexed literature [6], [25], [11]. Methodologically, the study demonstrates a CBT model that produces layered diagnostic profiles—from the total, dimension, and indicator levels to the individual student—thus addressing the previously identified gap in using CBT to measure pedagogical competence [22], [28]. Achievement profiles that can be decomposed to the individual level, as shown in the case of Student 1, strengthen the function of the CBT as an instrument that generates actionable diagnostic feedback [18]. In policy terms, the mean PK of 51.33% and the absence of any student in the High category provide urgent empirical evidence to drive data-based curriculum reform rather than normative assumptions [5], [23], particularly in the context of Indonesia's 2022 PISA results [3].

This study opens a new perspective by positioning the CBT not merely as an examination tool but as a technology-based diagnostic-assessment ecosystem that produces rich cognitive data, decomposed per dimension and per individual, and actionable for program managers [23], [30]. Several limitations should be noted, namely the relatively small sample size (n = 30) and the use of purposive sampling, which limit generalizability, as well as the still-bivariate scope of the analysis, which has not yet modeled other determinants simultaneously.

CONCLUSION AND RECOMMENDATIONS

This study succeeded in identifying the professional-knowledge profile of mathematics pre-service teachers at an LPTK through a CBT instrument. The mean CK (50.00%) and PCK (52.67%) scores were both below the ideal competency threshold (75%), with students distributed across the Low (60.00%) and Medium (40.00%) categories and none reaching the High category, with CCK as the highest CK dimension (73.33%), SCK the weakest (36.91%), and KCC as the weakest PCK dimension (46.25%). There was a strong positive correlation between CK and PCK ($r = 0.72$; $p < 0.001$), and CK predicted 52% of the variance in PCK ($PCK = 1.708 \times CK - 32.734$). Layered diagnostic analysis—down to the individual level as exemplified in the case of Student 1 (CK 52.00%; PCK 36.00%)—shows that the CBT is able to generate specific and actionable achievement profiles. Overall, these findings reinforce the view that CK and PCK are interrelated yet asymmetrically developing constructs, and affirm the feasibility of the CBT as an efficient and standardized diagnostic-assessment instrument in LPTKs.

Recommendations. In practical terms, the Mathematics Education (Tadris Matematika) study program is recommended to design remedial interventions targeted not only at the weakest dimensions (SCK in CK and KCC in PCK) but also down to the weakest indicators as read in Table 5. In the CK domain, the top priority is strengthening mathematical proof ability—namely the indicators proving properties or theorems (SCK4; 20.00%) and applying proof strategies (SCK5; 23.33%)—through gradually structured proof practice (scaffolded proof), analysis of proof errors, and habituation to writing deductive arguments; followed by strengthening mathematical modeling (SCK6; 30.00%) and complex problem solving (SCK7; 33.33%) through problem-based learning and mathematization projects. In the PCK domain, the focus is directed at the indicators

Commented [AM6]: Here is the **structure of a good Conclusion section** in an international scientific article:

1.**Restatement of the Research Objective**
Begin by briefly restating the main purpose of the study. Avoid copying the objective exactly from the Introduction.
2.**Summary of the Main Findings**
Present the most important findings in a concise way. Do not include too many numbers, tables, or detailed statistical results.
3.**Answer to the Research Question**
Make sure the conclusion directly answers the research question or research objective.
4.**Theoretical Contribution**
Explain how the study contributes to academic knowledge, theory, or existing literature.
5.**Practical or Policy Implications**
Explain how the findings can be useful for practitioners, institutions, policymakers, farmers, managers, or other stakeholders.
6.**Limitations of the Study**
Briefly mention the limitations of the study. This shows academic honesty.
7.**Recommendations for Future Research**
Suggest what future studies should explore.

analyzing the weaknesses and strengths of representations (KCT4; 36.67%), applying standards and curricula (KCC2; 43.33%), and designing measurement instruments (KCC3; 45.00%)—through case-based learning, analysis of teaching videos, examination of curriculum documents, and assignments to develop test blueprints and assessment tools. This indicator-targeting strategy can also be applied at the individual level: the profile of Student 1, for example, who falls into the Low category, calls for prioritizing the untouched indicators—the four advanced-reasoning SCK indicators (proving, proof strategies, modeling, and complex problem solving; SCK4–SCK7) and connecting school mathematics with advanced academic mathematics (HCK1) in the CK domain, as well as analyzing students' preconceptions (KCS1), identifying students' level of understanding (KCS2), identifying sources of students' learning difficulties (KCS4), analyzing students' literacy issues and errors (KCS5), and applying standards and curricula (KCC2) in the PCK domain—all of which are still at 0.00%. For LPTK managers, these results encourage the integration of a diagnostic CBT system—capable of mapping achievement down to the indicator and individual levels—into the quality-assurance cycle so that specific weaknesses can be monitored and followed up continuously. Theoretically and methodologically, further research is recommended to use a larger and more representative sample, a longitudinal design, cross-institution comparison, and multivariate models involving other determinants (academic background, learning motivation, and the quality of lecturers' teaching) to strengthen external validity and deepen understanding of the formation of mathematics pre-service teachers' professional knowledge.

ACKNOWLEDGMENTS

The authors thank the Mathematics Education (Tadris Matematika) study program of the LPTK for facilitating the implementation of the research, as well as all the students who participated as respondents.

REFERENCES

[1] X. Yang and G. Kaiser, "The impact of mathematics teachers' professional competence on instructional quality and students' mathematical learning outcomes," Teach. Teach. Educ., vol. 110, art. 103582, 2022, doi: 10.1016/j.tate.2021.103582

[2] S. Krauss et al., "Pedagogical content knowledge and content knowledge of secondary mathematics teachers," Journal of Educational Psychology, 100(3), 716–725, 2008, doi: 10.1037/0022-0663.100.3.716.

[3] A. Schleicher, PISA 2022 Results: The State of Learning and Equity in Education. Paris: OECD Publishing, 2023.

[4] M. T. Tatto, J. Schwille, S. L. Senk, L. Ingvarson, R. Peck, G. Rowley, K. Bankov, M. Rodriguez, and M. Reckase, "Policy, Practice, and Readiness to Teach Primary and Secondary Mathematics in 17 Countries: Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)". Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement, 2012.

[5] H. Qian and P. Youngs, "The effect of teacher education programs on future elementary mathematics teachers' knowledge: A five-country analysis using TEDS-M data". J. Math. Teacher Educ., vol. 19, no. 4, pp. 371–396, 2016, doi: 10.1007/s10857-014-9297-9.

[6] D. L. Ball, M. H. Thames, and G. Phelps, "Content knowledge for teaching: What makes it special?," J. Teacher Educ., vol. 59, no. 5, pp. 389–407, 2008, doi: 10.1177/0022487108324554.

[7] L. S. Shulman, "Those who understand: Knowledge growth in teaching," Educ. Res., vol. 15, no. 2, pp. 4–14, 1986, doi: 10.3102/0013189X015002004.

[8] M. Grigaliūnienė, E. Lehtinen, L. Verschaffel, and F. Depaepe, "Systematic review of research on pedagogical content knowledge in mathematics: Insights from a topic-specific approach," *ZDM Mathematics Education*, vol. 57, pp. 777–794, 2025, doi: 10.1007/s11858-025-01684-1.

[9] F. Depaepe, L. Verschaffel, and G. Kelchtermans, "Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research," Teach. Teach. Educ., vol. 34, pp. 12–25, 2013, doi: 10.1016/j.tate.2013.03.001.

[10] Ningsih, S. Y., Turmudi, & Juandi, D. (2020). *Pedagogical content knowledge (PCK) profile of prospective teachers in mathematics learning*. Journal of Physics: Conference Series, 1521(3), 032057, doi: 10.1088/1742-6596/1521/3/032057

[11] S. Fiangga, S. Khabibah, S. M. Amin, dan R. Ekawati, "A Learning Design Analysis of the Pre-service Teachers' Mathematics Pedagogical Content Knowledge," *J. Phys.: Conf. Ser.*, vol. 1899, no. 1, p. 012153, 2021, doi: 10.1088/1742-6596/1899/1/012153.

[12] Yang, X., Schwarz, B. & Leung, I.K.C. Pre-service mathematics teachers' professional modeling competencies: a comparative study between Germany, Mainland China, and Hong Kong. *Educ Stud Math* **109**, 409–429 (2022). doi: 10.1007/s10649-021-10064-x.

[13] Njiku, J. (2025). Pre-service teachers manifested mathematics pedagogical content knowledge: The role of the teaching practicum. Pedagogical Research, 10(1), em0229. doi: 10.29333/pr/15647.

[14] Boevé AJ, Meijer RR, Albers CJ, Beetsma Y, Bosker RJ (2015) Introducing Computer-Based Testing in High-Stakes Exams in Higher Education: Results of a Field Experiment. PLOS ONE 10(12): e0143616. https://doi.org/10.1371/journal.pone.0143616

[15] C. Y. Piaw, "Replacing paper-based testing with computer-based testing in assessment: Are we doing wrong?." *Procedia-Social and Behavioral Sciences* 64 (2012): 655-664, doi: 10.1016/j.sbspro.2012.11.077

[16] Julie, Gess-Newsome. "A model of teacher professional knowledge and skill including PCK: Results of the thinking from the PCK Summit." *Re-examining pedagogical content knowledge in science education*. Routledge, 2015. 28-42.

[17] Schmid, Mirjam, Eliana Brianza, and Dominik Petko. "Self-reported technological pedagogical content knowledge (TPACK) of pre-service teachers in relation to digital technology use in lesson plans." *Computers in human behavior* 115 (2021): 106586, doi: 10.1016/j.chb.2020.106586

[18] Maas, L.; Brinkhuis, M.J.S.; Kester, L.; Wijngaards-de Meij, L. Cognitive Diagnostic Assessment in University Statistics Education: Valid and Reliable Skill Measurement for Actionable Feedback Using Learning Dashboards. *Appl. Sci.* **2022**, *12*, 4809. doi: 10.3390/app12104809.

[19] A. Barana, M. Marchisio, and M. Sacchet, "Interactive feedback for learning mathematics in a digital learning environment," Educ. Sci., vol. 11, no. 6, art. 279, 2021, doi: 10.3390/educsci11060279.

[20] Rezat, S. How automated feedback from a digital mathematics textbook affects primary students' conceptual development: two case studies. *ZDM Mathematics Education* **53**, 1433–1445 (2021). doi: 10.1007/s11858-021-01263-0.

[21] L. N. Yaqin, L. D. Prasojo, N. A. Haji-Othman, N. Yusof, and A. Habibi, "Addressing the Digital Divide in Indonesian Higher Education: Insights, Implications, and Potential Solutions," in *From Digital Divide to Digital Inclusion*, Singapore: Springer Nature Singapore, 2024, pp. 291–307. doi: 10.1007/978-981-99-7645-4_13

[22] N. Rajiha and S. Rini, "Computer-Based Standardized Testing (CBT) Summative Assessment: One Institution Experience," *JISAE (Journal of Indonesian Student Assessment and Evaluation)*, vol. 9, no. 1, 2023. doi: 10.21009/jisae.v9i1.33233.

[23] G. Kaiser and J. König, "Competence measurement in (mathematics) teacher education and beyond: Implications for policy," High. Educ. Policy, vol. 32, no. 4, pp. 597–615, 2019, doi: 10.1057/s41307-019-00139-z.

[24] T. Kleickmann et al., "Teachers' content knowledge and pedagogical content knowledge: The role of structural differences in teacher education," J. Teacher Educ., vol. 64, no. 1, pp. 90–106, 2013, doi: 10.1177/0022487112460398.

[25] J. Baumert and M. Kunter, "The COACTIV model of teachers' professional competence," in Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers, M. Kunter et al., Eds. New York: Springer, 2013, pp. 25–48, doi: 10.1007/978-1-4614-5149-5_2.

[26] B. Rowan, R. Correnti, and R. J. Miller, “What large-scale, survey research tells us about teacher effects on student achievement: Insights from the Prospects Study of Elementary Schools,” J. Teacher Educ., vol. 104, no. 8, pp. 1525–1567, 2002, doi: 10.1177/00224871211001323.

[27] S. Blömeke, J. E. Gustafsson, and R. J. Shavelson, “Beyond dichotomies: Competence viewed as a continuum,” Z. Psychol., vol. 223, no. 1, pp. 3–13, 2015, doi: 10.1027/2151-2604/a000398.

[28] E. Mohammadpour *et al.*, "A performance-based test to measure teachers' mathematics and science content knowledge (CK) and pedagogical knowledge (PK) of teachers," *Heliyon*, vol. 9, no. 3, 2023, doi: 10.1007/s10763-021-10198-6.

[29] T. Clark and H. Endres, "Computer-based diagnostic assessment of high school students' grammar skills with automated feedback – an international trial," *Assessment in Education: Principles, Policy & Practice*, vol. 28, no. 5-6, pp. 602–632, 2021, doi: 10.1080/0969594X.2021.1970513

[30] G. A. Nortvedt and N. Buchholtz, “Assessment in mathematics education: Responding to issues regarding methodology, policy, and equity,” ZDM Math. Educ., vol. 50, no. 4, pp. 555–570, 2018, doi: 10.1007/s11858-018-0932-3.

Bukti Revisi I

Journal of Mathematics Instruction, Social Research and Opinion
Vol. xx, No. x, March/June/September/December 202x, pp. 1-1x, https://doi.org/
ISSN 2962-7842

Assessing Mathematics Pre-Service Teachers' Content Knowledge and Pedagogical Content Knowledge through a Diagnostic Computer-Based Test

Article Info

Article history:

Received mm dd, yyyy
Revised mm dd, yyyy
Accepted mm dd, yyyy

Keywords:

Computer-based test
Content knowledge
Pedagogical content knowledge
Mathematics pre-service teachers
Diagnostic assessment
Mathematical knowledge for teaching

ABSTRACT

Ensuring an adequate professional-knowledge profile of mathematics pre-service teachers is a strategic issue for teacher education. However, Content Knowledge (CK) and Pedagogical Content Knowledge (PCK) are rarely measured simultaneously through technology-based diagnostic assessment in the Indonesian context. This study aims to measure and profile the CK and PCK of mathematics pre-service teachers using a diagnostic Computer-Based Test (CBT). A quantitative descriptive-correlational design was applied to 30 fifth- and seventh-semester students of a Mathematics Education (Tadris Matematika) study program at a teacher education institution (LPTK), selected purposively. Each construct was measured with 25 multiple-choice items whose content validity was established through expert judgment and whose reliability was confirmed before use (KR-20 = 0.81 for CK and 0.79 for PCK). The data were analyzed using descriptive statistics, Pearson correlation, and simple linear regression after the underlying assumptions had been checked. The ideal competency threshold was set at 75%, a researcher-defined mastery criterion consistent with mastery-learning standards. The results show mean scores of 50.00% for CK and 52.67% for PCK, both below the 75% threshold, with 60.00% of the students in the low category; Specialized Content Knowledge (SCK, 36.91%) and Knowledge of Content and Curriculum (KCC, 46.25%) were the weakest dimensions. CK and PCK were strongly and positively correlated ($r = 0.72$; $p < 0.001$), and CK accounted for about 52% of the variance in PCK as a predictive, non-causal association. Diagnostic analysis at the dimension, indicator, and individual-student levels shows that the CBT produces specific and actionable profiles. These findings underscore the need for data-driven LPTK curriculum reform that integrates content mastery with pedagogical capacity.

Corresponding Author:

INTRODUCTION

The quality of a nation's mathematics education depends heavily on teachers' professional competence. Cross-national studies have shown that mathematics teachers' professional knowledge, comprising Content Knowledge (CK) and Pedagogical Content

Knowledge (PCK), is a primary predictor of instructional effectiveness and students' academic achievement [1], [2]. The 2022 PISA results placed Indonesian students' mathematics literacy 68th out of 81 countries, below the OECD average, indicating a fundamental problem in the quality of mathematics instruction [3]. This condition is inseparable from the quality of professional knowledge formed in teacher education institutions (LPTK), so ensuring an adequate professional-knowledge profile of pre-service teachers becomes a strategic imperative [4], [5].

Theoretically, the Mathematical Knowledge for Teaching (MKT) framework distinguishes CK—comprising common, specialized, and horizon content knowledge—from PCK, which encompasses knowledge of content and students, content and teaching, and content and curriculum [6], [7]. Systematic reviews affirm that PCK is the most frequently studied yet the most difficult component to measure consistently because of its contextual and multidimensional nature [8], [9]. Studies in Indonesia show that the professional knowledge of mathematics pre-service teachers varies widely across institutions and curricula, with a notable weakness in the PCK aspect [10], [11]. Similar findings have been reported in cross-national contexts [12], [13].

Alongside the rapid digitalization of education, competency-evaluation systems are now shifting massively from traditional approaches toward technology-based assessment. Various studies indicate that CBT is often more valid and reliable than conventional examination methods [14], [15]; however, its use to objectively measure both the content knowledge and the pedagogical content knowledge of pre-service teachers simultaneously remains very limited because of the complexity of these constructs [16], [17]. The potential of CBT to generate diagnostic feedback based on learning analytics has also not been optimally exploited [18], even though interactive digital feedback has been shown to improve mathematical understanding [19], [20].

The urgency of this study is heightened in the context of Indonesia's uneven higher-education digital transformation [21], which is not yet supported by adequate standardization of assessment systems [22]. Measuring teacher competence in the digital era is required to produce data that are actionable for study-program managers and policymakers [23]. A number of studies affirm that CK and PCK are relatively independent constructs that require different intervention pathways [24], [25], and that teacher effects on student achievement are substantial [26].

The novelty of this study lies in three aspects. First, theoretically, the study integrates the MKT framework and technology-based assessment within a single coherent quantitative measurement design to identify CK and PCK profiles simultaneously [27], [28]. Second, methodologically, the CBT is used as an autonomous, automated assessment medium that generates layered diagnostic feedback—at the total, dimension, indicator, and individual-student levels—so that test results can be acted upon directly for instructional improvement [29]. Third, contextually, the study targets students at a religiously affiliated LPTK in Indonesia, a context that—although not entirely absent from the literature—has received only limited attention in in-depth, technology-based diagnostic studies [30]. This study aims to: (1) measure students' CK and PCK achievement and profiles; (2) identify strengths and weaknesses in each dimension and indicator; and (3) demonstrate the use of diagnostic profiles—including at the individual-student level—as a basis for designing instructional interventions.

METHOD

This study employed a quantitative approach with a descriptive-survey and correlational design. The population comprised students of the Mathematics Education (Tadris Matematika) study program at an LPTK who were in their fifth and seventh semesters. A sample of 30 students was determined through purposive sampling with the inclusion criteria: (1) active students who had adequately completed mathematics-content and pedagogy courses; and (2) willing to take the entire series of CBT tests. Given the limited sample size, all data were analyzed as a single intact group (n = 30) and were not directed toward inter-semester comparison.

Research setting and participants. The research was conducted in the odd semester of the 2024/2025 academic year at the Mathematics Education (Tadris Matematika) study program of a state Islamic teacher education institution (LPTK) in Indonesia. Because every eligible student who met the inclusion criteria (having completed the core mathematics-content and pedagogy courses) took part, the 30 participants were treated as a single intact group; this modest sample size is acknowledged as a constraint on statistical power and on the generalizability of the correlational and regression results, as elaborated in the Discussion.

The operationalization of the variables comprised two main constructs. CK was defined as mastery of the mathematics content to be taught, measured through 25 multiple-choice items; PCK was defined as the ability to integrate content knowledge with pedagogical strategies, also measured through 25 multiple-choice items. Scores were converted into percentages of achievement relative to the ideal maximum score. The CBT instrument was designed to generate layered diagnostic reports, namely the total score, achievement per dimension, achievement per indicator, and an individual achievement profile for each student. Each item was mapped to a specific indicator and dimension (see the notes below Table 1) so that the CBT system could produce achievement per dimension and per indicator for all students at once, as well as their individual diagnostic profiles. The composition of the instrument is presented in Table 1.

Table 1. Composition of CBT Instrument Items by MKT Dimension

Domain	Dimension	Number of Items
CK	Common Content Knowledge (CCK)	7
CK	Specialized Content Knowledge (SCK)	14
CK	Horizon Content Knowledge (HCK)	4
PCK	Knowledge of Content and Students (KCS)	10
PCK	Knowledge of Content and Teaching (KCT)	9
PCK	Knowledge of Content and Curriculum (KCC)	6
Total		50

Notes on dimension codes and indicator coverage (based on the developed CK and PCK instruments). **CK – CCK** (Common Content Knowledge, items 1–7): understanding basic concepts (e.g., lines and angles), performing calculations and measurements accurately, and applying mathematical concepts in everyday life. **CK – SCK** (Specialized Content Knowledge, items 8–21): explaining concepts and principles, analyzing and connecting mathematical objects, proving properties, building models, and solving complex mathematical problems. **CK – HCK** (Horizon Content Knowledge, items 22–25): connecting school mathematics with advanced academic mathematics and with other disciplines. **PCK – KCS** (Knowledge of Content and Students, items 1–10): analyzing students' preconceptions and understanding as well as predicting and analyzing students' misconceptions, errors, and difficulties. **PCK – KCT** (Knowledge of Content and Teaching, items 11–19): selecting and organizing learning materials, selecting sources, strategies, and techniques, and selecting and analyzing representations for presenting material. **PCK – KCC** (Knowledge of Content and Curriculum, items 20–25): designing content and topic combinations, applying standards and curricula, and designing instruments to measure students' abilities.

Instrument development and content validity. The CBT instrument was developed through a systematic procedure. First, a test blueprint was constructed by mapping each item to a specific MKT dimension and indicator (Table 1). Second, the draft items were reviewed by three experts—two mathematics-education lecturers and one educational-assessment specialist—to establish content validity; the level of expert agreement was quantified using Aiken's V, and items with $V \geq 0.80$ were retained while the remaining items were revised or replaced. Third, the revised items were trialed and analyzed empirically before being administered in the present study.

Reliability and item analysis. Based on the item analysis, the instrument demonstrated adequate psychometric quality. Internal-consistency reliability computed with the Kuder–Richardson formula (KR-20) reached 0.81 for the CK test and 0.79 for the PCK test, both within the acceptable-to-good range. The item difficulty indices ranged from 0.20 to 0.87 (a balanced mix of easy, medium, and hard items), and the discrimination indices were largely satisfactory ($D \geq 0.30$); only items meeting the validity, difficulty, and discrimination criteria were retained in the final CBT.

Item format. Multiple-choice items were chosen because the CBT was designed for automated, standardized, and scalable diagnosis across many participants and indicators simultaneously. To capture pedagogical reasoning rather than mere recall, the PCK items were built around short classroom scenarios, samples of student work, and instructional-decision situations, with distractors constructed from typical student misconceptions and common teaching errors. It is nevertheless acknowledged that the multiple-choice format cannot fully capture the depth of pedagogical reasoning that open-ended or performance-based tasks would reveal, a limitation revisited in the Discussion.

The analysis techniques comprised descriptive and inferential analyses. Descriptive analysis produced the mean (M), standard deviation (SD), minimum, maximum, median, and percentage of achievement at the total, dimension, indicator, and individual levels, as well as a three-level categorization (High, Medium, Low). The categorization used the following cut-off points: High (≥ 75%), Medium (60–74.99%), and Low (< 60%), where the 75% mark represents the ideal mastery criterion adopted in this study. Inferential analysis was conducted through two techniques: (1) Pearson correlation to test the CK–PCK relationship and the corresponding inter-dimension relationships (CCK–KCS, SCK–KCT, HCK–KCC) at $\alpha = 0.05$; and (2) simple linear regression with CK as the predictor and PCK as the criterion. Prior to the inferential tests, the underlying assumptions were examined, namely the normality of the CK and PCK distributions (Shapiro–Wilk), the linearity of their relationship, the absence of influential outliers, and homoscedasticity; all assumptions were adequately met, so that the parametric analyses were appropriate.

All scoring, achievement computation, correlation, and linear regression were performed using Microsoft Excel through standardized formulas (including CORREL, SLOPE, INTERCEPT, and RSQ) so that every figure can be traced back to the raw data. The study also observed standard research-ethics principles: participation was entirely voluntary, informed consent was obtained from all respondents before data collection, and the participants' anonymity and confidentiality were maintained throughout the recording, processing, and reporting of data by using codes instead of names, with the individual-profile example presented under the pseudonym Student 1.

RESULTS AND DISCUSSION

3.1. Results

This section reports the findings of the data analysis in the following order: respondent characteristics, descriptive statistics, the distribution of achievement categories, dimension- and indicator-level achievement, the correlation and regression results, and an individual diagnostic profile. The 30 respondents were fifth- and seventh-semester students of the Mathematics Education (Tadris Matematika) study program who had completed the core mathematics-content and pedagogy courses; all of them completed the full series of CK and PCK tests through the CBT system.

Overall, students' professional-knowledge achievement had not yet reached the ideal competency threshold (≥ 75%). The mean CK score was 50.00% (SD = 10.16), PCK 52.67% (SD = 23.96), and the combined Professional Knowledge (PK) score 51.33% (SD = 16.04). The complete descriptive statistics are presented in Table 2.

Table 2. Descriptive Statistics of CK, PCK, and Professional Knowledge Scores (n = 30)

Variable	Mean (%)	SD	Min	Max	Median
Content Knowledge (CK)	50.00	10.16	24	68	52.00
Pedagogical Content Knowledge (PCK)	52.67	23.96	8	96	54.00
Professional Knowledge (PK)	51.33	16.04	26	78	50.00

Table 2 shows that the mean PCK is slightly higher than CK (a difference of 2.67 points), an interesting pattern indicating that pedagogical-transformation capacity slightly exceeds content mastery. The large standard deviation of PCK (SD = 23.96) and the much smaller one for CK (SD = 10.16) reflect substantial heterogeneity in knowledge profiles, with score ranges of 24–68% for CK and 8–96% for PCK. The distribution of Professional Knowledge categories is presented in Table 3.

Table 3. Distribution of Professional Knowledge Categories (n = 30)

Category	Frequency (f)	Percentage (%)
High	0	0.00
Medium	12	40.00
Low	18	60.00
Total	30	100.00

All students (100%) were in the Low or Medium category, with 60.00% in the Low category and 40.00% in the Medium category, and no student reaching the High category. This indicates that the majority of pre-service teachers have not attained an optimal level of professional-knowledge mastery. Achievement per dimension is presented in Table 4 and Figure 1.

Table 4. Average Achievement per CK and PCK Dimension

Domain	Dimension	Items	Mean (%)	SD
CK	CCK	7	73.33	20.47
CK	SCK	14	36.91	24.16
CK	HCK	4	55.00	35.60
PCK	KCS	10	54.07	33.18
PCK	KCT	9	57.50	26.14
PCK	KCC	6	46.25	32.30

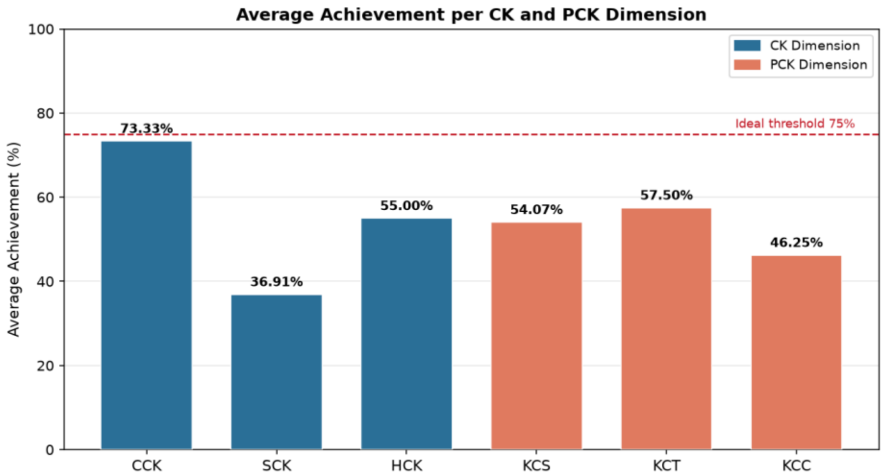


Figure 1. Average Achievement per CK and PCK Dimension

In the CK domain, the CCK dimension was the highest (73.33%) and SCK the lowest (36.91%); in the PCK domain, KCT was the highest (57.50%) while KCC was the lowest of all six dimensions (46.25%). Diagnostically, the low SCK and KCC signal a weak ability to relate content to the structure and organization of the curriculum—an area that needs to be prioritized in intervention. Achievement at the indicator level is presented in Table 5 and visualized in Figure 2.

Table 5. Average Achievement per CK and PCK Indicator

Domain	Indicator	Mean (%)
CK	CCK1	86.67
CK	CCK2	80.00
CK	CCK3	46.67
CK	SCK1	63.33
CK	SCK2	46.67
CK	SCK3	43.33
CK	SCK4	20.00
CK	SCK5	23.33
CK	SCK6	30.00
CK	SCK7	33.33
CK	HCK1	53.33
CK	HCK2	56.67
PCK	KCS1	58.33
PCK	KCS2	48.33
PCK	KCS3	58.33
PCK	KCS4	51.67
PCK	KCS5	53.33
PCK	KCT1	55.00
PCK	KCT2	66.67
PCK	KCT3	60.00
PCK	KCT4	36.67

PCK	KCC1	51.67
PCK	KCC2	43.33
PCK	KCC3	45.00

Notes on indicator codes (based on the CK and PCK instruments). CK indicators — **CCK1**: understanding basic school-mathematics concepts, such as lines, angles, sequences, and basic statistics (items 1–3); **CCK2**: performing calculations and measurements accurately (items 4–5); **CCK3**: applying mathematical concepts in everyday life (items 6–7); **SCK1**: explaining academic mathematical concepts and principles (items 8–9); **SCK2**: analyzing mathematical objects (items 10–11); **SCK3**: connecting mathematical properties and objects (items 12–13); **SCK4**: proving mathematical properties or theorems (items 14–16); **SCK5**: applying proof strategies (item 17); **SCK6**: building and constructing mathematical models (items 18–19); **SCK7**: solving complex mathematical problems (items 20–21); **HCK1**: connecting school mathematics with advanced academic mathematics (items 22–23); **HCK2**: connecting mathematics with other disciplines (items 24–25). PCK indicators — **KCS1**: analyzing students' preconceptions (items 1–2); **KCS2**: identifying students' level of understanding (items 3–4); **KCS3**: predicting students' misconceptions (items 5–6); **KCS4**: identifying sources of students' learning difficulties (items 7–8); **KCS5**: analyzing students' literacy issues and errors (items 9–10); **KCT1**: organizing and sequencing learning materials (items 11–12); **KCT2**: selecting learning sources, strategies, and techniques (items 13–14); **KCT3**: selecting representations and illustrations for presenting material (items 15–17); **KCT4**: analyzing the weaknesses and strengths of representations (items 18–19); **KCC1**: composing content and topic combinations (items 20–21); **KCC2**: applying standards and curricula (items 22–23); **KCC3**: designing instruments to measure students' abilities (items 24–25).

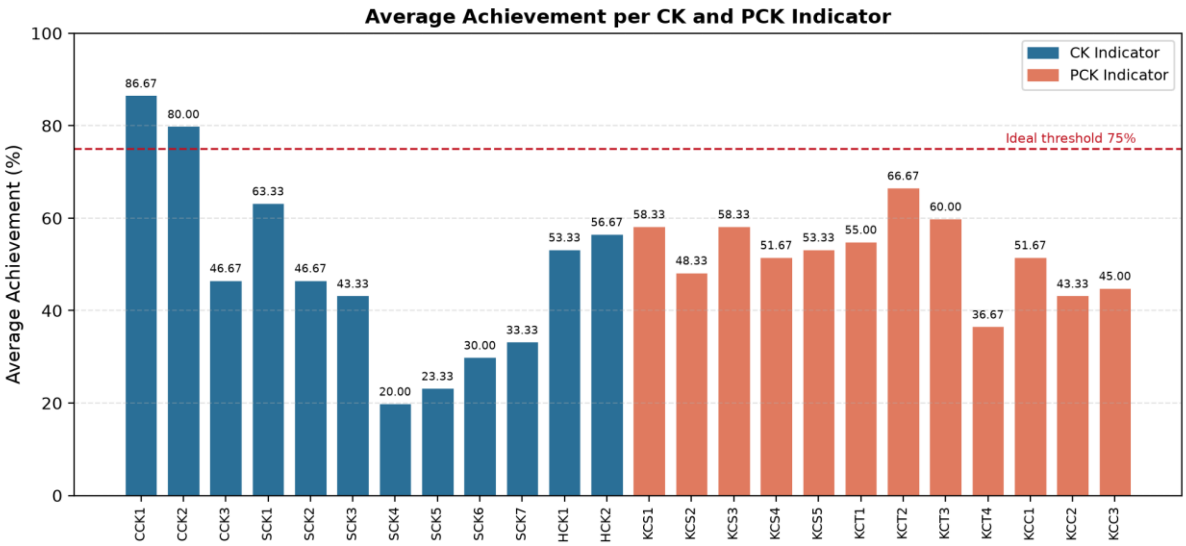


Figure 2. Average Achievement per CK and PCK Indicator

Figure 2 visualizes the average achievement of each indicator exactly as listed in Table 5. The range of achievement across indicators spans 20.00% to 86.67%. The highest achievement is on the CK indicator CCK1 (86.67%; understanding basic school-mathematics concepts), which is fundamental and therefore the most widely mastered by students. Conversely, the lowest overall achievement is on indicator SCK4 (20.00%; proving mathematical properties or theorems)—the skill with the highest cognitive demand in the CK domain—followed by SCK5 (23.33%; applying proof strategies) and SCK6 (30.00%; building and constructing mathematical models). The pattern in the SCK dimension declines gradually with complexity, from explaining concepts and principles (SCK1; 63.33%), analyzing and connecting mathematical objects (SCK2 46.67%; SCK3 43.33%), solving complex problems (SCK7; 33.33%), to the most difficult modeling and proving (SCK6 30.00%; SCK5 23.33%; SCK4 20.00%)—a pattern consistent with cognitive load theory. Because each dimension's achievement remains the mean of its constituent indicators—for example, KCS (54.07%) is the mean of KCS1–KCS5 and CCK (73.33%) the mean of CCK1–CCK3—this pattern maps precisely which items and indicators are the sources of weakness, so that lecturers can design remediation targeted at concrete rather than generic weaknesses. The results of the Pearson correlation test are presented in Table 6.

Table 6. Pearson Correlation between Domains and between Parallel Dimensions

Relationship	r	p	Interpretation
CK – PCK	0.72	< 0.001	Strong
CCK – KCS	0.42	< 0.05	Moderate
SCK – KCT	0.54	< 0.01	Moderate
HCK – KCC	0.86	< 0.001	Very strong

The strong positive correlation between CK and PCK ($r = 0.72$; $p < 0.001$) confirms that the two domains are constructively interdependent, reinforced by the correlations between parallel dimensions, especially HCK–KCC ($r = 0.86$) and SCK–KCT ($r = 0.54$). The predictive relationship of CK on PCK was tested through simple linear regression (Table 7 and Figure 3).

Table 7. Results of the Simple Linear Regression of CK on PCK

Parameter	Value
Slope (b)	1.708
Intercept (a)	−32.734
Coefficient of determination (R^2)	0.5245
Pearson r	0.72
Equation	$PCK = 1.708 \times CK - 32.734$

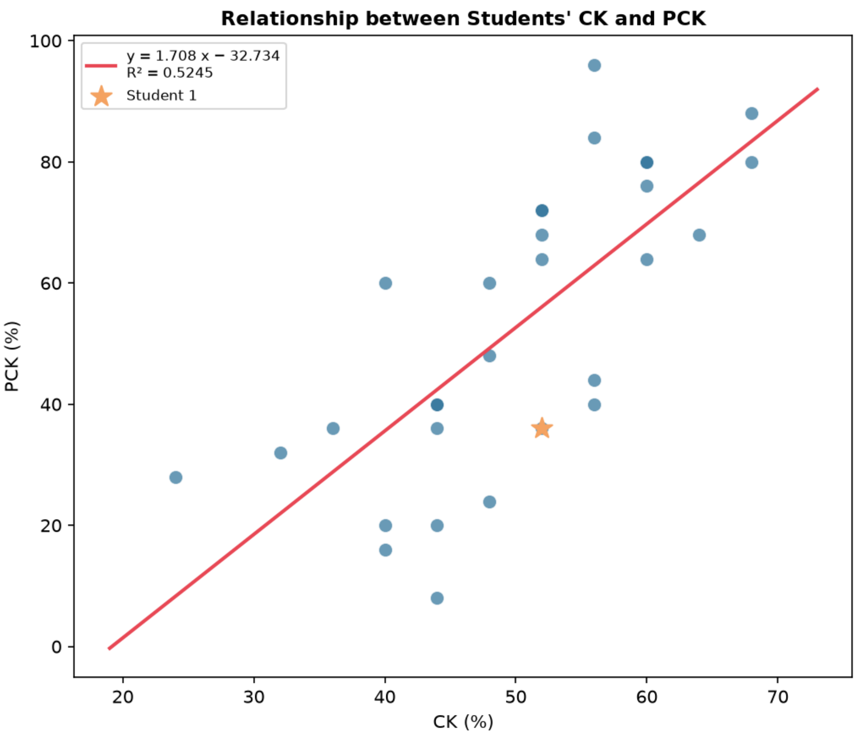


Figure 3. Scatter Plot and Regression Line of the CK–PCK Relationship

The equation $PCK = 1.708 \times CK - 32.734$ with $R^2 = 0.52$ indicates that about 52% of the variance in PCK is statistically associated with CK. Because the design is correlational, this relationship should be interpreted as a predictive association rather than a causal effect; the remaining 48% of the variance signals that PCK does not form automatically from content mastery but requires an explicit pedagogical learning pathway.

The main advantage of the CBT instrument lies in its ability to break down aggregate achievement into individual diagnostic profiles automatically. As an illustration, Table 8 and Figure 4 summarize the achievement of one student disguised as Student 1, while Figure 5 and Figure 6 display authentic screenshots of the diagnostic scoring report generated directly by the CBT system for Student 1. Such individual reports are produced automatically by the system for every participant without any manual scoring.

Table 8. Individual Achievement Profile of Student 1 Compared with the Class Average

Domain	Student 1 Score (%)	Class Average (%)	Difference	Category
Content Knowledge (CK)	52.00	50.00	+2.00	Low
Pedagogical Content Knowledge (PCK)	36.00	52.67	−16.67	Low

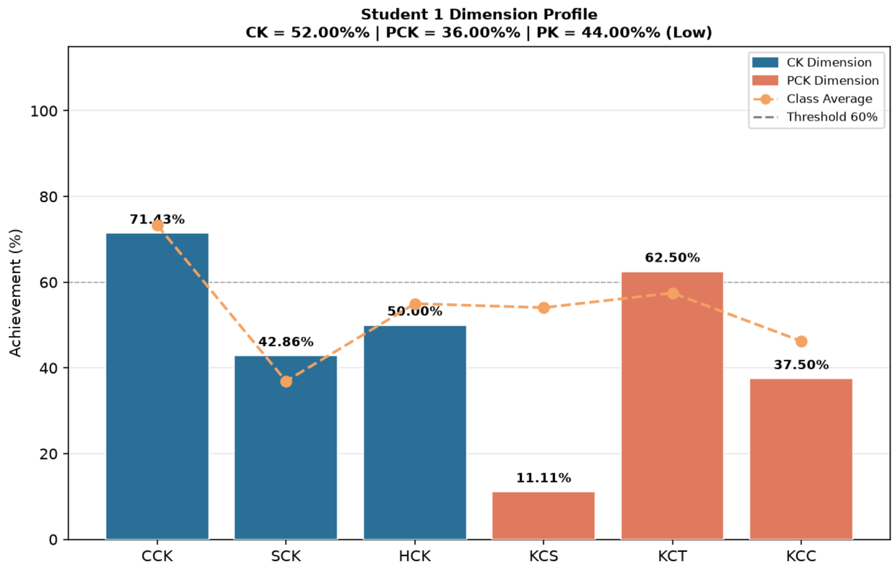


Figure 4. CK and PCK Achievement Profile of Student 1 Compared with the Class Average

At the individual level, the CBT system automatically scores the answers and generates a diagnostic scorecard for each participant. Student 1, for example, answered 13 of 25 CK items correctly (52.00%) and 9 of 25 PCK items correctly (36.00%); the CK achievement was slightly above the class average (a difference of +2.00 points) while the PCK achievement was far below it (a difference of −16.67 points), and neither reached the ideal competency threshold (75%), placing the student in the Low category for both domains. Authentic screenshots of the report are presented in Figure 5 (CK) and Figure 6 (PCK). The system-generated screenshots were anonymized so that no identifiable student information is displayed, and the pseudonym Student 1 is used throughout.

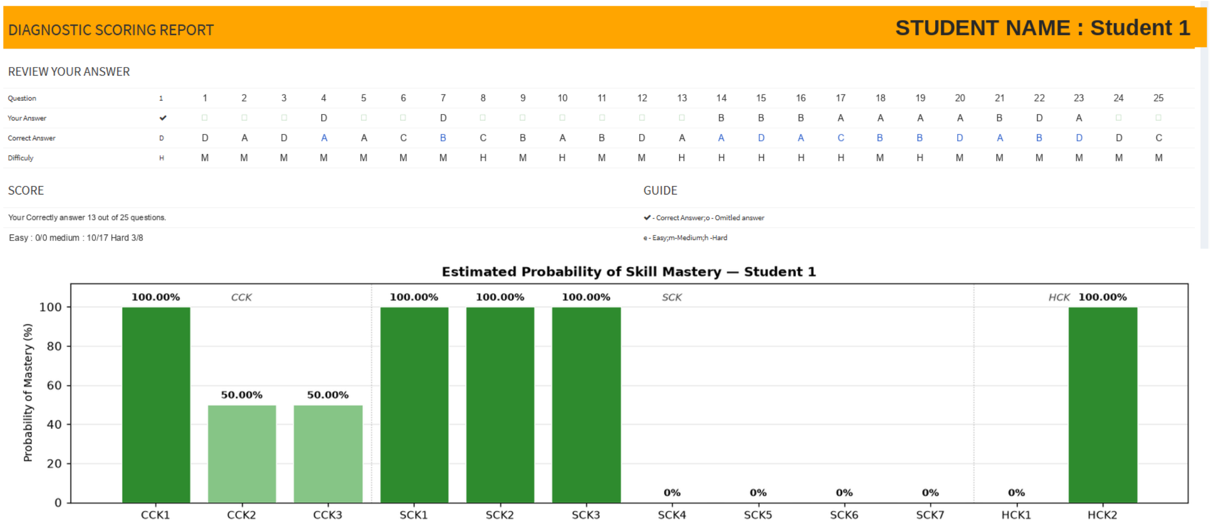


Figure 5. CBT Diagnostic Scorecard for the CK Domain for Student 1 (original system screenshot)

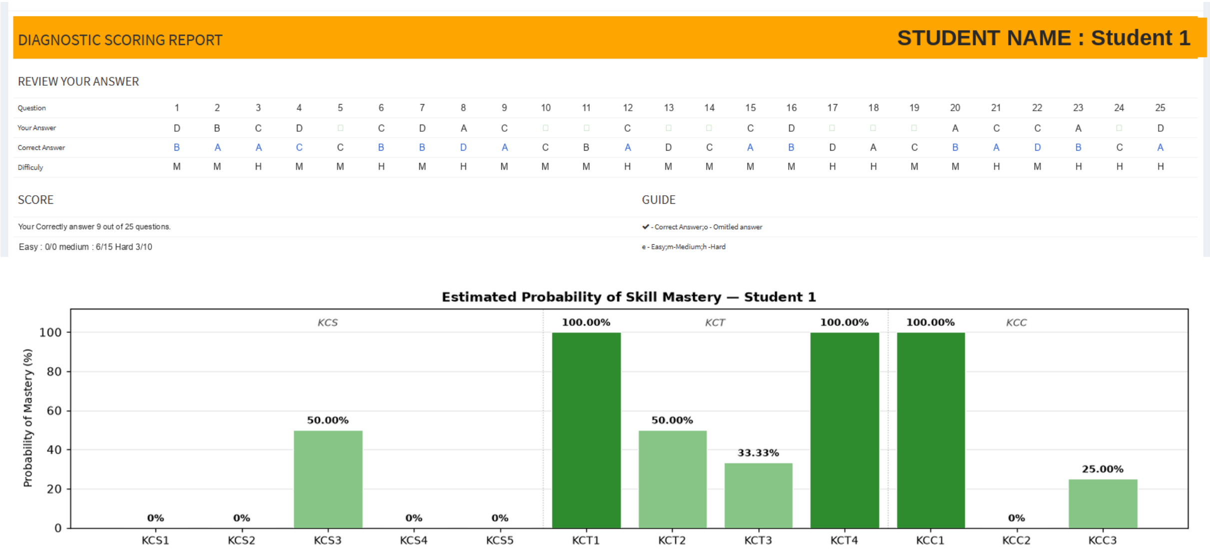


Figure 6. CBT Diagnostic Scorecard for the PCK Domain for Student 1 (system screenshot)

Figures 5 and 6 reveal diagnostic functions of the CBT that are more difficult to implement efficiently in conventional paper-based tests. In addition to recapping correct/incorrect answers per item along with their difficulty level, a combination of the Medium (M) and Hard (H) categories according to the item analysis of the instrument, the system presents an Estimated Probability of Skill Mastery graph that maps Student 1's mastery probability for each skill indicator (CCK1–HCK2 for CK and KCS1–KCC3 for PCK), complete with a mapping of sample items to each indicator. Thus, the CBT does not merely provide a final score but directly identifies which indicators have been mastered and which remain weak, in Student 1's case, several indicators show a mastery probability close to zero, so that lecturers can follow up with specific and immediate remediation. It is this automation of scoring, standardization of reporting, and diagnostic depth that affirms the added value of the CBT in assessing the professional knowledge of mathematics pre-service teachers.

Furthermore, the two scorecards allow Student 1's achievement to be traced down to the indicator level, not stopping at dimension averages. In the CK domain (Figure 5), Common Content Knowledge (CCK) was the strongest dimension (71.43%; 5 of 7 items): understanding basic school-mathematics concepts (CCK1) was fully mastered (100.00%), whereas performing calculations and measurements (CCK2) and applying mathematical concepts in everyday life (CCK3) were each only partially mastered (50.00%). Specialized Content Knowledge (SCK) stood at 42.86% (6 of 14 items) with a highly contrasting pattern: three basic indicators—explaining academic mathematical concepts and principles (SCK1), analyzing mathematical objects (SCK2), and connecting mathematical properties and objects (SCK3)—were all fully mastered (100.00%), while four indicators demanding higher-order reasoning—proving mathematical properties or theorems (SCK4), applying proof strategies (SCK5), building and constructing mathematical models (SCK6), and solving complex mathematical problems (SCK7)—were all at 0.00%, marking a total deficit in advanced mathematical reasoning. Horizon Content Knowledge (HCK) stood at 50.00% (2 of 4 items): connecting mathematics with other disciplines (HCK2) was fully mastered (100.00%), but connecting school mathematics with advanced academic mathematics (HCK1) remained at 0.00%. In the PCK domain (Figure 6), the indicator profile was also uneven within each dimension. Knowledge of Content and Teaching (KCT; dimension mean 62.50%) was the strongest: organizing and sequencing learning materials (KCT1) and analyzing the weaknesses and strengths of representations (KCT4) were fully mastered (100.00%), followed by selecting learning sources, strategies, and techniques (KCT2; 50.00%) and selecting representations and illustrations for presenting material (KCT3; 33.33%). Knowledge of Content and Curriculum (KCC; dimension mean 37.50%) rested on composing content and topic combinations (KCC1), which was fully mastered (100.00%), while designing instruments to measure students' abilities (KCC3) was only 25.00% and applying standards and curricula (KCC2) was still 0.00%. Conversely, Knowledge of Content and Students (KCS) was the weakest PCK dimension and the weakest point of the entire profile (11.11%): only predicting students' misconceptions (KCS3) was partially mastered (50.00%), while analyzing students' preconceptions (KCS1), identifying students' level of understanding (KCS2), identifying sources of students' learning difficulties (KCS4), and analyzing students' literacy issues and errors (KCS5) were all at 0.00%. This pattern confirms that Student 1's weaknesses are concentrated in a number of indicators that are genuinely untouched—SCK4, SCK5, SCK6, SCK7, and HCK1 in the CK domain and KCS1, KCS2, KCS4, KCS5, and KCC2 in the PCK domain, all at 0.00%—so that these ten indicators are the most urgent remediation priorities, while the indicators already fully mastered (CCK1, SCK1, SCK2, SCK3, HCK2, KCT1, KCT4, and KCC1) can serve as a foundation for reinforcement. Thus, indicator-based diagnosis provides a far sharper and more specific direction for individual intervention than reading dimension averages alone.

Summary of key findings. In sum, both CK (50.00%) and PCK (52.67%) fell below the 75% mastery threshold; 60.00% of the students were in the low category and none reached the high category; SCK (36.91%) and KCC (46.25%) were the weakest dimensions; CK and PCK were strongly and positively associated ($r = 0.72$; $p < 0.001$); and the CBT was able to decompose these results down to the indicator and individual-student levels, as illustrated by the profile of Student 1.

3.2. Discussion

The main findings reveal three key facts. **First**, neither CK (50.00%) nor PCK (52.67%) reached the minimum competency threshold; interestingly, PCK was even slightly higher than CK, reflecting the knowledge-transformation challenge articulated by Shulman [7] and reinforced in the MKT framework [6]. **Second**, the SCK dimension was the weakest point of CK (36.91%) and KCC the weakest point of PCK (46.25%), indicating a double deficit in the mastery of advanced mathematical content and curriculum, which in the COACTIV model strongly determines the quality of instructional planning [2], [25]. **Third**, the lower standard deviation of CK (10.16) compared with PCK (23.96) shows that CK achievement is more uniform, whereas PCK is highly heterogeneous, indicating large variation in students' pedagogical capacity.

The significant positive correlations, especially HCK–KCC ($r = 0.86$) and SCK–KCT ($r = 0.54$), show that content mastery in particular dimensions goes hand in hand with the relevant pedagogical capacity. However, the regression equation ($R^2 = 0.52$) affirms that content mastery explains only part of the variance in PCK, consistent with findings that pre-service teachers in developing countries tend to have lower PCK scores than CK because LPTK curricula place greater emphasis on content mastery [1], [10]. The case of Student 1 reinforces this at the individual level: Student 1 (CK = 52%, PCK = 36%) falls into the Low category with a highly uneven profile, namely CK achievement slightly above the class average while PCK is far below it. In the PCK domain, KCT is relatively the highest (62.50%) while KCS is the lowest (11.11%) and KCC is low (37.50%), indicating an ability to select teaching strategies that is not yet accompanied by an adequate understanding of student characteristics or curriculum mastery. Intervention needs to target KCS and KCC simultaneously, especially the ability to analyze preconceptions, identify levels of understanding, and identify sources of students' learning difficulties.

These results are broadly consistent with prior Indonesian studies reporting that the professional knowledge of mathematics pre-service teachers is generally moderate and uneven, with PCK frequently emerging as the more fragile component [10], [11]. The finding that PCK was marginally higher than CK is noteworthy and can be interpreted cautiously: the PCK items, which draw on familiar classroom situations, may be more accessible to students than the CK items, which demand advanced and abstract mathematical reasoning; moreover, the far larger dispersion of PCK (SD = 23.96) indicates that this apparent advantage is unevenly distributed across the cohort. The particularly low SCK (36.91%) is plausibly linked to teacher-education curricula that emphasize procedural and elementary content over advanced mathematical reasoning—such as formal proof, mathematical modeling, and complex problem solving—so that the more cognitively demanding indicators (SCK4–SCK7) remain underdeveloped [24], [25].

Theoretically, this study validates and extends the application of the MKT framework and the COACTIV model to the context of religiously affiliated LPTKs in Indonesia, which are still underrepresented in the indexed literature [6], [25], [11]. Methodologically, the study demonstrates a CBT model that produces layered diagnostic profiles—from the total, dimension, and indicator levels to the individual student—thus addressing the previously identified gap in using CBT to measure pedagogical competence [22], [28]. Achievement profiles that can be decomposed to the individual level, as shown in the case of Student 1, strengthen the function of the CBT as an instrument that generates actionable diagnostic feedback [18]. In policy terms, the mean PK of 51.33% and the absence of any student in the High category provide urgent empirical evidence to drive data-based curriculum reform rather than normative assumptions [5], [23], particularly in the context of Indonesia's 2022 PISA results [3].

This study opens a new perspective by positioning the CBT not merely as an examination tool but as a technology-based diagnostic-assessment ecosystem that produces rich cognitive data, decomposed per dimension and per individual, and actionable for program managers [23], [30]. Several limitations should nonetheless be noted, and they caution against generalizing these findings beyond the single institution and the 30 participants studied here: (a) the relatively small sample size ($n = 30$), which limits statistical power and generalizability; (b) the use of purposive sampling; (c) the single-institution context; (d) the still-bivariate scope of the analysis, which has not modeled other determinants simultaneously; (e) the absence of qualitative data to explain why students performed poorly; (f) the limited evidence of the CBT instrument's validity and reliability beyond the internal analyses reported here; and (g) the inherent constraint of the multiple-choice format, since PCK often requires the analysis of classroom scenarios, student work, teaching decisions, and pedagogical explanations that selected-response items cannot fully capture. Building on these limitations, future research should employ larger and more representative multi-institution samples, longitudinal and mixed-methods designs, and multivariate models to test additional determinants and to strengthen both internal and external validity.

CONCLUSION AND RECOMMENDATIONS

This study set out to measure and profile the CK and PCK of mathematics pre-service teachers at an LPTK through a diagnostic CBT. Both mean CK (50.00%) and PCK (52.67%) fell below the ideal competency threshold (75%), with 60.00% of the students in the low category and none in the high category; CCK was the strongest CK dimension (73.33%), whereas SCK (36.91%) and KCC (46.25%) were the weakest. CK and PCK were strongly and positively associated ($r = 0.72$; $p < 0.001$), with CK accounting for about 52% of the variance in PCK as a predictive, non-causal association. The layered diagnostic analysis—down to the individual level, as illustrated by Student 1—shows that the CBT can generate specific, actionable achievement profiles. Theoretically, the study extends the MKT and COACTIV frameworks to the Indonesian LPTK context; practically, it offers a standardized and efficient diagnostic instrument for teacher education. These conclusions should be read in light of the study's limitations—a small, single-institution, purposively selected sample and a bivariate, multiple-choice-based design—which future research should address through larger multi-institution samples, longitudinal and mixed-methods designs, and multivariate modeling.

Recommendations. Based on these findings, the following recommendations are organized by stakeholder group.

For lecturers. Remedial instruction should target not only the weakest dimensions (SCK in CK and KCC in PCK) but also the weakest indicators identified in Table 5. In the CK domain, the priority is strengthening mathematical proof—proving properties or theorems (SCK4; 20.00%) and applying proof strategies (SCK5; 23.33%)—through scaffolded proof practice, analysis of proof errors, and habituation to deductive argumentation, followed by mathematical modeling (SCK6; 30.00%) and complex problem solving (SCK7; 33.33%) via problem-based learning. In the PCK domain, the focus is on analyzing the strengths and weaknesses of representations (KCT4; 36.67%), applying standards and curricula (KCC2; 43.33%), and designing measurement instruments (KCC3; 45.00%) through case-based learning, analysis of teaching videos, and curriculum-document study. The same indicator-targeting logic applies at the individual level, as shown in the profile of Student 1.

For LPTK curriculum managers. The results encourage integrating a diagnostic CBT—capable of mapping achievement down to the indicator and individual levels—into the study program's quality-assurance cycle, and rebalancing the curriculum so that advanced mathematical reasoning and pedagogical-content competence receive as much emphasis as procedural content mastery.

For future researchers. Subsequent studies are encouraged to use larger and more representative multi-institution samples, longitudinal designs, mixed-methods approaches, and multivariate models involving additional determinants (such as academic background, learning motivation, and the quality of lecturers' teaching) in order to strengthen external validity and to explain the causes of the observed weaknesses.

For CBT system development. The CBT itself can be enhanced by adding automatic item-level feedback with worked explanations, direct links to remedial learning materials for each weak indicator, and adaptive item delivery, so that the system functions not only as a diagnostic tool but also as an integrated learning-support environment.

ACKNOWLEDGMENTS

The authors thank the Mathematics Education (Tadris Matematika) study program of the LPTK for facilitating the implementation of the research, as well as all the students who participated as respondents.

REFERENCES

[1] X. Yang and G. Kaiser, “The impact of mathematics teachers’ professional competence on instructional quality and students’ mathematical learning outcomes,” Teach. Teach. Educ., vol. 110, art. 103582, 2022, doi: 10.1016/j.tate.2021.103582

[2] S. Krauss et al., “Pedagogical content knowledge and content knowledge of secondary mathematics teachers,” Journal of Educational Psychology, 100(3), 716–725, 2008, doi: 10.1037/0022-0663.100.3.716.

[3] A. Schleicher, PISA 2022 Results: The State of Learning and Equity in Education. Paris: OECD Publishing, 2023.

[4] M. T. Tatto, J. Schwille, S. L. Senk, L. Ingvarson, R. Peck, G. Rowley, K. Bankov, M. Rodriguez, and M. Reckase, “Policy, Practice, and Readiness to Teach Primary and Secondary Mathematics in 17 Countries: Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)”. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement, 2012.

[5] H. Qian and P. Youngs, “The effect of teacher education programs on future elementary mathematics teachers’ knowledge: A five-country analysis using TEDS-M data”. J. Math. Teacher Educ., vol. 19, no. 4, pp. 371–396, 2016, doi: 10.1007/s10857-014-9297-9.

[6] D. L. Ball, M. H. Thames, and G. Phelps, “Content knowledge for teaching: What makes it special?,” J. Teacher Educ., vol. 59, no. 5, pp. 389–407, 2008, doi: 10.1177/0022487108324554.

[7] L. S. Shulman, “Those who understand: Knowledge growth in teaching,” Educ. Res., vol. 15, no. 2, pp. 4–14, 1986, doi: 10.3102/0013189X015002004.

[8] M. Grigaliūnienė, E. Lehtinen, L. Verschaffel, and F. Depaepe, “Systematic review of research on pedagogical content knowledge in mathematics: Insights from a topic-specific approach,” *ZDM Mathematics Education*, vol. 57, pp. 777–794, 2025, doi: 10.1007/s11858-025-01684-1.

[9] F. Depaepe, L. Verschaffel, and G. Kelchtermans, “Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research,” Teach. Teach. Educ., vol. 34, pp. 12–25, 2013, doi: 10.1016/j.tate.2013.03.001.

[10] Ningsih, S. Y., Turmudi, & Juandi, D. (2020). *Pedagogical content knowledge (PCK) profile of prospective teachers in mathematics learning*. Journal of Physics: Conference Series, 1521(3), 032057, doi: 10.1088/1742-6596/1521/3/032057

[11] S. Fiangga, S. Khabibah, S. M. Amin, dan R. Ekawati, "A Learning Design Analysis of the Pre-service Teachers’ Mathematics Pedagogical Content Knowledge," *J. Phys.: Conf. Ser.*, vol. 1899, no. 1, p. 012153, 2021, doi: 10.1088/1742-6596/1899/1/012153.

[12] Yang, X., Schwarz, B. & Leung, I.K.C. Pre-service mathematics teachers’ professional modeling competencies: a comparative study between Germany, Mainland China, and Hong Kong. *Educ Stud Math* **109**, 409–429 (2022). doi: 10.1007/s10649-021-10064-x.

[13] Njiku, J. (2025). Pre-service teachers manifested mathematics pedagogical content knowledge: The role of the teaching practicum. Pedagogical Research, 10(1), em0229. doi: 10.29333/pr/15647.

[14] Boevé AJ, Meijer RR, Albers CJ, Beetsma Y, Bosker RJ (2015) Introducing Computer-Based Testing in High-Stakes Exams in Higher Education: Results of a Field Experiment. PLOS ONE 10(12): e0143616. https://doi.org/10.1371/journal.pone.0143616

[15] C. Y. Piaw, "Replacing paper-based testing with computer-based testing in assessment: Are we doing wrong?." *Procedia-Social and Behavioral Sciences* 64 (2012): 655-664, doi: 10.1016/j.sbspro.2012.11.077

[16] Julie, Gess-Newsome. "A model of teacher professional knowledge and skill including PCK: Results of the thinking from the PCK Summit." *Re-examining pedagogical content knowledge in science education*. Routledge, 2015. 28-42.

[17] Schmid, Mirjam, Eliana Brianza, and Dominik Petko. "Self-reported technological pedagogical content knowledge (TPACK) of pre-service teachers in relation to digital technology use in lesson plans." *Computers in human behavior* 115 (2021): 106586, doi: 10.1016/j.chb.2020.106586

[18] Maas, L.; Brinkhuis, M.J.S.; Kester, L.; Wijngaards-de Meij, L. Cognitive Diagnostic Assessment in University Statistics Education: Valid and Reliable Skill Measurement for Actionable Feedback Using Learning Dashboards. *Appl. Sci.* **2022**, *12*, 4809. doi: 10.3390/app12104809.

[19] A. Barana, M. Marchisio, and M. Sacchet, “Interactive feedback for learning mathematics in a digital learning environment,” Educ. Sci., vol. 11, no. 6, art. 279, 2021, doi: 10.3390/educsci11060279.

[20] Rezat, S. How automated feedback from a digital mathematics textbook affects primary students’ conceptual development: two case studies. *ZDM Mathematics Education* **53**, 1433–1445 (2021). doi: 10.1007/s11858-021-01263-0.

[21] L. N. Yaqin, L. D. Prasojo, N. A. Haji-Othman, N. Yusof, and A. Habibi, "Addressing the Digital Divide in Indonesian Higher Education: Insights, Implications, and Potential Solutions," in *From Digital Divide to Digital Inclusion*, Singapore: Springer Nature Singapore, 2024, pp. 291–307. doi: 10.1007/978-981-99-7645-4_13

[22] N. Rajiha and S. Rini, "Computer-Based Standardized Testing (CBT) Summative Assessment: One Institution Experience," *JISAE (Journal of Indonesian Student Assessment and Evaluation)*, vol. 9, no. 1, 2023. doi: 10.21009/jisae.v9i1.33233.

[23] G. Kaiser and J. König, “Competence measurement in (mathematics) teacher education and beyond: Implications for policy,” High. Educ. Policy, vol. 32, no. 4, pp. 597–615, 2019, doi: 10.1057/s41307-019-00139-z.

[24] T. Kleickmann et al., “Teachers’ content knowledge and pedagogical content knowledge: The role of structural differences in teacher education,” J. Teacher Educ., vol. 64, no. 1, pp. 90–106, 2013, doi: 10.1177/0022487112460398.

[25] J. Baumert and M. Kunter, “The COACTIV model of teachers’ professional competence,” in Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers, M. Kunter et al., Eds. New York: Springer, 2013, pp. 25–48, doi: 10.1007/978-1-4614-5149-5_2.

[26] B. Rowan, R. Correnti, and R. J. Miller, “What large-scale, survey research tells us about teacher effects on student achievement**:** Insights from the Prospects Study of Elementary Schools,” J. Teacher Educ., vol. 104, no. 8, pp. 1525–1567, 2002, doi: 10.1177/00224871211001323.

[27] S. Blömeke, J. E. Gustafsson, and R. J. Shavelson, “Beyond dichotomies: Competence viewed as a continuum,” Z. Psychol., vol. 223, no. 1, pp. 3–13, 2015, doi: 10.1027/2151-2604/a000398.

[28] E. Mohammadpour *et al.*, "A performance-based test to measure teachers' mathematics and science content knowledge (CK) and pedagogical knowledge (PK) of teachers," *Heliyon*, vol. 9, no. 3, 2023, doi: 10.1007/s10763-021-10198-6.

[29] T. Clark and H. Endres, "Computer-based diagnostic assessment of high school students’ grammar skills with automated feedback – an international trial," *Assessment in Education: Principles, Policy & Practice*, vol. 28, no. 5-6, pp. 602–632, 2021, doi: 10.1080/0969594X.2021.1970513

[30] G. A. Nortvedt and N. Buchholtz, “Assessment in mathematics education: Responding to issues regarding methodology, policy, and equity,” ZDM Math. Educ., vol. 50, no. 4, pp. 555–570, 2018, doi: 10.1007/s11858-018-0932-3.

Bukti Review 2 oleh Reviewer

Journal of Mathematics Instruction, Social Research and Opinion
Vol. xx, No. x, March/June/September/December 202x, pp. 1-1x, https://doi.org/
ISSN 2962-7842

Assessing Mathematics Pre-Service Teachers' Content Knowledge and Pedagogical Content Knowledge through a Diagnostic Computer-Based Test

Article Info

Article history:

Received mm dd, yyyy
Revised mm dd, yyyy
Accepted mm dd, yyyy

Keywords:

Computer-based test
Content knowledge
Pedagogical content knowledge
Mathematics pre-service teachers
Diagnostic assessment
Mathematical knowledge for teaching

ABSTRACT

Ensuring an adequate professional-knowledge profile of mathematics pre-service teachers is a strategic issue for teacher education. However, Content Knowledge (CK) and Pedagogical Content Knowledge (PCK) are rarely measured simultaneously through technology-based diagnostic assessment in the Indonesian context. This study aims to measure and profile the CK and PCK of mathematics pre-service teachers using a diagnostic Computer-Based Test (CBT). A quantitative descriptive-correlational design was applied to 30 fifth- and seventh-semester students of a Mathematics Education (Tadris Matematika) study program at a teacher education institution (LPTK), selected purposively. Each construct was measured with 25 multiple-choice items whose content validity was established through expert judgment and whose reliability was confirmed before use (KR-20 = 0.81 for CK and 0.79 for PCK). The data were analyzed using descriptive statistics, Pearson correlation, and simple linear regression after the underlying assumptions had been checked. The ideal competency threshold was set at 75%, a researcher-defined mastery criterion consistent with mastery-learning standards. The results show mean scores of 50.00% for CK and 52.67% for PCK, both below the 75% threshold, with 60.00% of the students in the low category; Specialized Content Knowledge (SCK, 36.91%) and Knowledge of Content and Curriculum (KCC, 46.25%) were the weakest dimensions. CK and PCK were strongly and positively correlated ($r = 0.72$; $p < 0.001$), and CK accounted for about 52% of the variance in PCK as a predictive, non-causal association. Diagnostic analysis at the dimension, indicator, and individual-student levels shows that the CBT produces specific and actionable profiles. These findings underscore the need for data-driven LPTK curriculum reform that integrates content mastery with pedagogical capacity.

Corresponding Author:

INTRODUCTION

The quality of a nation's mathematics education depends heavily on teachers' professional competence. Cross-national studies have shown that mathematics teachers' professional knowledge, comprising Content Knowledge (CK) and Pedagogical Content Knowledge

Commented [a5]: alphabetical order

Commented [a1]: No need

Commented [a2]: In english

Commented [a3]: In english

Commented [a4]: shorten to a maximum of 200 words

Knowledge (PCK), is a primary predictor of instructional effectiveness and students' academic achievement [1], [2]. The 2022 PISA results placed Indonesian students' mathematics literacy 68th out of 81 countries, below the OECD average, indicating a fundamental problem in the quality of mathematics instruction [3]. This condition is inseparable from the quality of professional knowledge formed in teacher education institutions (LPTK), so ensuring an adequate professional-knowledge profile of pre-service teachers becomes a strategic imperative [4], [5].

Theoretically, the Mathematical Knowledge for Teaching (MKT) framework distinguishes CK—comprising common, specialized, and horizon content knowledge—from PCK, which encompasses knowledge of content and students, content and teaching, and content and curriculum [6], [7]. Systematic reviews affirm that PCK is the most frequently studied yet the most difficult component to measure consistently because of its contextual and multidimensional nature [8], [9]. Studies in Indonesia show that the professional knowledge of mathematics pre-service teachers varies widely across institutions and curricula, with a notable weakness in the PCK aspect [10], [11]. Similar findings have been reported in cross-national contexts [12], [13].

Alongside the rapid digitalization of education, competency-evaluation systems are now shifting massively from traditional approaches toward technology-based assessment. Various studies indicate that CBT is often more valid and reliable than conventional examination methods [14], [15]; however, its use to objectively measure both the content knowledge and the pedagogical content knowledge of pre-service teachers simultaneously remains very limited because of the complexity of these constructs [16], [17]. The potential of CBT to generate diagnostic feedback based on learning analytics has also not been optimally exploited [18], even though interactive digital feedback has been shown to improve mathematical understanding [19], [20].

The urgency of this study is heightened in the context of Indonesia's uneven higher-education digital transformation [21], which is not yet supported by adequate standardization of assessment systems [22]. Measuring teacher competence in the digital era is required to produce data that are actionable for study-program managers and policymakers [23]. A number of studies affirm that CK and PCK are relatively independent constructs that require different intervention pathways [24], [25], and that teacher effects on student achievement are substantial [26].

The novelty of this study lies in three aspects. First, theoretically, the study integrates the MKT framework and technology-based assessment within a single coherent quantitative measurement design to identify CK and PCK profiles simultaneously [27], [28]. Second, methodologically, the CBT is used as an autonomous, automated assessment medium that generates layered diagnostic feedback—at the total, dimension, indicator, and individual-student levels—so that test results can be acted upon directly for instructional improvement [29]. Third, contextually, the study targets students at a religiously affiliated LPTK in Indonesia, a context that—although not entirely absent from the literature—has received only limited attention in in-depth, technology-based diagnostic studies [30]. This study aims to: (1) measure students' CK and PCK achievement and profiles; (2) identify strengths and weaknesses in each dimension and indicator; and (3) demonstrate the use of diagnostic profiles—including at the individual-student level—as a basis for designing instructional interventions.

Commented [a6]: use the Mendeley tool, the Mendeley tool is not active

Commented [a7]: In english

Commented [a8]: need to clarify research gaps through a critical synthesis of previous studies, avoid general claims of novelty, and emphasize the theoretical, methodological, and practical contributions that truly distinguish this research from previous research.

Commented [a9]: In emnglish

Commented [a10]: the article's scientific contribution will still appear more like an application of CBT to a local context than the development of new knowledge relevant to the international research community.

METHOD

This study employed a quantitative approach with a descriptive-survey and correlational design. The population comprised students of the Mathematics Education (Tadris Matematika) study program at an LPTK who were in their fifth and seventh semesters. A sample of 30 students was determined through purposive sampling with the inclusion criteria: (1) active students who had adequately completed mathematics-content and pedagogy courses; and (2) willing to take the entire series of CBT tests. Given the limited sample size, all data were analyzed as a single intact group ($n = 30$) and were not directed toward inter-semester comparison.

Research setting and participants. The research was conducted in the odd semester of the 2024/2025 academic year at the Mathematics Education (Tadris Matematika) study program of a state Islamic teacher education institution (LPTK) in Indonesia. Because every eligible student who met the inclusion criteria (having completed the core mathematics-content and pedagogy courses) took part, the 30 participants were treated as a single intact group; this modest sample size is acknowledged as a constraint on statistical power and on the generalizability of the correlational and regression results, as elaborated in the Discussion.

The operationalization of the variables comprised two main constructs. CK was defined as mastery of the mathematics content to be taught, measured through 25 multiple-choice items; PCK was defined as the ability to integrate content knowledge with pedagogical strategies, also measured through 25 multiple-choice items. Scores were converted into percentages of achievement relative to the ideal maximum score. The CBT instrument was designed to generate layered diagnostic reports, namely the total score, achievement per dimension, achievement per indicator, and an individual achievement profile for each student. Each item was mapped to a specific indicator and dimension (see the notes below Table 1) so that the CBT system could produce achievement per dimension and per indicator for all students at once, as well as their individual diagnostic profiles. The composition of the instrument is presented in Table 1.

Table 1. Composition of CBT Instrument Items by MKT Dimension

Domain	Dimension	Number of Items
CK	Common Content Knowledge (CCK)	7
CK	Specialized Content Knowledge (SCK)	14
CK	Horizon Content Knowledge (HCK)	4
PCK	Knowledge of Content and Students (KCS)	10
PCK	Knowledge of Content and Teaching (KCT)	9
PCK	Knowledge of Content and Curriculum (KCC)	6
Total		50

Notes on dimension codes and indicator coverage (based on the developed CK and PCK instruments). **CK – CCK** (Common Content Knowledge, items 1–7): understanding basic concepts (e.g., lines and angles), performing calculations and measurements accurately, and applying mathematical concepts in everyday life. **CK – SCK** (Specialized Content Knowledge, items 8–21): explaining concepts and principles, analyzing and connecting mathematical objects, proving properties, building models, and solving complex mathematical problems. **CK – HCK** (Horizon Content Knowledge, items 22–25): connecting school mathematics with advanced academic mathematics and with other disciplines. **PCK – KCS** (Knowledge of Content and Students, items 1–10): analyzing students' preconceptions and understanding as well as

Commented [a11]: In english

Commented [a12]: Explain the teknik sampling dan why just 30 responden?

Commented [a13]: All tadris matematika and LPTK translete in english

Commented [a14]: why use multiple choice, give reasons predicting and analyzing students' misconceptions, errors, and difficulties. **PCK – KCT** (Knowledge of Content and Teaching, items 11–19): selecting and organizing learning materials, selecting sources, strategies, and techniques, and selecting and analyzing representations for presenting material. **PCK – KCC** (Knowledge of Content and Curriculum, items 20–25): designing content and topic combinations, applying standards and curricula, and designing instruments to measure students' abilities.

Commented [a15]: You can add the form of the instrument used, if there are many you can include a link that readers can see.

Instrument development and content validity. The CBT instrument was developed through a systematic procedure. First, a test blueprint was constructed by mapping each item to a specific MKT dimension and indicator (Table 1). Second, the draft items were reviewed by three experts—two mathematics-education lecturers and one educational-assessment specialist—to establish content validity; the level of expert agreement was quantified using Aiken's V, and items with $V \geq 0.80$ were retained while the remaining items were revised or replaced. Third, the revised items were trialed and analyzed empirically before being administered in the present study.

Reliability and item analysis. Based on the item analysis, the instrument demonstrated adequate psychometric quality. Internal-consistency reliability computed with the Kuder–Richardson formula (KR-20) reached 0.81 for the CK test and 0.79 for the PCK test, both within the acceptable-to-good range. The item difficulty indices ranged from 0.20 to 0.87 (a balanced mix of easy, medium, and hard items), and the discrimination indices were largely satisfactory ($D \geq 0.30$); only items meeting the validity, difficulty, and discrimination criteria were retained in the final CBT.

Item format. Multiple-choice items were chosen because the CBT was designed for automated, standardized, and scalable diagnosis across many participants and indicators simultaneously. To capture pedagogical reasoning rather than mere recall, the PCK items were built around short classroom scenarios, samples of student work, and instructional-decision situations, with distractors constructed from typical student misconceptions and common teaching errors. It is nevertheless acknowledged that the multiple-choice format cannot fully capture the depth of pedagogical reasoning that open-ended or performance-based tasks would reveal, a limitation revisited in the Discussion.

Commented [a16]: There was no syllabus analysis.

There was no observation.

There were no interviews.? For to support quantitative data?

The analysis techniques comprised descriptive and inferential analyses. Descriptive analysis produced the mean (M), standard deviation (SD), minimum, maximum, median, and percentage of achievement at the total, dimension, indicator, and individual levels, as well as a three-level categorization (High, Medium, Low). The categorization used the following cut-off points: High (≥ 75%), Medium (60–74.99%), and Low (< 60%), where the 75% mark represents the ideal mastery criterion adopted in this study. Inferential analysis was conducted through two techniques: (1) Pearson correlation to test the CK–PCK relationship and the corresponding inter-dimension relationships (CCK–KCS, SCK–KCT, HCK–KCC) at $\alpha = 0.05$; and (2) simple linear regression with CK as the predictor and PCK as the criterion. Prior to the inferential tests, the underlying assumptions were examined, namely the normality of the CK and PCK distributions (Shapiro–Wilk), the linearity of their relationship, the absence of influential outliers, and homoscedasticity; all assumptions were adequately met, so that the parametric analyses were appropriate.

All scoring, achievement computation, correlation, and linear regression were performed using Microsoft Excel through standardized formulas (including CORREL, SLOPE, INTERCEPT, and RSQ) so that every figure can be traced back to the raw data. The study also observed standard research-ethics principles: participation was entirely voluntary, informed consent was obtained from all respondents before data collection, and the participants' anonymity and confidentiality were maintained throughout the recording, processing, and reporting of data by using codes instead of names, with the individual-profile example presented under the pseudonym Student 1.

RESULTS AND DISCUSSION

3.1. Results

This section reports the findings of the data analysis in the following order: respondent characteristics, descriptive statistics, the distribution of achievement categories, dimension- and indicator-level achievement, the correlation and regression results, and an individual diagnostic profile. The 30 respondents were fifth- and seventh-semester students of the Mathematics Education (Tadris Matematika) study program who had completed the core mathematics-content and pedagogy courses; all of them completed the full series of CK and PCK tests through the CBT system.

Overall, students' professional-knowledge achievement had not yet reached the ideal competency threshold (≥ 75%). The mean CK score was 50.00% (SD = 10.16), PCK 52.67% (SD = 23.96), and the combined Professional Knowledge (PK) score 51.33% (SD = 16.04). The complete descriptive statistics are presented in Table 2.

Table 2. Descriptive Statistics of CK, PCK, and Professional Knowledge Scores (n = 30)

Variable	Mean (%)	SD	Min	Max	Median
Content Knowledge (CK)	50.00	10.16	24	68	52.00
Pedagogical Content Knowledge (PCK)	52.67	23.96	8	96	54.00
Professional Knowledge (PK)	51.33	16.04	26	78	50.00

Table 2 shows that the mean PCK is slightly higher than CK (a difference of 2.67 points), an interesting pattern indicating that pedagogical-transformation capacity slightly exceeds content mastery. The large standard deviation of PCK (SD = 23.96) and the much smaller one for CK (SD = 10.16) reflect substantial heterogeneity in knowledge profiles, with score ranges of 24–68% for CK and 8–96% for PCK. The distribution of Professional Knowledge categories is presented in Table 3.

Table 3. Distribution of Professional Knowledge Categories (n = 30)

Category	Frequency (f)	Percentage (%)
High	0	0.00
Medium	12	40.00
Low	18	60.00
Total	30	100.00

All students (100%) were in the Low or Medium category, with 60.00% in the Low category and 40.00% in the Medium category, and no student reaching the High category. This indicates that the majority of pre-service teachers have not attained an optimal level of professional-knowledge mastery. Achievement per dimension is presented in Table 4 and Figure 1.

Table 4. Average Achievement per CK and PCK Dimension

Domain	Dimension	Items	Mean (%)	SD
CK	CCK	7	73.33	20.47
CK	SCK	14	36.91	24.16
CK	HCK	4	55.00	35.60
PCK	KCS	10	54.07	33.18
PCK	KCT	9	57.50	26.14
PCK	KCC	6	46.25	32.30

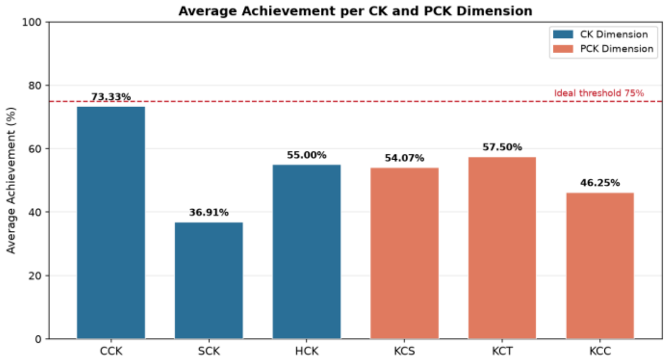


Figure 1. Average Achievement per CK and PCK Dimension

In the CK domain, the CCK dimension was the highest (73.33%) and SCK the lowest (36.91%); in the PCK domain, KCT was the highest (57.50%) while KCC was the lowest of all six dimensions (46.25%). Diagnostically, the low SCK and KCC signal a weak ability to relate content to the structure and organization of the curriculum—an area that needs to be prioritized in intervention. Achievement at the indicator level is presented in Table 5 and visualized in Figure 2.

Table 5. Average Achievement per CK and PCK Indicator

Domain	Indicator	Mean (%)
CK	CCK1	86.67
CK	CCK2	80.00
CK	CCK3	46.67
CK	SCK1	63.33
CK	SCK2	46.67
CK	SCK3	43.33
CK	SCK4	20.00
CK	SCK5	23.33
CK	SCK6	30.00
CK	SCK7	33.33
CK	HCK1	53.33
CK	HCK2	56.67
PCK	KCS1	58.33
PCK	KCS2	48.33
PCK	KCS3	58.33
PCK	KCS4	51.67
PCK	KCS5	53.33
PCK	KCT1	55.00
PCK	KCT2	66.67
PCK	KCT3	60.00
PCK	KCT4	36.67

PCK	KCC1	51.67
PCK	KCC2	43.33
PCK	KCC3	45.00

Notes on indicator codes (based on the CK and PCK instruments). CK indicators — **CCK1**: understanding basic school-mathematics concepts, such as lines, angles, sequences, and basic statistics (items 1–3); **CCK2**: performing calculations and measurements accurately (items 4–5); **CCK3**: applying mathematical concepts in everyday life (items 6–7); **SCK1**: explaining academic mathematical concepts and principles (items 8–9); **SCK2**: analyzing mathematical objects (items 10–11); **SCK3**: connecting mathematical properties and objects (items 12–13); **SCK4**: proving mathematical properties or theorems (items 14–16); **SCK5**: applying proof strategies (item 17); **SCK6**: building and constructing mathematical models (items 18–19); **SCK7**: solving complex mathematical problems (items 20–21); **HCK1**: connecting school mathematics with advanced academic mathematics (items 22–23); **HCK2**: connecting mathematics with other disciplines (items 24–25). PCK indicators — **KCS1**: analyzing students' preconceptions (items 1–2); **KCS2**: identifying students' level of understanding (items 3–4); **KCS3**: predicting students' misconceptions (items 5–6); **KCS4**: identifying sources of students' learning difficulties (items 7–8); **KCS5**: analyzing students' literacy issues and errors (items 9–10); **KCT1**: organizing and sequencing learning materials (items 11–12); **KCT2**: selecting learning sources, strategies, and techniques (items 13–14); **KCT3**: selecting representations and illustrations for presenting material (items 15–17); **KCT4**: analyzing the weaknesses and strengths of representations (items 18–19); **KCC1**: composing content and topic combinations (items 20–21); **KCC2**: applying standards and curricula (items 22–23); **KCC3**: designing instruments to measure students' abilities (items 24–25).

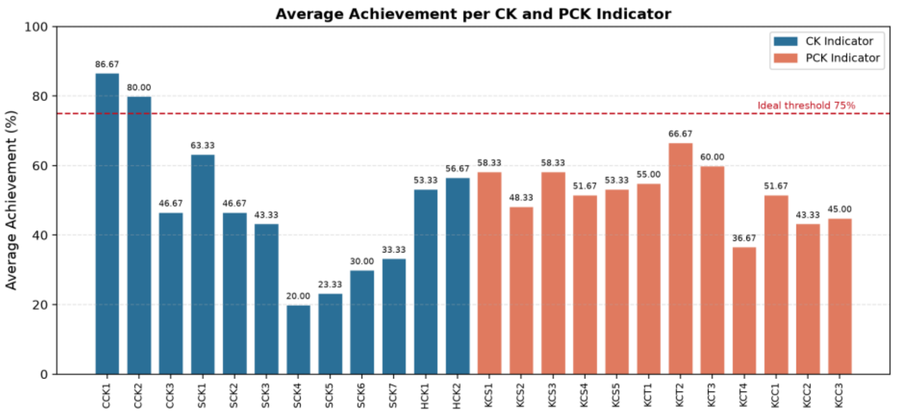


Figure 2. Average Achievement per CK and PCK Indicator

Figure 2 visualizes the average achievement of each indicator exactly as listed in Table 5. The range of achievement across indicators spans 20.00% to 86.67%. The highest achievement is on the CK indicator CCK1 (86.67%; understanding basic school-mathematics concepts), which is fundamental and therefore the most widely mastered by students. Conversely, the lowest overall achievement is on indicator SCK4 (20.00%; proving mathematical properties or theorems)—the skill with the highest cognitive demand in the CK domain—followed by SCK5 (23.33%; applying proof strategies) and SCK6 (30.00%; building and constructing mathematical models). The pattern in the SCK dimension declines gradually with complexity, from explaining concepts and principles (SCK1; 63.33%), analyzing and connecting mathematical objects (SCK2 46.67%; SCK3 43.33%), solving complex problems (SCK7; 33.33%), to the most difficult modeling and proving (SCK6 30.00%; SCK5 23.33%; SCK4 20.00%)—a pattern consistent with cognitive load theory. Because each dimension's achievement remains the mean of its constituent indicators—for example, KCS (54.07%) is the mean of KCS1–KCS5 and CCK (73.33%) the mean of CCK1–CCK3—this pattern maps precisely which items and indicators are the sources of weakness, so that lecturers can design remediation targeted at concrete rather than generic weaknesses. The results of the Pearson correlation test are presented in Table 6.

Table 6. Pearson Correlation between Domains and between Parallel Dimensions

Relationship	R	P	Interpretation
CK – PCK	0.72	< 0.001	Strong
CCK – KCS	0.42	< 0.05	Moderate
SCK – KCT	0.54	< 0.01	Moderate
HCK – KCC	0.86	< 0.001	Very strong

The strong positive correlation between CK and PCK ($r = 0.72$; $p < 0.001$) confirms that the two domains are constructively interdependent, reinforced by the correlations between parallel dimensions, especially HCK–KCC ($r = 0.86$) and SCK–KCT ($r = 0.54$). The predictive relationship of CK on PCK was tested through simple linear regression (Table 7 and Figure 3).

Table 7. Results of the Simple Linear Regression of CK on PCK

Parameter	Value
Slope (b)	1.708
Intercept (a)	−32.734
Coefficient of determination (R^2)	0.5245
Pearson r	0.72
Equation	PCK = 1.708 × CK − 32.734

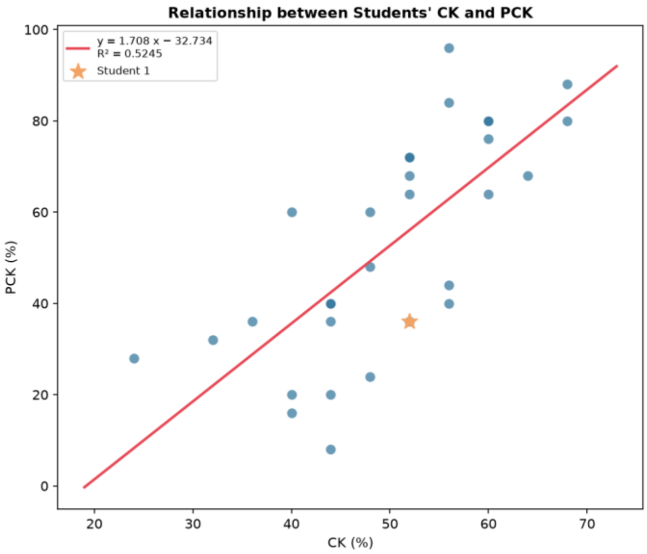


Figure 3. Scatter Plot and Regression Line of the CK–PCK Relationship

The equation PCK = 1.708 × CK − 32.734 with $R^2 = 0.52$ indicates that about 52% of the variance in PCK is statistically associated with CK. Because the design is correlational, this relationship should be interpreted as a predictive association rather than a causal effect; the remaining 48% of the variance signals that PCK does not form automatically from content mastery but requires an explicit pedagogical learning pathway.

The main advantage of the CBT instrument lies in its ability to break down aggregate achievement into individual diagnostic profiles automatically. As an illustration, Table 8 and Figure 4 summarize the achievement of one student disguised as Student 1, while Figure 5 and Figure 6 display authentic screenshots of the diagnostic scoring report generated directly by the CBT system for Student 1. Such individual reports are produced automatically by the system for every participant without any manual scoring.

Table 8. Individual Achievement Profile of Student 1 Compared with the Class Average

Domain	Student 1 Score (%)	Class Average (%)	Difference	Category
Content Knowledge (CK)	52.00	50.00	+2.00	Low
Pedagogical Content Knowledge (PCK)	36.00	52.67	−16.67	Low

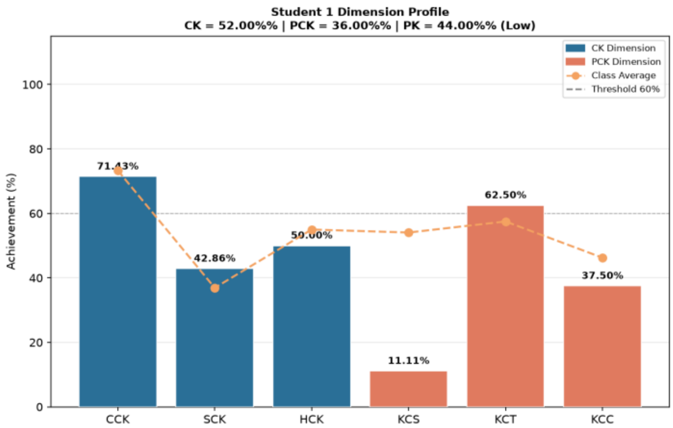


Figure 4. CK and PCK Achievement Profile of Student 1 Compared with the Class Average

At the individual level, the CBT system automatically scores the answers and generates a diagnostic scorecard for each participant. Student 1, for example, answered 13 of 25 CK items correctly (52.00%) and 9 of 25 PCK items correctly (36.00%); the CK achievement was slightly above the class average (a difference of +2.00 points) while the PCK achievement was far below it (a difference of −16.67 points), and neither reached the ideal competency threshold (75%), placing the student in the Low category for both domains. Authentic screenshots of the report are presented in Figure 5 (CK) and Figure 6 (PCK). The system-generated screenshots were anonymized so that no identifiable student information is displayed, and the pseudonym Student 1 is used throughout.

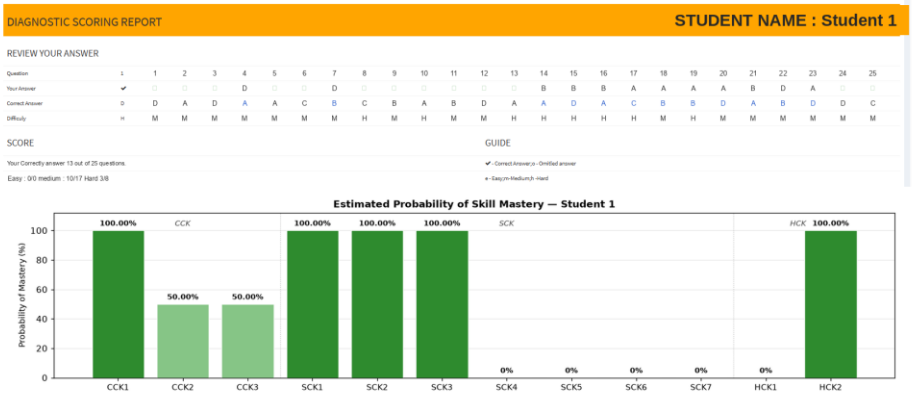


Figure 5. CBT Diagnostic Scorecard for the CK Domain for Student 1 (original system screenshot)

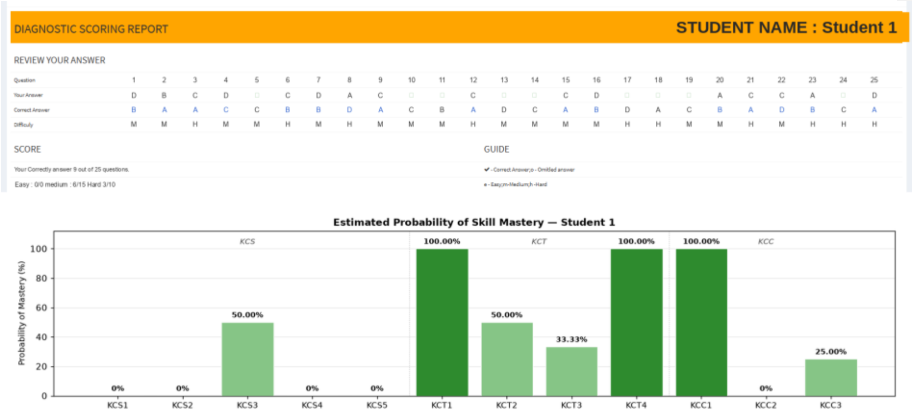


Figure 6. CBT Diagnostic Scorecard for the PCK Domain for Student 1 (system screenshot)

Figures 5 and 6 reveal diagnostic functions of the CBT that are more difficult to implement efficiently in conventional paper-based tests. In addition to recapping correct/incorrect answers per item along with their difficulty level, a combination of the Medium (M) and Hard (H) categories according to the item analysis of the instrument, the system presents an Estimated Probability of Skill Mastery graph that maps Student 1's mastery probability for each skill indicator (CCK1–HCK2 for CK and KCS1–KCC3 for PCK), complete with a mapping of sample items to each indicator. Thus, the CBT does not merely provide a final score but directly identifies which indicators have been mastered and which remain weak, in Student 1's case, several indicators show a mastery probability close to zero, so that lecturers can follow up with specific and immediate remediation. It is this automation of scoring, standardization of reporting, and diagnostic depth that affirms the added value of the CBT in assessing the professional knowledge of mathematics pre-service teachers.

Furthermore, the two scorecards allow Student 1's achievement to be traced down to the indicator level, not stopping at dimension averages. In the CK domain (Figure 5), Common Content Knowledge (CCK) was the strongest dimension (71.43%; 5 of 7 items): understanding basic school-mathematics concepts (CCK1) was fully mastered (100.00%), whereas performing calculations and measurements (CCK2) and applying mathematical concepts in everyday life (CCK3) were each only partially mastered (50.00%). Specialized Content Knowledge (SCK) stood at 42.86% (6 of 14 items) with a highly contrasting pattern: three basic indicators—explaining academic mathematical concepts and principles (SCK1), analyzing mathematical objects (SCK2), and connecting mathematical properties and objects (SCK3)—were all fully mastered (100.00%), while four indicators demanding higher-order reasoning—proving mathematical properties or theorems (SCK4), applying proof strategies (SCK5), building and constructing mathematical models (SCK6), and solving complex mathematical problems (SCK7)—were all at 0.00%, marking a total deficit in advanced mathematical reasoning. Horizon Content Knowledge (HCK) stood at 50.00% (2 of 4 items): connecting mathematics with other disciplines (HCK2) was fully mastered (100.00%), but connecting school mathematics with advanced academic mathematics (HCK1) remained at 0.00%. In the PCK domain (Figure 6), the indicator profile was also uneven within each dimension. Knowledge of Content and Teaching (KCT; dimension mean 62.50%) was the strongest: organizing and sequencing learning materials (KCT1) and analyzing the weaknesses and strengths of representations (KCT4) were fully mastered (100.00%), followed by selecting learning sources, strategies, and techniques (KCT2; 50.00%) and selecting representations and illustrations for presenting material (KCT3; 33.33%). Knowledge of Content and Curriculum (KCC; dimension mean 37.50%) rested on composing content and topic combinations (KCC1), which was fully mastered (100.00%), while designing instruments to measure students' abilities (KCC3) was only 25.00% and applying standards and curricula (KCC2) was still 0.00%. Conversely, Knowledge of Content and Students (KCS) was the weakest PCK dimension and the weakest point of the entire profile (11.11%): only predicting students' misconceptions (KCS3) was partially mastered (50.00%), while analyzing students' preconceptions (KCS1), identifying students' level of understanding (KCS2), identifying sources of students' learning difficulties (KCS4), and analyzing students' literacy issues and errors (KCS5) were all at 0.00%. This pattern confirms that Student 1's weaknesses are concentrated in a number of indicators that are genuinely untouched—SCK4, SCK5, SCK6, SCK7, and HCK1 in the CK domain and KCS1, KCS2, KCS4, KCS5, and KCC2 in the PCK domain, all at 0.00%—so that these ten indicators are the most urgent remediation priorities, while the indicators already fully mastered (CCK1, SCK1, SCK2, SCK3, HCK2, KCT1, KCT4, and KCC1) can serve as a foundation for reinforcement. Thus, indicator-based diagnosis provides a far sharper and more specific direction for individual intervention than reading dimension averages alone.

Summary of key findings. In sum, both CK (50.00%) and PCK (52.67%) fell below the 75% mastery threshold; 60.00% of the students were in the low category and none reached the high category; SCK (36.91%) and KCC (46.25%) were the weakest dimensions; CK and PCK were strongly and positively associated ($r = 0.72$; $p < 0.001$); and the CBT was able to decompose these results down to the indicator and individual-student levels, as illustrated by the profile of Student 1.

Commented [a17]: Student 1

Quite interesting. However, the discussion is too long. One student's point cannot be the basis for a general interpretation.
It would be better to use it as an illustrative example rather than primary evidence.

3.2. Discussion

The main findings reveal three key facts. **First**, neither CK (50.00%) nor PCK (52.67%) reached the minimum competency threshold; interestingly, PCK was even slightly higher than CK, reflecting the knowledge-transformation challenge articulated by Shulman [7] and reinforced in the MKT framework [6]. **Second**, the SCK dimension was the weakest point of CK (36.91%) and KCC the weakest point of PCK (46.25%), indicating a double deficit in the mastery of advanced mathematical content and curriculum, which in the COACTIV model strongly determines the quality of instructional planning [2], [25]. **Third**, the lower standard deviation of CK (10.16) compared with PCK (23.96) shows that CK achievement is more uniform, whereas PCK is highly heterogeneous, indicating large variation in students' pedagogical capacity.

The significant positive correlations, especially HCK–KCC ($r = 0.86$) and SCK–KCT ($r = 0.54$), show that content mastery in particular dimensions goes hand in hand with the relevant pedagogical capacity. However, the regression equation ($R^2 = 0.52$) affirms that content mastery explains only part of the variance in PCK, consistent with findings that pre-service teachers in developing countries tend to have lower PCK scores than CK because LPTK curricula place greater emphasis on content mastery [1], [10]. The case of Student 1 reinforces this at the individual level: Student 1 (CK = 52%, PCK = 36%) falls into the Low category with a highly uneven profile, namely CK achievement slightly above the class average while PCK is far below it. In the PCK domain, KCT is relatively the highest (62.50%) while KCS is the lowest (11.11%) and KCC is low (37.50%), indicating an ability to select teaching strategies that is not yet accompanied by an adequate understanding of student characteristics or curriculum mastery. Intervention needs to target KCS and KCC simultaneously, especially the ability to analyze preconceptions, identify levels of understanding, and identify sources of students' learning difficulties.

These results are broadly consistent with prior Indonesian studies reporting that the professional knowledge of mathematics pre-service teachers is generally moderate and uneven, with PCK frequently emerging as the more fragile component [10], [11]. The finding that PCK was marginally higher than CK is noteworthy and can be interpreted cautiously: the PCK items, which draw on familiar classroom situations, may be more accessible to students than the CK items, which demand advanced and abstract mathematical reasoning; moreover, the far larger dispersion of PCK (SD = 23.96) indicates that this apparent advantage is unevenly distributed across the cohort. The particularly low SCK (36.91%) is plausibly linked to teacher-education curricula that emphasize procedural and elementary content over advanced mathematical reasoning—such as formal proof, mathematical modeling, and complex problem solving—so that the more cognitively demanding indicators (SCK4–SCK7) remain underdeveloped [24], [25].

Theoretically, this study validates and extends the application of the MKT framework and the COACTIV model to the context of religiously affiliated LPTKs in Indonesia, which are still underrepresented in the indexed literature [6], [25], [11]. Methodologically, the study demonstrates a CBT model that produces layered diagnostic profiles—from the total, dimension, and indicator levels to the individual student—thus addressing the previously identified gap in using CBT to measure pedagogical competence [22], [28]. Achievement profiles that can be decomposed to the individual level, as shown in the case

Commented [a18]: In the discussion section, there is no need to include results in the form of numbers, focus more on discussing the results based on objectives linked to the results of previous research and expert theories.

Commented [a19]: What new theoretical insight does this study provide?

of Student 1, strengthen the function of the CBT as an instrument that generates actionable diagnostic feedback [18]. In policy terms, the mean PK of 51.33% and the absence of any student in the High category provide urgent empirical evidence to drive data-based curriculum reform rather than normative assumptions [5], [23], particularly in the context of Indonesia's 2022 PISA results [3].

This study opens a new perspective by positioning the CBT not merely as an examination tool but as a technology-based diagnostic-assessment ecosystem that produces rich cognitive data, decomposed per dimension and per individual, and actionable for program managers [23], [30]. Several limitations should nonetheless be noted, and they caution against generalizing these findings beyond the single institution and the 30 participants studied here: (a) the relatively small sample size (n = 30), which limits statistical power and generalizability; (b) the use of purposive sampling; (c) the single-institution context; (d) the still-bivariate scope of the analysis, which has not modeled other determinants simultaneously; (e) the absence of qualitative data to explain why students performed poorly; (f) the limited evidence of the CBT instrument's validity and reliability beyond the internal analyses reported here; and (g) the inherent constraint of the multiple-choice format, since PCK often requires the analysis of classroom scenarios, student work, teaching decisions, and pedagogical explanations that selected-response items cannot fully capture. Building on these limitations, future research should employ larger and more representative multi-institution samples, longitudinal and mixed-methods designs, and multivariate models to test additional determinants and to strengthen both internal and external validity.

Commented [a20]: more specific practical implications for curriculum development and the education of prospective teachers

CONCLUSION AND RECOMMENDATIONS

This study set out to measure and profile the CK and PCK of mathematics pre-service teachers at an LPTK through a diagnostic CBT. Both mean CK (50.00%) and PCK (52.67%) fell below the ideal competency threshold (75%), with 60.00% of the students in the low category and none in the high category; CCK was the strongest CK dimension (73.33%), whereas SCK (36.91%) and KCC (46.25%) were the weakest. CK and PCK were strongly and positively associated ($r = 0.72$; $p < 0.001$), with CK accounting for about 52% of the variance in PCK as a predictive, non-causal association. The layered diagnostic analysis—down to the individual level, as illustrated by Student 1—shows that the CBT can generate specific, actionable achievement profiles. Theoretically, the study extends the MKT and COACTIV frameworks to the Indonesian LPTK context; practically, it offers a standardized and efficient diagnostic instrument for teacher education. These conclusions should be read in light of the study's limitations—a small, single-institution, purposively selected sample and a bivariate, multiple-choice-based design—which future research should address through larger multi-institution samples, longitudinal and mixed-methods designs, and multivariate modeling.

Commented [a21]: reducing numerical details, avoiding overclaims, removing individual case illustrations (Student 1), clarifying theoretical, methodological and practical contributions proportionally, and closing the conclusion with a main message that emphasizes the value of research for the development of diagnostic assessment and education of prospective mathematics teachers

Recommendations. Based on these findings, the following recommendations are organized by stakeholder group.

For lecturers. Remedial instruction should target not only the weakest dimensions (SCK in CK and KCC in PCK) but also the weakest indicators identified in Table 5. In the CK domain, the priority is strengthening mathematical proof—proving properties or theorems (SCK4; 20.00%) and applying proof strategies (SCK5; 23.33%)—through scaffolded proof practice, analysis of proof errors, and habituation to deductive argumentation, followed by mathematical modeling (SCK6; 30.00%) and complex problem solving (SCK7; 33.33%) via problem-based learning. In the PCK domain, the focus is on analyzing the strengths and weaknesses of representations (KCT4; 36.67%), applying standards and curricula (KCC2; 43.33%), and designing measurement instruments (KCC3; 45.00%) through case-based learning, analysis of teaching videos, and curriculum-document study. The same indicator-targeting logic applies at the individual level, as shown in the profile of Student 1.

For LPTK curriculum managers. The results encourage integrating a diagnostic CBT—capable of mapping achievement down to the indicator and individual levels—into the study program's quality-assurance cycle, and rebalancing the curriculum so that advanced mathematical reasoning and pedagogical-content competence receive as much emphasis as procedural content mastery.

For future researchers. Subsequent studies are encouraged to use larger and more representative multi-institution samples, longitudinal designs, mixed-methods approaches, and multivariate models involving additional determinants (such as academic background, learning motivation, and the quality of lecturers' teaching) in order to strengthen external validity and to explain the causes of the observed weaknesses.

For CBT system development. The CBT itself can be enhanced by adding automatic item-level feedback with worked explanations, direct links to remedial learning materials for each weak indicator, and adaptive item delivery, so that the system functions not only as a diagnostic tool but also as an integrated learning-support environment.

Commented [a22]: These recommendations are presented as potentially effective alternatives based on the literature, not as actions directly justified by the results of this study. Furthermore, the individual case illustration (Student 1) should be removed from the recommendations section, and the proposed development of an untested CBT system should be moved to a further research agenda.

ACKNOWLEDGMENTS

The authors thank the Mathematics Education (Tadris Matematika) study program of the LPTK for facilitating the implementation of the research, as well as all the students who participated as respondents.

REFERENCES

[1] X. Yang and G. Kaiser, "The impact of mathematics teachers' professional competence on instructional quality and students' mathematical learning outcomes," Teach. Teach. Educ., vol. 110, art. 103582, 2022, doi: 10.1016/j.tate.2021.103582

[2] S. Krauss et al., "Pedagogical content knowledge and content knowledge of secondary mathematics teachers," Journal of Educational Psychology, 100(3), 716–725, 2008, doi: 10.1037/0022-0663.100.3.716.

[3] A. Schleicher, PISA 2022 Results: The State of Learning and Equity in Education. Paris: OECD Publishing, 2023.

[4] M. T. Tatto, J. Schwille, S. L. Senk, L. Ingvarson, R. Peck, G. Rowley, K. Bankov, M. Rodriguez, and M. Reckase, "Policy, Practice, and Readiness to Teach Primary and Secondary Mathematics in 17 Countries: Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)". Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement, 2012.

[5] H. Qian and P. Youngs, "The effect of teacher education programs on future elementary mathematics teachers' knowledge: A five-country analysis using TEDS-M data". J. Math. Teacher Educ., vol. 19, no. 4, pp. 371–396, 2016, doi: 10.1007/s10857-014-9297-9.

[6] D. L. Ball, M. H. Thames, and G. Phelps, "Content knowledge for teaching: What makes it special?," J. Teacher Educ., vol. 59, no. 5, pp. 389–407, 2008, doi: 10.1177/0022487108324554.

[7] L. S. Shulman, "Those who understand: Knowledge growth in teaching," Educ. Res., vol. 15, no. 2, pp. 4–14, 1986, doi: 10.3102/0013189X015002004.

[8] M. Grigaliūnienė, E. Lehtinen, L. Verschaffel, and F. Depaepe, "Systematic review of research on pedagogical content knowledge in mathematics: Insights from a topic-specific approach," *ZDM Mathematics Education*, vol. 57, pp. 777–794, 2025, doi: 10.1007/s11858-025-01684-1.

[9] F. Depaepe, L. Verschaffel, and G. Kelchtermans, "Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research," Teach. Teach. Educ., vol. 34, pp. 12–25, 2013, doi: 10.1016/j.tate.2013.03.001.

[10] Ningsih, S. Y., Turmudi, & Juandi, D. (2020). *Pedagogical content knowledge (PCK) profile of prospective teachers in mathematics learning*. Journal of Physics: Conference Series, 1521(3), 032057, doi: 10.1088/1742-6596/1521/3/032057

[11] S. Fiangga, S. Khabibah, S. M. Amin, dan R. Ekawati, "A Learning Design Analysis of the Pre-service Teachers' Mathematics Pedagogical Content Knowledge," *J. Phys.: Conf. Ser.*, vol. 1899, no. 1, p. 012153, 2021, doi: 10.1088/1742-6596/1899/1/012153.

[12] Yang, X., Schwarz, B. & Leung, I.K.C. Pre-service mathematics teachers' professional modeling competencies: a comparative study between Germany, Mainland China, and Hong Kong. *Educ Stud Math* **109**, 409–429 (2022). doi: 10.1007/s10649-021-10064-x.

[13] Njiku, J. (2025). Pre-service teachers manifested mathematics pedagogical content knowledge: The role of the teaching practicum. Pedagogical Research, 10(1), em0229. doi: 10.29333/pr/15647.

[14] Boevé AJ, Meijer RR, Albers CJ, Beetsma Y, Bosker RJ (2015) Introducing Computer-Based Testing in High-Stakes Exams in Higher Education: Results of a Field Experiment. PLOS ONE 10(12): e0143616. https://doi.org/10.1371/journal.pone.0143616

[15] C. Y. Piaw, "Replacing paper-based testing with computer-based testing in assessment: Are we doing wrong?." *Procedia-Social and Behavioral Sciences* 64 (2012): 655-664, doi: 10.1016/j.sbspro.2012.11.077

[16] Julie, Gess-Newsome. "A model of teacher professional knowledge and skill including PCK: Results of the thinking from the PCK Summit." *Re-examining pedagogical content knowledge in science education*. Routledge, 2015. 28-42.

[17] Schmid, Mirjam, Eliana Brianza, and Dominik Petko. "Self-reported technological pedagogical content knowledge (TPACK) of pre-service teachers in relation to digital technology use in lesson plans." *Computers in human behavior* 115 (2021): 106586, doi: 10.1016/j.chb.2020.106586

[18] Maas, L.; Brinkhuis, M.J.S.; Kester, L.; Wijngaards-de Meij, L. Cognitive Diagnostic Assessment in University Statistics Education: Valid and Reliable Skill Measurement for Actionable Feedback Using Learning Dashboards. *Appl. Sci.* **2022**, *12*, 4809. doi: 10.3390/app12104809.

[19] A. Barana, M. Marchisio, and M. Sacchet, "Interactive feedback for learning mathematics in a digital learning environment," Educ. Sci., vol. 11, no. 6, art. 279, 2021, doi: 10.3390/educsci11060279.

[20] Rezat, S. How automated feedback from a digital mathematics textbook affects primary students' conceptual development: two case studies. *ZDM Mathematics Education* **53**, 1433–1445 (2021). doi: 10.1007/s11858-021-01263-0.

[21] L. N. Yaqin, L. D. Prasojo, N. A. Haji-Othman, N. Yusof, and A. Habibi, "Addressing the Digital Divide in Indonesian Higher Education: Insights, Implications, and Potential Solutions," in *From Digital Divide to Digital Inclusion*, Singapore: Springer Nature Singapore, 2024, pp. 291–307. doi: 10.1007/978-981-99-7645-4_13

[22] N. Rajiha and S. Rini, "Computer-Based Standardized Testing (CBT) Summative Assessment: One Institution Experience," *JISAE (Journal of Indonesian Student Assessment and Evaluation)*, vol. 9, no. 1, 2023. doi: 10.21009/jisae.v9i1.33233.

[23] G. Kaiser and J. König, "Competence measurement in (mathematics) teacher education and beyond: Implications for policy," High. Educ. Policy, vol. 32, no. 4, pp. 597–615, 2019, doi: 10.1057/s41307-019-00139-z.

[24] T. Kleickmann et al., "Teachers' content knowledge and pedagogical content knowledge: The role of structural differences in teacher education," J. Teacher Educ., vol. 64, no. 1, pp. 90–106, 2013, doi: 10.1177/0022487112460398.

[25] J. Baumert and M. Kunter, "The COACTIV model of teachers' professional competence," in Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers, M. Kunter et al., Eds. New York: Springer, 2013, pp. 25–48, doi: 10.1007/978-1-4614-5149-5_2.

[26] B. Rowan, R. Correnti, and R. J. Miller, “What large-scale, survey research tells us about teacher effects on student achievement: Insights from the Prospects Study of Elementary Schools,” J. Teacher Educ., vol. 104, no. 8, pp. 1525–1567, 2002, doi: 10.1177/00224871211001323.

[27] S. Blömeke, J. E. Gustafsson, and R. J. Shavelson, “Beyond dichotomies: Competence viewed as a continuum,” Z. Psychol., vol. 223, no. 1, pp. 3–13, 2015, doi: 10.1027/2151-2604/a000398.

[28] E. Mohammadpour *et al.*, "A performance-based test to measure teachers' mathematics and science content knowledge (CK) and pedagogical knowledge (PK) of teachers," *Heliyon*, vol. 9, no. 3, 2023, doi: 10.1007/s10763-021-10198-6.

[29] T. Clark and H. Endres, "Computer-based diagnostic assessment of high school students’ grammar skills with automated feedback – an international trial," *Assessment in Education: Principles, Policy & Practice*, vol. 28, no. 5-6, pp. 602–632, 2021, doi: 10.1080/0969594X.2021.1970513

[30] G. A. Nortvedt and N. Buchholtz, “Assessment in mathematics education: Responding to issues regarding methodology, policy, and equity,” ZDM Math. Educ., vol. 50, no. 4, pp. 555–570, 2018, doi: 10.1007/s11858-018-0932-3.

Bukti Revisi ke 2

Journal of Mathematics Instruction, Social Research and Opinion
Vol. xx, No. x, March/June/September/December 202x, pp. 1-1x, https://doi.org/
ISSN 2962-7842

Assessing Mathematics Pre-Service Teachers' Content Knowledge and Pedagogical Content Knowledge through a Diagnostic Computer-Based Test

Article Info

Article history:

Received mm dd, yyyy
Revised mm dd, yyyy
Accepted mm dd, yyyy

Keywords:

Computer-based test
Content knowledge
Diagnostic assessment
Mathematical knowledge for teaching
Mathematics pre-service teachers
Pedagogical content knowledge

ABSTRACT

Content Knowledge (CK) and Pedagogical Content Knowledge (PCK) are core components of mathematics teachers' competence, yet they are rarely measured simultaneously through technology-based diagnostic assessment. This study measures and profiles the CK and PCK of mathematics pre-service teachers using a diagnostic Computer-Based Test (CBT). A quantitative descriptive-correlational design was applied to 30 purposively selected pre-service teachers. Each construct was measured with 25 multiple-choice items whose content validity was established by expert judgment and whose reliability was confirmed before use (KR-20 = 0.81 for CK; 0.79 for PCK). Data were analyzed with descriptive statistics, Pearson correlation, and simple linear regression after assumption checks. Mean CK (50.00%) and PCK (52.67%) were both below the 75% mastery threshold, with 60% of students in the low category; specialized content knowledge and knowledge of content and curriculum were the weakest dimensions. CK and PCK correlated strongly and positively ($r = 0.72$; $p < 0.001$), with CK accounting for about 52% of the variance in PCK as a predictive, non-causal association. The CBT generated actionable diagnostic profiles at the dimension and indicator levels. These findings support using diagnostic CBT to guide data-driven curriculum reform integrating content mastery and pedagogical capacity.

Corresponding Author:

INTRODUCTION

The quality of a nation's mathematics education depends heavily on teachers' professional competence. Cross-national studies have shown that mathematics teachers' professional knowledge—comprising Content Knowledge (CK) and Pedagogical Content Knowledge (PCK)—is a primary predictor of instructional effectiveness and students' academic achievement [1], [2]. The 2022 PISA results placed Indonesian students' mathematics literacy 68th out of 81 countries, below the OECD average, indicating a fundamental problem in the quality of mathematics instruction [3]. This condition is inseparable from the quality of professional knowledge formed in teacher education

institutions, so ensuring an adequate professional-knowledge profile of pre-service teachers becomes a strategic imperative [4], [5].

Theoretically, the Mathematical Knowledge for Teaching (MKT) framework distinguishes CK—comprising common, specialized, and horizon content knowledge—from PCK, which encompasses knowledge of content and students, content and teaching, and content and curriculum [6], [7]. Systematic reviews affirm that PCK is the most frequently studied yet the most difficult component to measure consistently because of its contextual and multidimensional nature [8], [9]. Studies in Indonesia show that the professional knowledge of mathematics pre-service teachers varies widely across institutions and curricula, with a notable weakness in the PCK aspect [10], [11]. Similar findings have been reported in cross-national contexts [12], [13].

Alongside the rapid digitalization of education, competency-evaluation systems are now shifting massively from traditional approaches toward technology-based assessment. Various studies indicate that CBT is often more valid and reliable than conventional examination methods [14], [15]; however, its use to objectively measure both the content knowledge and the pedagogical content knowledge of pre-service teachers simultaneously remains very limited because of the complexity of these constructs [16], [17]. The potential of CBT to generate diagnostic feedback based on learning analytics has also not been optimally exploited [18], even though interactive digital feedback has been shown to improve mathematical understanding [19], [20].

The urgency of this study is heightened in the context of Indonesia's uneven higher-education digital transformation [21], which is not yet supported by adequate standardization of assessment systems [22]. Measuring teacher competence in the digital era is required to produce data that are actionable for study-program managers and policymakers [23]. A number of studies affirm that CK and PCK are relatively independent constructs that require different intervention pathways [24], [25], and that teacher effects on student achievement are substantial [26]. Taken together, however, this body of work remains fragmented: CK and PCK are rarely measured simultaneously with a single validated instrument, the diagnostic potential of CBT is seldom exploited at the indicator level, and the evidence is concentrated in a few well-studied national settings, leaving a clear gap for integrated, technology-based diagnosis of both constructs.

Building on this gap, the present study makes three specific contributions rather than a general claim to novelty. Theoretically, it provides empirical evidence on how CK and PCK co-vary at low overall proficiency, showing that their interdependence persists even when neither reaches mastery, an insight that refines, rather than merely applies, the MKT and COACTIV frameworks [6], [25], [27]. Methodologically, it operationalizes a CBT that produces layered, indicator-level diagnostic profiles for both constructs simultaneously, addressing the scarcity of validated instruments that jointly diagnose CK and PCK [28], [29]. Practically, it demonstrates how such diagnostic data can be translated into targeted curriculum and instructional decisions. By drawing on an underrepresented context, the study also contributes comparative evidence to the international literature on teacher knowledge rather than a purely local application of CBT [12], [30]. Accordingly, the study aims to: (1) measure and profile pre-service teachers' CK and PCK; (2) identify strengths and weaknesses at the dimension and indicator levels; and (3) demonstrate how diagnostic profiles can inform instructional intervention.

METHOD

This study employed a quantitative approach with a descriptive-survey and correlational design. The population comprised students of the Mathematics Education study program at a teacher education institution who were in their fifth and seventh semesters. A sample of 30 students was determined through purposive sampling with the inclusion criteria: (1) active students who had adequately completed mathematics-content and pedagogy courses; and (2) willing to take the entire series of CBT tests. Given the limited sample size, all data were analyzed as a single intact group (n = 30) and were not directed toward inter-semester comparison.

Research setting and participants. The research was conducted in the odd semester of the 2024/2025 academic year in the Mathematics Education study program of a state Islamic teacher education institution in Indonesia. Participants were selected using total (saturated) purposive sampling: all 30 students who met the inclusion criterion—having completed the core mathematics-content and pedagogy courses—were included, so that the sample comprised the entire eligible cohort rather than a probability sample. The sample size (n = 30) therefore reflects the actual number of eligible students at the single participating institution and suits the study's diagnostic, exploratory aim of profiling CK and PCK in depth; it is nonetheless acknowledged as a constraint on statistical power and on the generalizability of the correlational and regression results, as elaborated in the Discussion.

The operationalization of the variables comprised two main constructs. CK was defined as mastery of the mathematics content to be taught, measured through 25 multiple-choice items; PCK was defined as the ability to integrate content knowledge with pedagogical strategies, also measured through 25 multiple-choice items. Scores were converted into percentages of achievement relative to the ideal maximum score. The CBT instrument was designed to generate layered diagnostic reports, namely the total score, achievement per dimension, achievement per indicator, and an individual achievement profile for each student. Each item was mapped to a specific indicator and dimension (see the notes below Table 1) so that the CBT system could produce achievement per dimension and per indicator for all students at once, as well as their individual diagnostic profiles. The composition of the instrument is presented in Table 1. The multiple-choice format was chosen to enable fully automated, standardized, and scalable scoring across all participants and indicators at once, a prerequisite for the layered diagnostic reporting; its rationale and limitations are elaborated under ‘Item format’.

Table 1. Composition of CBT Instrument Items by MKT Dimension

Domain	Dimension	Number of Items
CK	Common Content Knowledge (CCK)	7
CK	Specialized Content Knowledge (SCK)	14
CK	Horizon Content Knowledge (HCK)	4
PCK	Knowledge of Content and Students (KCS)	10
PCK	Knowledge of Content and Teaching (KCT)	9
PCK	Knowledge of Content and Curriculum (KCC)	6

Total	50

Notes on dimension codes and indicator coverage (based on the developed CK and PCK instruments). **CK – CCK** (Common Content Knowledge, items 1–7): understanding basic concepts (e.g., lines and angles), performing calculations and measurements accurately, and applying mathematical concepts in everyday life. **CK – SCK** (Specialized Content Knowledge, items 8–21): explaining concepts and principles, analyzing and connecting mathematical objects, proving properties, building models, and solving complex mathematical problems. **CK – HCK** (Horizon Content Knowledge, items 22–25): connecting school mathematics with advanced academic mathematics and with other disciplines. **PCK – KCS** (Knowledge of Content and Students, items 1–10): analyzing students' preconceptions and understanding as well as predicting and analyzing students' misconceptions, errors, and difficulties. **PCK – KCT** (Knowledge of Content and Teaching, items 11–19): selecting and organizing learning materials, selecting sources, strategies, and techniques, and selecting and analyzing representations for presenting material. **PCK – KCC** (Knowledge of Content and Curriculum, items 20–25): designing content and topic combinations, applying standards and curricula, and designing instruments to measure students' abilities.

Instrument development and content validity. The CBT instrument was developed through a systematic procedure. First, a test blueprint was constructed by mapping each item to a specific MKT dimension and indicator (Table 1). Second, the draft items were reviewed by three experts—two mathematics-education lecturers and one educational-assessment specialist—to establish content validity; the level of expert agreement was quantified using Aiken's V, and items with $V \geq 0.80$ were retained while the remaining items were revised or replaced. Third, the revised items were trialed and analyzed empirically before being administered in the present study. To illustrate the item format, each CK item presented a mathematical task with four to five options, whereas each PCK item was framed as a short classroom scenario or a sample of student work followed by options representing alternative pedagogical decisions. The complete test blueprint and sample items are available from the authors upon reasonable request.

Reliability and item analysis. Based on the item analysis, the instrument demonstrated adequate psychometric quality. Internal-consistency reliability computed with the Kuder–Richardson formula (KR-20) reached 0.81 for the CK test and 0.79 for the PCK test, both within the acceptable-to-good range. The item difficulty indices ranged from 0.20 to 0.87 (a balanced mix of easy, medium, and hard items), and the discrimination indices were largely satisfactory ($D \geq 0.30$); only items meeting the validity, difficulty, and discrimination criteria were retained in the final CBT.

Item format. Multiple-choice items were chosen because the CBT was designed for automated, standardized, and scalable diagnosis across many participants and indicators simultaneously. To capture pedagogical reasoning rather than mere recall, the PCK items were built around short classroom scenarios, samples of student work, and instructional-decision situations, with distractors constructed from typical student misconceptions and common teaching errors. It is nevertheless acknowledged that the multiple-choice format cannot fully capture the depth of pedagogical reasoning that open-ended or performance-based tasks would reveal, a limitation revisited in the Discussion.

The analysis techniques comprised descriptive and inferential analyses. Descriptive analysis produced the mean (M), standard deviation (SD), minimum, maximum, median, and percentage of achievement at the total, dimension, indicator, and individual levels, as well as a three-level categorization (High, Medium, Low). The categorization used the following cut-off points: High (≥ 75%), Medium (60–74.99%), and Low (< 60%), where the 75% mark represents the ideal mastery criterion adopted in this study. Inferential analysis was conducted through two techniques: (1) Pearson correlation to test the CK–PCK relationship and the corresponding inter-dimension relationships (CCK–KCS, SCK–

KCT, HCK–KCC) at $\alpha = 0.05$; and (2) simple linear regression with CK as the predictor and PCK as the criterion. Prior to the inferential tests, the underlying assumptions were examined, namely the normality of the CK and PCK distributions (Shapiro–Wilk), the linearity of their relationship, the absence of influential outliers, and homoscedasticity; all assumptions were adequately met, so that the parametric analyses were appropriate.

All scoring, achievement computation, correlation, and linear regression were performed using Microsoft Excel through standardized formulas (including CORREL, SLOPE, INTERCEPT, and RSQ) so that every figure can be traced back to the raw data. The study also observed standard research-ethics principles: participation was entirely voluntary, informed consent was obtained from all respondents before data collection, and the participants' anonymity and confidentiality were maintained throughout the recording, processing, and reporting of data by using codes instead of names, with the individual-profile example presented under the pseudonym Student 1.

RESULTS AND DISCUSSION

3.1. Results

This section reports the findings of the data analysis in the following order: respondent characteristics, descriptive statistics, the distribution of achievement categories, dimension- and indicator-level achievement, the correlation and regression results, and an individual diagnostic profile. The 30 respondents were fifth- and seventh-semester students of the Mathematics Education study program who had completed the core mathematics-content and pedagogy courses; all of them completed the full series of CK and PCK tests through the CBT system.

Overall, students' professional-knowledge achievement had not yet reached the ideal competency threshold (≥ 75%). The mean CK score was 50.00% (SD = 10.16), PCK 52.67% (SD = 23.96), and the combined Professional Knowledge (PK) score 51.33% (SD = 16.04). The complete descriptive statistics are presented in Table 2.

Table 2. Descriptive Statistics of CK, PCK, and Professional Knowledge Scores (n = 30)

Variable	Mean (%)	SD	Min	Max	Median
Content Knowledge (CK)	50.00	10.16	24	68	52.00
Pedagogical Content Knowledge (PCK)	52.67	23.96	8	96	54.00
Professional Knowledge (PK)	51.33	16.04	26	78	50.00

Table 2 shows that the mean PCK is slightly higher than CK (a difference of 2.67 points), an interesting pattern indicating that pedagogical-transformation capacity slightly exceeds content mastery. The large standard deviation of PCK (SD = 23.96) and the much smaller one for CK (SD = 10.16) reflect substantial heterogeneity in knowledge profiles, with score ranges of 24–68% for CK and 8–96% for PCK. The distribution of Professional Knowledge categories is presented in Table 3.

Table 3. Distribution of Professional Knowledge Categories (n = 30)

Category	Frequency (f)	Percentage (%)

High	0	0.00
Medium	12	40.00
Low	18	60.00
Total	30	100.00

All students (100%) were in the Low or Medium category, with 60.00% in the Low category and 40.00% in the Medium category, and no student reaching the High category. This indicates that the majority of pre-service teachers have not attained an optimal level of professional-knowledge mastery. Achievement per dimension is presented in Table 4 and Figure 1.

Table 4. Average Achievement per CK and PCK Dimension

Domain	Dimension	Items	Mean (%)	SD
CK	CCK	7	73.33	20.47
CK	SCK	14	36.91	24.16
CK	HCK	4	55.00	35.60
PCK	KCS	10	54.07	33.18
PCK	KCT	9	57.50	26.14
PCK	KCC	6	46.25	32.30

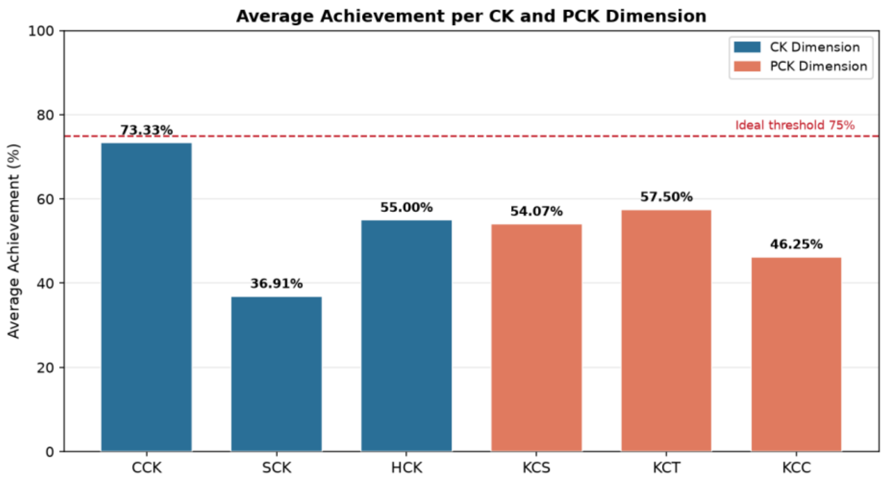


Figure 1. Average Achievement per CK and PCK Dimension

In the CK domain, the CCK dimension was the highest (73.33%) and SCK the lowest (36.91%); in the PCK domain, KCT was the highest (57.50%) while KCC was the lowest of all six dimensions (46.25%). Diagnostically, the low SCK and KCC signal a weak ability to relate content to the structure and organization of the curriculum—an area that needs to be prioritized in intervention. Achievement at the indicator level is presented in Table 5 and visualized in Figure 2.

Table 5. Average Achievement per CK and PCK Indicator

Domain	Indicator	Mean (%)
CK	CCK1	86.67
CK	CCK2	80.00
CK	CCK3	46.67
CK	SCK1	63.33
CK	SCK2	46.67
CK	SCK3	43.33
CK	SCK4	20.00
CK	SCK5	23.33

CK	SCK6	30.00
CK	SCK7	33.33
CK	HCK1	53.33
CK	HCK2	56.67
PCK	KCS1	58.33
PCK	KCS2	48.33
PCK	KCS3	58.33
PCK	KCS4	51.67
PCK	KCS5	53.33
PCK	KCT1	55.00
PCK	KCT2	66.67
PCK	KCT3	60.00
PCK	KCT4	36.67
PCK	KCC1	51.67
PCK	KCC2	43.33
PCK	KCC3	45.00

Notes on indicator codes (based on the CK and PCK instruments). CK indicators — **CCK1**: understanding basic school-mathematics concepts, such as lines, angles, sequences, and basic statistics (items 1–3); **CCK2**: performing calculations and measurements accurately (items 4–5); **CCK3**: applying mathematical concepts in everyday life (items 6–7); **SCK1**: explaining academic mathematical concepts and principles (items 8–9); **SCK2**: analyzing mathematical objects (items 10–11); **SCK3**: connecting mathematical properties and objects (items 12–13); **SCK4**: proving mathematical properties or theorems (items 14–16); **SCK5**: applying proof strategies (item 17); **SCK6**: building and constructing mathematical models (items 18–19); **SCK7**: solving complex mathematical problems (items 20–21); **HCK1**: connecting school mathematics with advanced academic mathematics (items 22–23); **HCK2**: connecting mathematics with other disciplines (items 24–25). PCK indicators — **KCS1**: analyzing students' preconceptions (items 1–2); **KCS2**: identifying students' level of understanding (items 3–4); **KCS3**: predicting students' misconceptions (items 5–6); **KCS4**: identifying sources of students' learning difficulties (items 7–8); **KCS5**: analyzing students' literacy issues and errors (items 9–10); **KCT1**: organizing and sequencing learning materials (items 11–12); **KCT2**: selecting learning sources, strategies, and techniques (items 13–14); **KCT3**: selecting representations and illustrations for presenting material (items 15–17); **KCT4**: analyzing the weaknesses and strengths of representations (items 18–19); **KCC1**: composing content and topic combinations (items 20–21); **KCC2**: applying standards and curricula (items 22–23); **KCC3**: designing instruments to measure students' abilities (items 24–25).

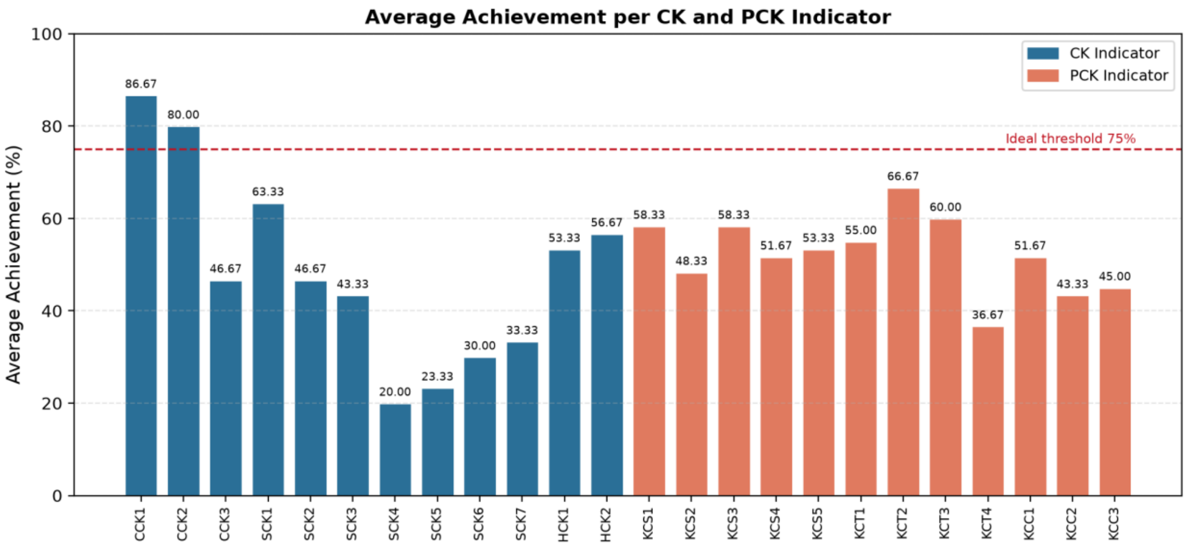


Figure 2. Average Achievement per CK and PCK Indicator

Figure 2 visualizes the average achievement of each indicator exactly as listed in Table 5. The range of achievement across indicators spans 20.00% to 86.67%. The highest achievement is on the CK indicator CCK1 (86.67%; understanding basic school-mathematics concepts), which is fundamental and therefore the most widely mastered by students. Conversely, the lowest overall achievement is on indicator SCK4 (20.00%; proving mathematical properties or theorems)—the skill with the highest cognitive demand in the CK domain—followed by SCK5 (23.33%; applying proof strategies) and SCK6 (30.00%; building and constructing mathematical models). The pattern in the SCK dimension declines gradually with complexity, from explaining concepts and principles (SCK1; 63.33%), analyzing and connecting mathematical objects (SCK2 46.67%; SCK3 43.33%), solving complex problems (SCK7; 33.33%), to the most difficult modeling and proving (SCK6 30.00%; SCK5 23.33%; SCK4 20.00%)—a pattern consistent with cognitive load theory. Because each dimension's achievement remains the mean of its constituent indicators—for example, KCS (54.07%) is the mean of KCS1–KCS5 and CCK (73.33%) the mean of CCK1–CCK3—this pattern maps precisely which items and indicators are the sources of weakness, so that lecturers can design remediation targeted at concrete rather than generic weaknesses. The results of the Pearson correlation test are presented in Table 6.

Table 6. Pearson Correlation between Domains and between Parallel Dimensions

Relationship	r	p	Interpretation
CK – PCK	0.72	< 0.001	Strong
CCK – KCS	0.42	< 0.05	Moderate
SCK – KCT	0.54	< 0.01	Moderate
HCK – KCC	0.86	< 0.001	Very strong

The strong positive correlation between CK and PCK ($r = 0.72$; $p < 0.001$) confirms that the two domains are constructively interdependent, reinforced by the correlations between parallel dimensions, especially HCK–KCC ($r = 0.86$) and SCK–KCT ($r = 0.54$). The predictive relationship of CK on PCK was tested through simple linear regression (Table 7 and Figure 3).

Table 7. Results of the Simple Linear Regression of CK on PCK

Parameter	Value
Slope (b)	1.708
Intercept (a)	−32.734
Coefficient of determination (R^2)	0.5245
Pearson r	0.72
Equation	$PCK = 1.708 \times CK - 32.734$

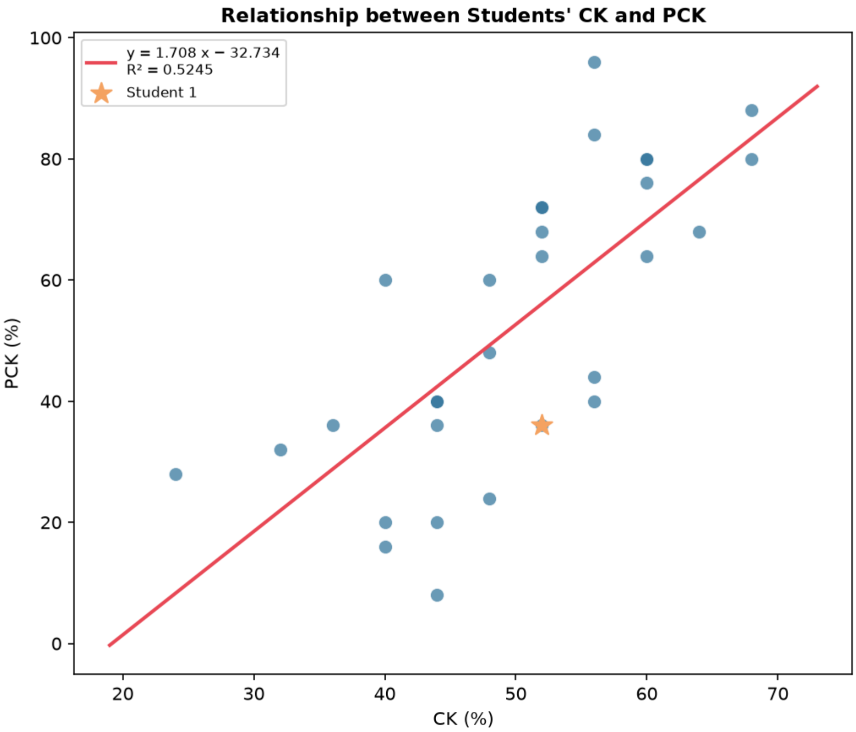


Figure 3. Scatter Plot and Regression Line of the CK–PCK Relationship

The equation PCK = 1.708 × CK − 32.734 with $R^2 = 0.52$ indicates that about 52% of the variance in PCK is statistically associated with CK. Because the design is correlational, this relationship should be interpreted as a predictive association rather than a causal effect; the remaining 48% of the variance signals that PCK does not form automatically from content mastery but requires an explicit pedagogical learning pathway.

The CBT also breaks aggregate achievement down into individual diagnostic profiles automatically. To illustrate this capability—rather than to draw general conclusions from a single case—Table 8 and Figure 4 summarize the profile of one participant (pseudonym Student 1), and Figures 5 and 6 show anonymized screenshots of the corresponding diagnostic report generated by the CBT. Such reports are produced automatically for every participant.

Table 8. Individual Achievement Profile of Student 1 Compared with the Class Average

Domain	Student 1 Score (%)	Class Average (%)	Difference	Category
Content Knowledge (CK)	52.00	50.00	+2.00	Low
Pedagogical Content Knowledge (PCK)	36.00	52.67	−16.67	Low

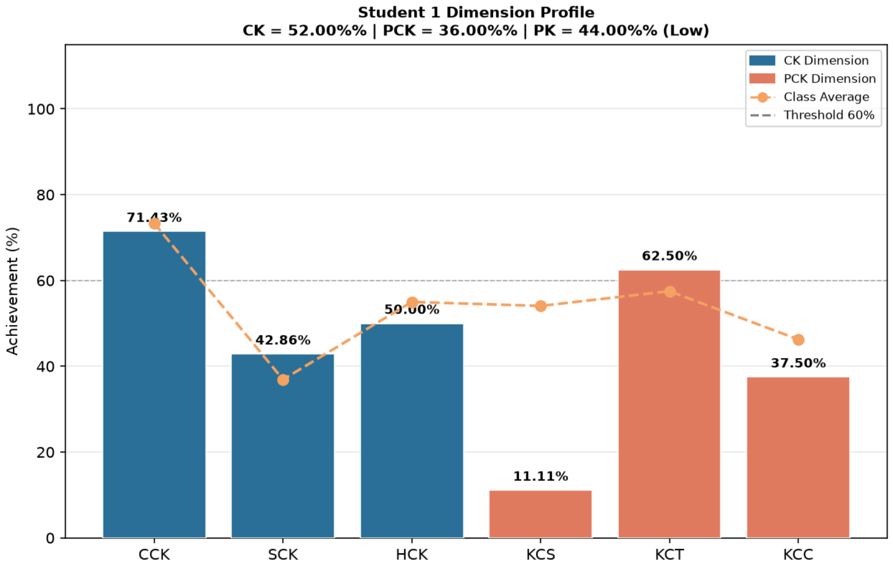
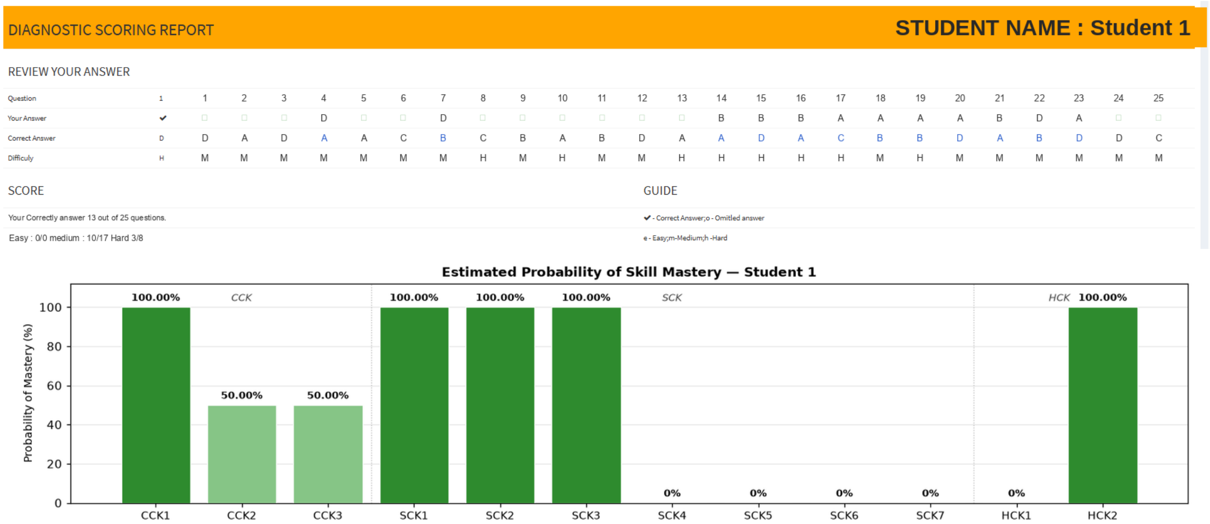


Figure 4. CK and PCK Achievement Profile of Student 1 Compared with the Class Average

As an illustrative example, Student 1 answered 13 of 25 CK items and 9 of 25 PCK items correctly, falling into the Low category for both domains, with CK slightly above and PCK well below the class average. Anonymized screenshots of the report are shown in Figures 5 and 6, with no identifiable student information displayed. This single case is used only to demonstrate the diagnostic output of the CBT and is not intended as a basis for general interpretation.


DIAGNOSTIC SCORING REPORT — STUDENT NAME : Student 1

REVIEW YOUR ANSWER

Question	1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Your Answer	✔	□	□	□	D	□	□	D	□	□	□	□	□	□	B	B	B	A	A	A	A	B	D	A	□	□
Correct Answer	D	D	A	D	A	A	C	B	C	B	A	B	D	A	A	D	A	C	B	B	D	A	B	D	D	C
Difficuly	H	M	M	M	M	M	M	M	H	M	H	M	M	H	H	H	H	H	M	H	M	M	M	M	M	M

SCORE

Your Correctly answer 13 out of 25 questions.

Easy : 0/0 medium : 10/17 Hard 3/8

GUIDE

✔ - Correct Answer;o - Omitled answer

e - Easy;m-Medium;h -Hard

Figure 5. CBT Diagnostic Scorecard for the CK Domain for Student 1 (original system screenshot)

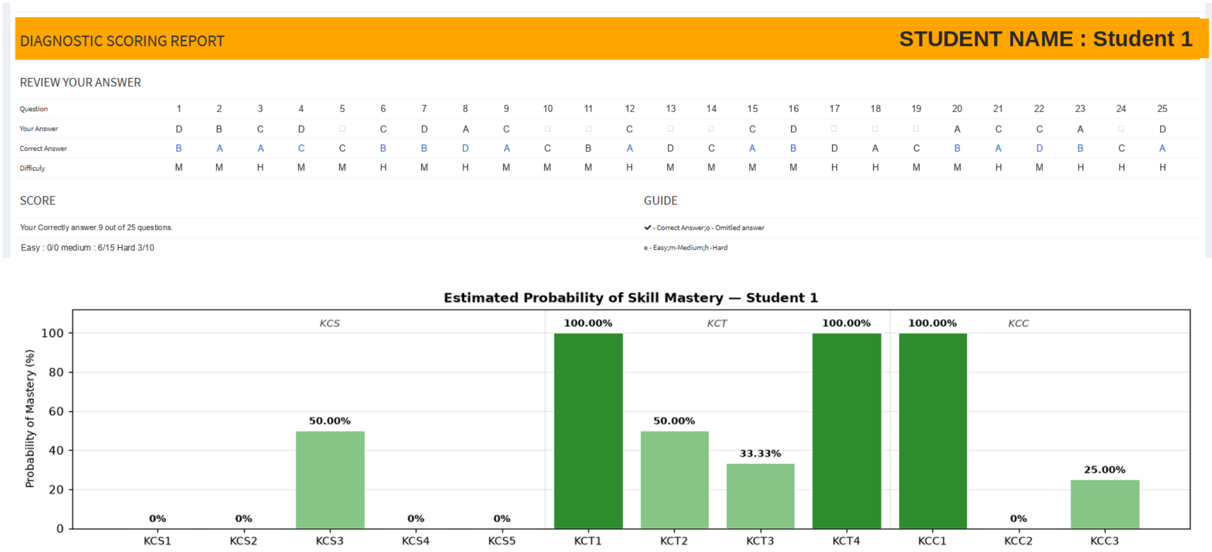


Figure 6. CBT Diagnostic Scorecard for the PCK Domain for Student 1 (system screenshot)

Figures 5 and 6 reveal diagnostic functions of the CBT that are more difficult to implement efficiently in conventional paper-based tests. In addition to recapping correct/incorrect answers per item along with their difficulty level—a combination of the Medium (M) and Hard (H) categories according to the item analysis of the instrument—the system presents an Estimated Probability of Skill Mastery graph that maps Student 1's mastery probability for each skill indicator (CCK1–HCK2 for CK and KCS1–KCC3 for PCK), complete with a mapping of sample items to each indicator. Thus, the CBT does not merely provide a final score but directly identifies which indicators have been mastered and which remain weak—in Student 1's case, several indicators show a mastery probability close to zero—so that lecturers can follow up with specific and immediate remediation. It is this automation of scoring, standardization of reporting, and diagnostic depth that affirms the added value of the CBT in assessing the professional knowledge of mathematics pre-service teachers.

At the indicator level, the same report pinpoints exactly which indicators a student has and has not mastered. For Student 1, for example, several higher-order indicators—advanced reasoning and proof within SCK, and knowledge of students within KCS—registered no mastery, whereas basic content indicators were fully mastered. This illustrates how indicator-level diagnosis can guide individualized remediation far more precisely than dimension averages alone, without implying that this individual profile generalizes to the cohort.

Summary of key findings. In sum, both CK (50.00%) and PCK (52.67%) fell below the 75% mastery threshold; 60.00% of the students were in the low category and none reached the high category; SCK (36.91%) and KCC (46.25%) were the weakest dimensions; CK and PCK were strongly and positively associated ($r = 0.72$; $p < 0.001$); and the CBT was able to decompose these results down to the indicator and individual-student levels, as illustrated by the profile of Student 1.

3.2. Discussion

Interpreted against the study's objectives, three points stand out. First, neither CK nor PCK reached the mastery threshold, and—contrary to the common assumption that content mastery precedes and exceeds pedagogical knowledge—PCK was marginally higher than CK; this pattern resonates with Shulman's argument that transforming content into teachable form is a distinct challenge [7] and with the MKT framework's treatment of CK and PCK as related but non-identical constructs [6]. Second, the weakest areas lay in advanced specialized content within CK and in knowledge of curriculum within PCK, echoing the COACTIV model's claim that these facets critically shape the quality of instructional planning [2], [25]. Third, CK achievement was far more homogeneous than PCK, suggesting that pedagogical capacity develops more unevenly across pre-service teachers, consistent with reviews describing PCK as the most context-dependent and hardest-to-cultivate component [8], [9].

The pattern of correlations indicates that stronger mastery in particular content dimensions tends to co-occur with the corresponding pedagogical capacity, yet content mastery accounts for only part of the variation in PCK—consistent with evidence that pre-service teachers in developing contexts often show comparatively weaker PCK because curricula emphasize content coverage over pedagogical transformation [1], [10]. Rather than forming automatically from content knowledge, PCK appears to require explicit, dedicated learning experiences. The profile of Student 1 is consistent with this interpretation but is treated here only as an illustrative example, not as evidence in its own right.

These results are broadly consistent with prior Indonesian studies reporting that the professional knowledge of mathematics pre-service teachers is generally moderate and uneven, with PCK frequently emerging as the more fragile component [10], [11]. The finding that PCK was marginally higher than CK is noteworthy and can be interpreted cautiously: the PCK items, which draw on familiar classroom situations, may be more accessible to students than the CK items, which demand advanced and abstract mathematical reasoning; moreover, the far larger dispersion of PCK (SD = 23.96) indicates that this apparent advantage is unevenly distributed across the cohort. The particularly low SCK (36.91%) is plausibly linked to teacher-education curricula that emphasize procedural and elementary content over advanced mathematical reasoning—such as formal proof, mathematical modeling, and complex problem solving—so that the more cognitively demanding indicators (SCK4–SCK7) remain underdeveloped [24], [25].

Theoretically, the study offers a specific insight rather than a mere application of existing models: in a low-proficiency, underrepresented setting, CK and PCK remain strongly interdependent even though neither reaches mastery and PCK can slightly exceed CK. This qualifies the intuitive assumption that CK necessarily leads PCK and suggests that, at early proficiency levels, the two constructs may develop in tandem rather than strictly sequentially—extending the MKT and COACTIV frameworks to conditions from which they were not originally derived [6], [25], [27]. Methodologically, the study contributes a CBT model that jointly diagnoses CK and PCK down to the indicator level, addressing the scarcity of instruments that do so simultaneously [22], [28]. In policy terms, the finding that no student reached the high category provides empirical grounds for data-driven rather than assumption-driven curriculum reform [5], [23], a concern underscored by Indonesia's recent international assessment results [3].

Positioned this way, the CBT functions not merely as an examination tool but as a diagnostic-assessment ecosystem that yields rich, decomposable cognitive data actionable for program managers [23], [30]. Several limitations should nonetheless temper these conclusions and caution against generalizing beyond the single institution and the 30 participants studied here: (a) the small sample size, which limits statistical power and generalizability; (b) the total-purposive, single-institution sampling; (c) the bivariate scope, which does not model other determinants simultaneously; (d) the reliance solely on quantitative CBT data, without syllabus and document analysis, classroom observation, or interviews that could triangulate and explain why students underperformed; (e) the limited external evidence of the instrument's validity and reliability beyond the internal analyses reported here; and (f) the multiple-choice format, which cannot fully capture the reasoning that open-ended or performance-based tasks would reveal. In practical terms, these findings imply that teacher-education curricula should allocate explicit, assessed learning time to advanced specialized content (formal proof, mathematical modeling, and complex problem solving) and to knowledge of students and curriculum, and should embed diagnostic CBT within course cycles so that indicator-level weaknesses are identified early and addressed through targeted, evidence-based remediation. Future research should employ larger multi-institution and longitudinal samples, mixed-methods designs that add qualitative triangulation, and multivariate models to strengthen internal and external validity.

CONCLUSION AND RECOMMENDATIONS

This study set out to measure and profile the CK and PCK of mathematics pre-service teachers through a diagnostic CBT. On average, both constructs remained below the mastery threshold, with the majority of students in the low category and none in the high category; the weakest areas were advanced specialized content within CK and knowledge of curriculum within PCK, while CK and PCK were strongly and positively associated. Theoretically, the study clarifies that CK and PCK can remain interdependent—and PCK can slightly exceed CK—even at low overall proficiency; methodologically, it demonstrates a CBT capable of layered, indicator-level diagnosis of both constructs; and practically, it provides a standardized, efficient means of generating actionable competence profiles for teacher education. These conclusions should be read in light of the study's limitations, namely a small, single-institution, purposively selected sample analyzed with bivariate, multiple-choice-based methods. Overall, the central message is that diagnostic computer-based assessment can meaningfully advance the monitoring and development of mathematics pre-service teachers' professional knowledge, provided its profiles are used to inform evidence-based curriculum and instructional reform.

Recommendations. The following recommendations are offered as literature-informed, potentially effective alternatives suggested by—rather than directly proven by—the present findings, and they should be validated in future studies.

For lecturers. Remedial instruction may target not only the weakest dimensions (SCK within CK and KCC within PCK) but also the weakest indicators identified in Table 5. In the CK domain, priority may be given to strengthening mathematical proof—proving properties or theorems (SCK4) and applying proof strategies (SCK5)—through scaffolded proof practice, analysis of proof errors, and habituation to deductive argumentation, followed by mathematical modeling (SCK6) and complex problem solving (SCK7) via problem-based learning. In the PCK domain, the focus may fall on analyzing the strengths and weaknesses of representations (KCT4), applying standards and curricula (KCC2), and designing measurement instruments (KCC3) through case-based learning, analysis of teaching videos, and curriculum-document study.

For teacher education institution curriculum managers. The results encourage integrating a diagnostic CBT—capable of mapping achievement down to the indicator and individual levels—into the study program's quality-assurance cycle, and rebalancing the curriculum so that advanced mathematical reasoning and pedagogical-content competence receive as much emphasis as procedural content mastery.

For future researchers. Future studies may use larger, multi-institution, and longitudinal samples, mixed-methods designs that add qualitative triangulation (syllabus analysis, classroom observation, and interviews), and multivariate models involving additional determinants (such as academic background, learning motivation, and quality of teaching) to strengthen external validity and to explain the causes of the observed weaknesses. A further agenda is to develop and empirically test enhancements to the CBT itself—automatic item-level feedback with worked explanations, direct links to remedial materials for each weak indicator, and adaptive item delivery—so that, once validated, the system can function not only as a diagnostic tool but also as an integrated learning-support environment.

ACKNOWLEDGMENTS

The authors thank the Mathematics Education study program of the teacher education institution for facilitating the implementation of the research, as well as all the students who participated as respondents.

REFERENCES

[1] X. Yang and G. Kaiser, "The impact of mathematics teachers' professional competence on instructional quality and students' mathematics learning outcomes," *Curr. Opin. Behav. Sci.*, vol. 48, p. 101225, Dec. 2022, doi: 10.1016/j.cobeha.2022.101225.

[2] S. Krauss *et al.*, "Pedagogical content knowledge and content knowledge of secondary mathematics teachers.," *J. Educ. Psychol.*, vol. 100, no. 3, pp. 716–725, Aug. 2008, doi: 10.1037/0022-0663.100.3.716.

[3] OECD, *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*. in PISA. OECD Publishing, 2023. doi: 10.1787/53f23881-en.

[4] International Association for the Evaluation of Educational Achievement, *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. Amsterdam: IEA, 2012.

[5] H. Qian and P. Youngs, "The effect of teacher education programs on future elementary mathematics teachers' knowledge: a five-country analysis using TEDS-M data," *J. Math. Teach. Educ.*, vol. 19, no. 4, pp. 371–396, Aug. 2016, doi: 10.1007/s10857-014-9297-0.

[6] D. Loewenberg Ball, M. H. Thames, and G. Phelps, "Content Knowledge for Teaching: What Makes It Special?," *J. Teach. Educ.*, vol. 59, no. 5, pp. 389–407, Nov. 2008, doi: 10.1177/0022487108324554.

[7] L. S. Shulman, "Those Who Understand: Knowledge Growth in Teaching," *Educ. Res.*, vol. 15, no. 2, pp. 4–14, Feb. 1986, doi: 10.3102/0013189X015002004.

[8] M. Grigaliūnienė, E. Lehtinen, L. Verschaffel, and F. Depaepe, "Systematic Review of Research on Pedagogical Content Knowledge in Mathematics: Insights from a Topic-Specific Approach," *ZDM – Math. Educ.*, vol. 57, no. 4, pp. 777–794, Aug. 2025, doi: 10.1007/s11858-025-01684-1.

[9] F. Depaepe, L. Verschaffel, and G. Kelchtermans, "Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research," *Teach. Teach. Educ.*, vol. 34, pp. 12–25, Aug. 2013, doi: 10.1016/j.tate.2013.03.001.

[10] S. Y. Ningsih, Turmudi, and D. Juandi, "Pedagogical content knowledge (PCK) profile of prospective teachers in mathematics learning," *J. Phys. Conf. Ser.*, vol. 1521, no. 3, p. 032057, Apr. 2020, doi: 10.1088/1742-6596/1521/3/032057.

[11] S. Fiangga, S. Khabibah, S. M. Amin, and R. Ekawati, “A Learning Design Analysis of the Pre-service Teachers’ Mathematics Pedagogical Content Knowledge,” *J. Phys. Conf. Ser.*, vol. 1899, no. 1, p. 012153, May 2021, doi: 10.1088/1742-6596/1899/1/012153.

[12] X. Yang, B. Schwarz, and I. K. C. Leung, “Pre-service mathematics teachers’ professional modeling competencies: a comparative study between Germany, Mainland China, and Hong Kong,” *Educ. Stud. Math.*, vol. 109, no. 2, pp. 409–429, Feb. 2022, doi: 10.1007/s10649-021-10064-x.

[13] J. Njiku, “Pre-service teachers manifested mathematics pedagogical content knowledge: The role of the teaching practicum,” *Pedagog. Res.*, vol. 10, no. 1, p. em0229, Jan. 2025, doi: 10.29333/pr/15647.

[14] A. J. Boevé, R. R. Meijer, C. J. Albers, Y. Beetsma, and R. J. Bosker, “Introducing Computer-Based Testing in High-Stakes Exams in Higher Education: Results of a Field Experiment,” *PLOS ONE*, vol. 10, no. 12, p. e0143616, Dec. 2015, doi: 10.1371/journal.pone.0143616.

[15] C. Y. Piaw, “Replacing Paper-based Testing with Computer-based Testing in Assessment: Are we Doing Wrong?,” *Procedia - Soc. Behav. Sci.*, vol. 64, pp. 655–664, Nov. 2012, doi: 10.1016/j.sbspro.2012.11.077.

[16] A. Berry, P. Friedrichsen, and J. Loughran, Eds., *Re-examining pedagogical content knowledge in science education*. in Teaching and learning in science series. London New York: Routledge, 2015. doi: 10.4324/9781315735665.

[17] M. Schmid, E. Brianza, and D. Petko, “Self-reported technological pedagogical content knowledge (TPACK) of pre-service teachers in relation to digital technology use in lesson plans,” *Comput. Hum. Behav.*, vol. 115, p. 106586, Feb. 2021, doi: 10.1016/j.chb.2020.106586.

[18] L. Maas, M. J. S. Brinkhuis, L. Kester, and L. Wijngaards-de Meij, “Cognitive Diagnostic Assessment in University Statistics Education: Valid and Reliable Skill Measurement for Actionable Feedback Using Learning Dashboards,” *Appl. Sci.*, vol. 12, no. 10, p. 4809, May 2022, doi: 10.3390/app12104809.

[19] A. Barana, M. Marchisio, and M. Sacchet, “Interactive Feedback for Learning Mathematics in a Digital Learning Environment,” *Educ. Sci.*, vol. 11, no. 6, p. 279, Jun. 2021, doi: 10.3390/educsci11060279.

[20] S. Rezat, “How automated feedback from a digital mathematics textbook affects primary students’ conceptual development: two case studies,” *ZDM – Math. Educ.*, vol. 53, no. 6, pp. 1433–1445, Nov. 2021, doi: 10.1007/s11858-021-01263-0.

[21] L. N. Yaqin, L. D. Prasojo, N. A. Haji-Othman, N. Yusof, and A. Habibi, “Addressing the Digital Divide in Indonesian Higher Education: Insights, Implications, and Potential Solutions,” in *From Digital Divide to Digital Inclusion*, Ł. Tomczyk, F. D. Guillén-Gámez, J. Ruiz-Palmero, and A. Habibi, Eds., in Lecture Notes in Educational Technology. , Singapore: Springer Nature Singapore, 2023, pp. 291–307. doi: 10.1007/978-981-99-7645-4_13.

[22] Naila Rajiha and S. Rini, “COMPUTER-BASED STANDARDIZED TESTING (CBT) SUMMATIVE ASSESSMENT: ONE INSTITUTION EXPERIENCE,” *JISAE J. Indones. Stud. Assess. Eval.*, vol. 9, no. 1, pp. 1–7, Jan. 2023, doi: 10.21009/jisae.v9i1.33233.

[23] G. Kaiser and J. König, “Competence Measurement in (Mathematics) Teacher Education and Beyond: Implications for Policy,” *High. Educ. Policy*, vol. 32, no. 4, pp. 597–615, Dec. 2019, doi: 10.1057/s41307-019-00139-z.

[24] T. Kleickmann *et al.*, “Teachers’ Content Knowledge and Pedagogical Content Knowledge: The Role of Structural Differences in Teacher Education,” *J. Teach. Educ.*, vol. 64, no. 1, pp. 90–106, Jan. 2013, doi: 10.1177/0022487112460398.

[25] J. Baumert and M. Kunter, “The COACTIV Model of Teachers’ Professional Competence,” in *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers*, M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, and M. Neubrand, Eds., Boston, MA: Springer US, 2013, pp. 25–48. doi: 10.1007/978-1-4614-5149-5_2.

[26] B. Rowan, R. Correnti, and R. J. Miller, “What Large-Scale, Survey Research Tells Us about Teacher Effects on Student Achievement: Insights from the Prospects Study of Elementary Schools,” *Teach. Coll. Rec. Voice Scholarsh. Educ.*, vol. 104, no. 8, pp. 1525–1567, Aug. 2002, doi: 10.1111/1467-9620.00212.

[27] S. Blömeke, J.-E. Gustafsson, and R. J. Shavelson, “Beyond Dichotomies: Competence Viewed as a Continuum,” *Z. Für Psychol.*, vol. 223, no. 1, pp. 3–13, Jan. 2015, doi: 10.1027/2151-2604/a000194.

[28] E. Mohammadpour and Y. Maroofi, “A performance-based test to measure teachers’ mathematics and science content and pedagogical knowledge,” *Heliyon*, vol. 9, no. 3, p. e13932, Mar. 2023, doi: 10.1016/j.heliyon.2023.e13932.

[29] T. Clark and H. Endres, “Computer-based diagnostic assessment of high school students’ grammar skills with automated feedback – an international trial,” *Assess. Educ. Princ. Policy Pract.*, vol. 28, no. 5–6, pp. 602–632, Nov. 2021, doi: 10.1080/0969594X.2021.1970513.

[30] G. A. Nortvedt and N. Buchholtz, “Assessment in mathematics education: responding to issues regarding methodology, policy, and equity,” *ZDM*, vol. 50, no. 4, pp. 555–570, Aug. 2018, doi: 10.1007/s11858-018-0963-z.

Bukti Review ke-3 oleh Reviewer

Journal of Mathematics Instruction, Social Research and Opinion
Vol. xx, No. x, March/June/September/December 202x, pp. 1-1x, https://doi.org/
ISSN 2962-7842

Assessing Mathematics Pre-Service Teachers' Content Knowledge and Pedagogical Content Knowledge through a Diagnostic Computer-Based Test

Article Info

Article history:

Received mm dd, yyyy
Revised mm dd, yyyy
Accepted mm dd, yyyy

Keywords:

Computer-based test
Content knowledge
Diagnostic assessment
Mathematical knowledge for teaching
Mathematics pre-service teachers
Pedagogical content knowledge

ABSTRACT

Content Knowledge (CK) and Pedagogical Content Knowledge (PCK) are core components of mathematics teachers' competence, yet they are rarely measured simultaneously through technology-based diagnostic assessment. This study measures and profiles the CK and PCK of mathematics pre-service teachers using a diagnostic Computer-Based Test (CBT). A quantitative descriptive-correlational design was applied to 30 purposively selected pre-service teachers. Each construct was measured with 25 multiple-choice items whose content validity was established by expert judgment and whose reliability was confirmed before use (KR-20 = 0.81 for CK; 0.79 for PCK). Data were analyzed with descriptive statistics, Pearson correlation, and simple linear regression after assumption checks. Mean CK (50.00%) and PCK (52.67%) were both below the 75% mastery threshold, with 60% of students in the low category; specialized content knowledge and knowledge of content and curriculum were the weakest dimensions. CK and PCK correlated strongly and positively ($r = 0.72$; $p < 0.001$), with CK accounting for about 52% of the variance in PCK as a predictive, non-causal association. The CBT generated actionable diagnostic profiles at the dimension and indicator levels. These findings support using diagnostic CBT to guide data-driven curriculum reform integrating content mastery and pedagogical capacity.

Corresponding Author:

Commented [AM1]: According to Table 3, the 60% figure appears to refer to the combined Professional Knowledge category rather than separate CK and PCK categories. This should not be presented as though 60% were low in both constructs.
The abstract reports a 75% mastery threshold but does not explain its basis.
The authors should avoid calling 75% an established mastery threshold unless it was determined through:
- an Angoff procedure;
- bookmark standard-setting;
- institutional competency standards;
- expert consensus;
- or external criterion validation

INTRODUCTION

The quality of a nation's mathematics education depends heavily on teachers' professional competence. Cross-national studies have shown that mathematics teachers' professional knowledge—comprising Content Knowledge (CK) and Pedagogical Content Knowledge (PCK)—is a primary predictor of instructional effectiveness and students' academic achievement [1], [2]. The 2022 PISA results placed Indonesian students' mathematics literacy 68th out of 81 countries, below the OECD average, indicating a fundamental problem in the quality of mathematics instruction [3]. This condition is inseparable from the quality of professional knowledge formed in teacher education

institutions, so ensuring an adequate professional-knowledge profile of pre-service teachers becomes a strategic imperative [4], [5].

Theoretically, the Mathematical Knowledge for Teaching (MKT) framework distinguishes CK—comprising common, specialized, and horizon content knowledge—from PCK, which encompasses knowledge of content and students, content and teaching, and content and curriculum [6], [7]. Systematic reviews affirm that PCK is the most frequently studied yet the most difficult component to measure consistently because of its contextual and multidimensional nature [8], [9]. Studies in Indonesia show that the professional knowledge of mathematics pre-service teachers varies widely across institutions and curricula, with a notable weakness in the PCK aspect [10], [11]. Similar findings have been reported in cross-national contexts [12], [13].

Alongside the rapid digitalization of education, competency-evaluation systems are now shifting massively from traditional approaches toward technology-based assessment. Various studies indicate that CBT is often more valid and reliable than conventional examination methods [14], [15]; however, its use to objectively measure both the content knowledge and the pedagogical content knowledge of pre-service teachers simultaneously remains very limited because of the complexity of these constructs [16], [17]. The potential of CBT to generate diagnostic feedback based on learning analytics has also not been optimally exploited [18], even though interactive digital feedback has been shown to improve mathematical understanding [19], [20].

The urgency of this study is heightened in the context of Indonesia's uneven higher-education digital transformation [21], which is not yet supported by adequate standardization of assessment systems [22]. Measuring teacher competence in the digital era is required to produce data that are actionable for study-program managers and policymakers [23]. A number of studies affirm that CK and PCK are relatively independent constructs that require different intervention pathways [24], [25], and that teacher effects on student achievement are substantial [26]. Taken together, however, this body of work remains fragmented: CK and PCK are rarely measured simultaneously with a single validated instrument, the diagnostic potential of CBT is seldom exploited at the indicator level, and the evidence is concentrated in a few well-studied national settings, leaving a clear gap for integrated, technology-based diagnosis of both constructs.

Building on this gap, the present study makes three specific contributions rather than a general claim to novelty. Theoretically, it provides empirical evidence on how CK and PCK co-vary at low overall proficiency, showing that their interdependence persists even when neither reaches mastery, an insight that refines, rather than merely applies, the MKT and COACTIV frameworks [6], [25], [27]. Methodologically, it operationalizes a CBT that produces layered, indicator-level diagnostic profiles for both constructs simultaneously, addressing the scarcity of validated instruments that jointly diagnose CK and PCK [28], [29]. Practically, it demonstrates how such diagnostic data can be translated into targeted curriculum and instructional decisions. By drawing on an underrepresented context, the study also contributes comparative evidence to the international literature on teacher knowledge rather than a purely local application of CBT [12], [30]. Accordingly, the study aims to: (1) measure and profile pre-service teachers' CK and PCK; (2) identify strengths and weaknesses at the dimension and indicator levels; and (3) demonstrate how diagnostic profiles can inform instructional intervention.

METHOD

This study employed a quantitative approach with a descriptive-survey and correlational design. The population comprised students of the Mathematics Education study program at a teacher education institution who were in their fifth and seventh semesters. A sample of 30 students was determined through purposive sampling with the inclusion criteria: (1) active students who had adequately completed mathematics-content and pedagogy courses; and (2) willing to take the entire series of CBT tests. Given the limited sample size, all data were analyzed as a single intact group (n = 30) and were not directed toward inter-semester comparison.

Research setting and participants. The research was conducted in the odd semester of the 2024/2025 academic year in the Mathematics Education study program of a state Islamic teacher education institution in Indonesia. Participants were selected using total (saturated) purposive sampling: all 30 students who met the inclusion criterion—having completed the core mathematics-content and pedagogy courses—were included, so that the sample comprised the entire eligible cohort rather than a probability sample. The sample size (n = 30) therefore reflects the actual number of eligible students at the single participating institution and suits the study's diagnostic, exploratory aim of profiling CK and PCK in depth; it is nonetheless acknowledged as a constraint on statistical power and on the generalizability of the correlational and regression results, as elaborated in the Discussion.

The operationalization of the variables comprised two main constructs. CK was defined as mastery of the mathematics content to be taught, measured through 25 multiple-choice items; PCK was defined as the ability to integrate content knowledge with pedagogical strategies, also measured through 25 multiple-choice items. Scores were converted into percentages of achievement relative to the ideal maximum score. The CBT instrument was designed to generate layered diagnostic reports, namely the total score, achievement per dimension, achievement per indicator, and an individual achievement profile for each student. Each item was mapped to a specific indicator and dimension (see the notes below Table 1) so that the CBT system could produce achievement per dimension and per indicator for all students at once, as well as their individual diagnostic profiles. The composition of the instrument is presented in Table 1. The multiple-choice format was chosen to enable fully automated, standardized, and scalable scoring across all participants and indicators at once, a prerequisite for the layered diagnostic reporting; its rationale and limitations are elaborated under ‘Item format’.

Table 1. Composition of CBT Instrument Items by MKT Dimension

Domain	Dimension	Number of Items
CK	Common Content Knowledge (CCK)	7
CK	Specialized Content Knowledge (SCK)	14
CK	Horizon Content Knowledge (HCK)	4
PCK	Knowledge of Content and Students (KCS)	10
PCK	Knowledge of Content and Teaching (KCT)	9
PCK	Knowledge of Content and Curriculum (KCC)	6

Total	50

Notes on dimension codes and indicator coverage (based on the developed CK and PCK instruments). **CK – CCK** (Common Content Knowledge, items 1–7): understanding basic concepts (e.g., lines and angles), performing calculations and measurements accurately, and applying mathematical concepts in everyday life. **CK – SCK** (Specialized Content Knowledge, items 8–21): explaining concepts and principles, analyzing and connecting mathematical objects, proving properties, building models, and solving complex mathematical problems. **CK – HCK** (Horizon Content Knowledge, items 22–25): connecting school mathematics with advanced academic mathematics and with other disciplines. **PCK – KCS** (Knowledge of Content and Students, items 1–10): analyzing students' preconceptions and understanding as well as predicting and analyzing students' misconceptions, errors, and difficulties. **PCK – KCT** (Knowledge of Content and Teaching, items 11–19): selecting and organizing learning materials, selecting sources, strategies, and techniques, and selecting and analyzing representations for presenting material. **PCK – KCC** (Knowledge of Content and Curriculum, items 20–25): designing content and topic combinations, applying standards and curricula, and designing instruments to measure students' abilities.

Instrument development and content validity. The CBT instrument was developed through a systematic procedure. First, a test blueprint was constructed by mapping each item to a specific MKT dimension and indicator (Table 1). Second, the draft items were reviewed by three experts—two mathematics-education lecturers and one educational-assessment specialist—to establish content validity; the level of expert agreement was quantified using Aiken's V, and items with $V \geq 0.80$ were retained while the remaining items were revised or replaced. Third, the revised items were trialed and analyzed empirically before being administered in the present study. To illustrate the item format, each CK item presented a mathematical task with four to five options, whereas each PCK item was framed as a short classroom scenario or a sample of student work followed by options representing alternative pedagogical decisions. The complete test blueprint and sample items are available from the authors upon reasonable request.

Reliability and item analysis. Based on the item analysis, the instrument demonstrated adequate psychometric quality. Internal-consistency reliability computed with the Kuder–Richardson formula (KR-20) reached 0.81 for the CK test and 0.79 for the PCK test, both within the acceptable-to-good range. The item difficulty indices ranged from 0.20 to 0.87 (a balanced mix of easy, medium, and hard items), and the discrimination indices were largely satisfactory ($D \geq 0.30$); only items meeting the validity, difficulty, and discrimination criteria were retained in the final CBT.

Item format. Multiple-choice items were chosen because the CBT was designed for automated, standardized, and scalable diagnosis across many participants and indicators simultaneously. To capture pedagogical reasoning rather than mere recall, the PCK items were built around short classroom scenarios, samples of student work, and instructional-decision situations, with distractors constructed from typical student misconceptions and common teaching errors. It is nevertheless acknowledged that the multiple-choice format cannot fully capture the depth of pedagogical reasoning that open-ended or performance-based tasks would reveal, a limitation revisited in the Discussion.

The analysis techniques comprised descriptive and inferential analyses. Descriptive analysis produced the mean (M), standard deviation (SD), minimum, maximum, median, and percentage of achievement at the total, dimension, indicator, and individual levels, as well as a three-level categorization (High, Medium, Low). The categorization used the following cut-off points: High (≥ 75%), Medium (60–74.99%), and Low (< 60%), where the 75% mark represents the ideal mastery criterion adopted in this study. Inferential analysis was conducted through two techniques: (1) Pearson correlation to test the CK–PCK relationship and the corresponding inter-dimension relationships (CCK–KCS, SCK–

KCT, HCK–KCC) at α = 0.05; and (2) simple linear regression with CK as the predictor and PCK as the criterion. Prior to the inferential tests, the underlying assumptions were examined, namely the normality of the CK and PCK distributions (Shapiro–Wilk), the linearity of their relationship, the absence of influential outliers, and homoscedasticity; all assumptions were adequately met, so that the parametric analyses were appropriate.

All scoring, achievement computation, correlation, and linear regression were performed using Microsoft Excel through standardized formulas (including CORREL, SLOPE, INTERCEPT, and RSQ) so that every figure can be traced back to the raw data. The study also observed standard research-ethics principles: participation was entirely voluntary, informed consent was obtained from all respondents before data collection, and the participants' anonymity and confidentiality were maintained throughout the recording, processing, and reporting of data by using codes instead of names, with the individual-profile example presented under the pseudonym Student 1.

RESULTS AND DISCUSSION

3.1. Results

This section reports the findings of the data analysis in the following order: respondent characteristics, descriptive statistics, the distribution of achievement categories, dimension- and indicator-level achievement, the correlation and regression results, and an individual diagnostic profile. The 30 respondents were fifth- and seventh-semester students of the Mathematics Education study program who had completed the core mathematics-content and pedagogy courses; all of them completed the full series of CK and PCK tests through the CBT system.

Overall, students' professional-knowledge achievement had not yet reached the ideal competency threshold (≥ 75%). The mean CK score was 50.00% (SD = 10.16), PCK 52.67% (SD = 23.96), and the combined Professional Knowledge (PK) score 51.33% (SD = 16.04). The complete descriptive statistics are presented in Table 2.

Table 2. Descriptive Statistics of CK, PCK, and Professional Knowledge Scores (n = 30)

Variable	Mean (%)	SD	Min	Max	Median
Content Knowledge (CK)	50.00	10.16	24	68	52.00
Pedagogical Content Knowledge (PCK)	52.67	23.96	8	96	54.00
Professional Knowledge (PK)	51.33	16.04	26	78	50.00

Table 2 shows that the mean PCK is slightly higher than CK (a difference of 2.67 points), an interesting pattern indicating that pedagogical-transformation capacity slightly exceeds content mastery. The large standard deviation of PCK (SD = 23.96) and the much smaller one for CK (SD = 10.16) reflect substantial heterogeneity in knowledge profiles, with score ranges of 24–68% for CK and 8–96% for PCK. The distribution of Professional Knowledge categories is presented in Table 3.

Table 3. Distribution of Professional Knowledge Categories (n = 30)

Category	Frequency (f)	Percentage (%)

High	0	0.00
Medium	12	40.00
Low	18	60.00
Total	30	100.00

All students (100%) were in the Low or Medium category, with 60.00% in the Low category and 40.00% in the Medium category, and no student reaching the High category. This indicates that the majority of pre-service teachers have not attained an optimal level of professional-knowledge mastery. Achievement per dimension is presented in Table 4 and Figure 1.

Table 4. Average Achievement per CK and PCK Dimension

Domain	Dimension	Items	Mean (%)	SD
CK	CCK	7	73.33	20.47
CK	SCK	14	36.91	24.16
CK	HCK	4	55.00	35.60
PCK	KCS	10	54.07	33.18
PCK	KCT	9	57.50	26.14
PCK	KCC	6	46.25	32.30

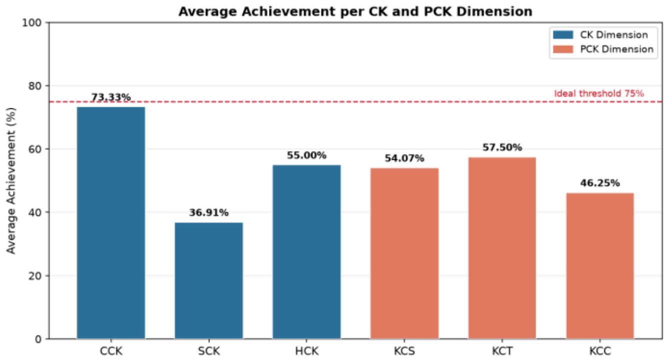


Figure 1. Average Achievement per CK and PCK Dimension

In the CK domain, the CCK dimension was the highest (73.33%) and SCK the lowest (36.91%); in the PCK domain, KCT was the highest (57.50%) while KCC was the lowest of all six dimensions (46.25%). Diagnostically, the low SCK and KCC signal a weak ability to relate content to the structure and organization of the curriculum—an area that needs to be prioritized in intervention. Achievement at the indicator level is presented in Table 5 and visualized in Figure 2.

Table 5. Average Achievement per CK and PCK Indicator

Domain	Indicator	Mean (%)
CK	CCK1	86.67
CK	CCK2	80.00
CK	CCK3	46.67
CK	SCK1	63.33
CK	SCK2	46.67
CK	SCK3	43.33
CK	SCK4	20.00
CK	SCK5	23.33

CK	SCK6	30.00
CK	SCK7	33.33
CK	HCK1	53.33
CK	HCK2	56.67
PCK	KCS1	58.33
PCK	KCS2	48.33
PCK	KCS3	58.33
PCK	KCS4	51.67
PCK	KCS5	53.33
PCK	KCT1	55.00
PCK	KCT2	66.67
PCK	KCT3	60.00
PCK	KCT4	36.67
PCK	KCC1	51.67
PCK	KCC2	43.33
PCK	KCC3	45.00

Notes on indicator codes (based on the CK and PCK instruments). CK indicators — **CCK1**: understanding basic school-mathematics concepts, such as lines, angles, sequences, and basic statistics (items 1–3); **CCK2**: performing calculations and measurements accurately (items 4–5); **CCK3**: applying mathematical concepts in everyday life (items 6–7); **SCK1**: explaining academic mathematical concepts and principles (items 8–9); **SCK2**: analyzing mathematical objects (items 10–11); **SCK3**: connecting mathematical properties and objects (items 12–13); **SCK4**: proving mathematical properties or theorems (items 14–16); **SCK5**: applying proof strategies (item 17); **SCK6**: building and constructing mathematical models (items 18–19); **SCK7**: solving complex mathematical problems (items 20–21); **HCK1**: connecting school mathematics with advanced academic mathematics (items 22–23); **HCK2**: connecting mathematics with other disciplines (items 24–25). PCK indicators — **KCS1**: analyzing students' preconceptions (items 1–2); **KCS2**: identifying students' level of understanding (items 3–4); **KCS3**: predicting students' misconceptions (items 5–6); **KCS4**: identifying sources of students' learning difficulties (items 7–8); **KCS5**: analyzing students' literacy issues and errors (items 9–10); **KCT1**: organizing and sequencing learning materials (items 11–12); **KCT2**: selecting learning sources, strategies, and techniques (items 13–14); **KCT3**: selecting representations and illustrations for presenting material (items 15–17); **KCT4**: analyzing the weaknesses and strengths of representations (items 18–19); **KCC1**: composing content and topic combinations (items 20–21); **KCC2**: applying standards and curricula (items 22–23); **KCC3**: designing instruments to measure students' abilities (items 24–25).

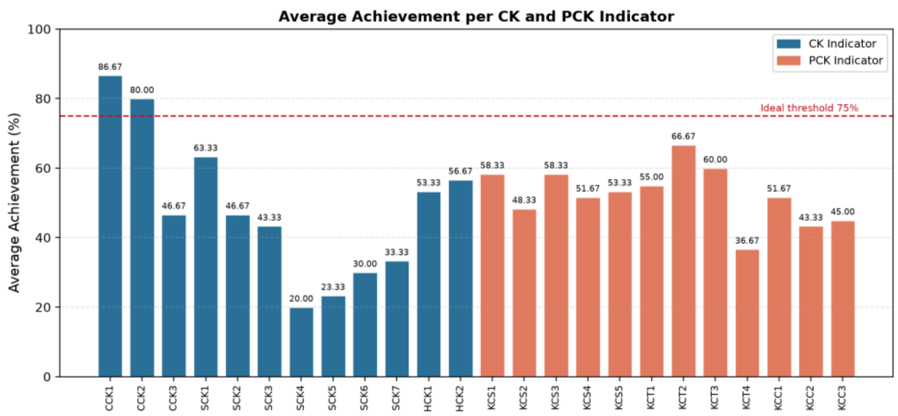


Figure 2. Average Achievement per CK and PCK Indicator

Figure 2 visualizes the average achievement of each indicator exactly as listed in Table 5. The range of achievement across indicators spans 20.00% to 86.67%. The highest achievement is on the CK indicator CCK1 (86.67%; understanding basic school-mathematics concepts), which is fundamental and therefore the most widely mastered by students. Conversely, the lowest overall achievement is on indicator SCK4 (20.00%; proving mathematical properties or theorems)—the skill with the highest cognitive demand in the CK domain—followed by SCK5 (23.33%; applying proof strategies) and SCK6 (30.00%; building and constructing mathematical models). The pattern in the SCK dimension declines gradually with complexity, from explaining concepts and principles (SCK1; 63.33%), analyzing and connecting mathematical objects (SCK2 46.67%; SCK3 43.33%), solving complex problems (SCK7; 33.33%), to the most difficult modeling and proving (SCK6 30.00%; SCK5 23.33%; SCK4 20.00%)—a pattern consistent with cognitive load theory. Because each dimension's achievement remains the mean of its constituent indicators—for example, KCS (54.07%) is the mean of KCS1–KCS5 and CCK (73.33%) the mean of CCK1–CCK3—this pattern maps precisely which items and indicators are the sources of weakness, so that lecturers can design remediation targeted at concrete rather than generic weaknesses. The results of the Pearson correlation test are presented in Table 6.

Table 6. Pearson Correlation between Domains and between Parallel Dimensions

Relationship	r	p	Interpretation
CK – PCK	0.72	< 0.001	Strong
CCK – KCS	0.42	< 0.05	Moderate
SCK – KCT	0.54	< 0.01	Moderate
HCK – KCC	0.86	< 0.001	Very strong

The strong positive correlation between CK and PCK ($r = 0.72$; $p < 0.001$) confirms that the two domains are constructively interdependent, reinforced by the correlations between parallel dimensions, especially HCK–KCC ($r = 0.86$) and SCK–KCT ($r = 0.54$). The predictive relationship of CK on PCK was tested through simple linear regression (Table 7 and Figure 3).

Table 7. Results of the Simple Linear Regression of CK on PCK

Parameter	Value
Slope (b)	1.708
Intercept (a)	−32.734
Coefficient of determination (R^2)	0.5245
Pearson r	0.72
Equation	$PCK = 1.708 \times CK - 32.734$

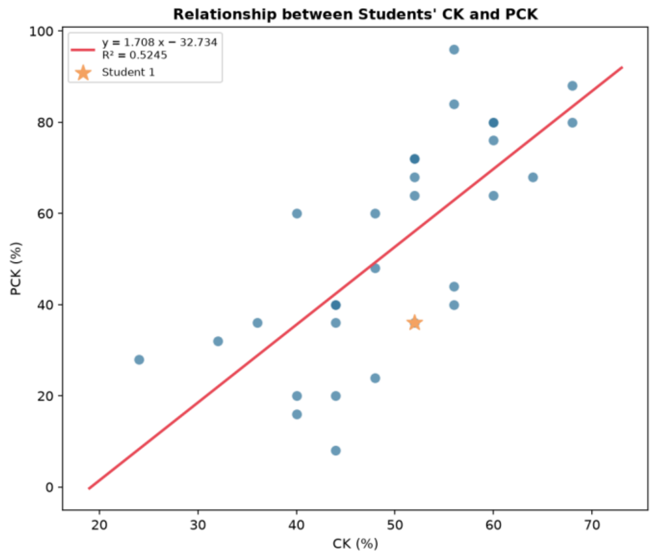


Figure 3. Scatter Plot and Regression Line of the CK–PCK Relationship

The equation PCK = 1.708 × CK − 32.734 with $R^2 = 0.52$ indicates that about 52% of the variance in PCK is statistically associated with CK. Because the design is correlational, this relationship should be interpreted as a predictive association rather than a causal effect; the remaining 48% of the variance signals that PCK does not form automatically from content mastery but requires an explicit pedagogical learning pathway.

The CBT also breaks aggregate achievement down into individual diagnostic profiles automatically. To illustrate this capability—rather than to draw general conclusions from a single case—Table 8 and Figure 4 summarize the profile of one participant (pseudonym Student 1), and Figures 5 and 6 show anonymized screenshots of the corresponding diagnostic report generated by the CBT. Such reports are produced automatically for every participant.

Table 8. Individual Achievement Profile of Student 1 Compared with the Class Average

Domain	Student 1 Score (%)	Class Average (%)	Difference	Category
Content Knowledge (CK)	52.00	50.00	+2.00	Low
Pedagogical Content Knowledge (PCK)	36.00	52.67	−16.67	Low

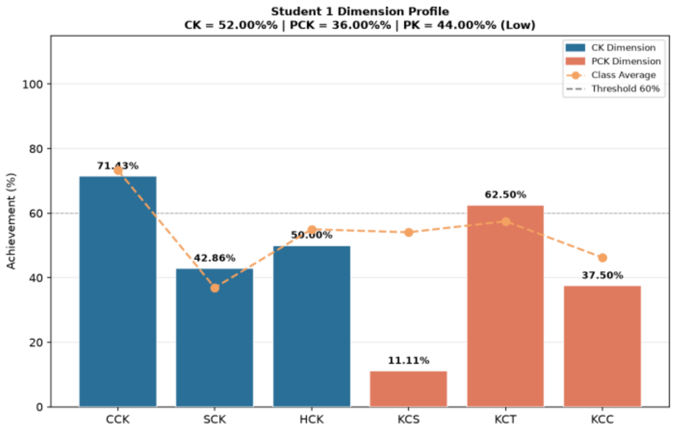


Figure 4. CK and PCK Achievement Profile of Student 1 Compared with the Class Average

As an illustrative example, Student 1 answered 13 of 25 CK items and 9 of 25 PCK items correctly, falling into the Low category for both domains, with CK slightly above and PCK well below the class average. Anonymized screenshots of the report are shown in Figures 5 and 6, with no identifiable student information displayed. This single case is used only to demonstrate the diagnostic output of the CBT and is not intended as a basis for general interpretation.

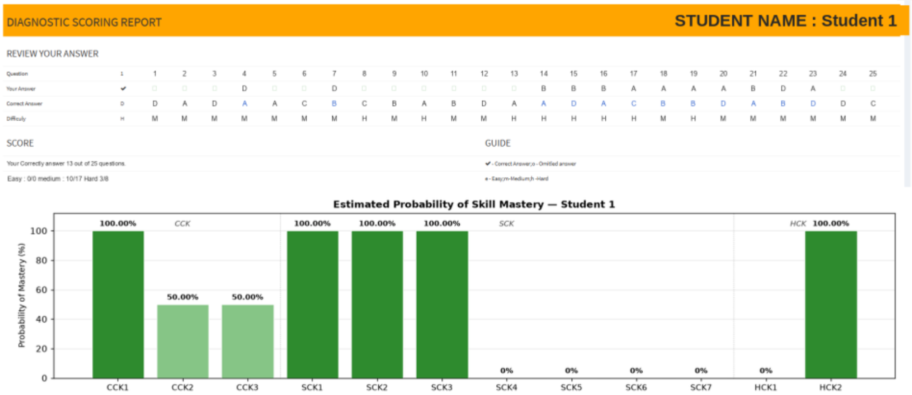


Figure 5. CBT Diagnostic Scorecard for the CK Domain for Student 1 (original system screenshot)

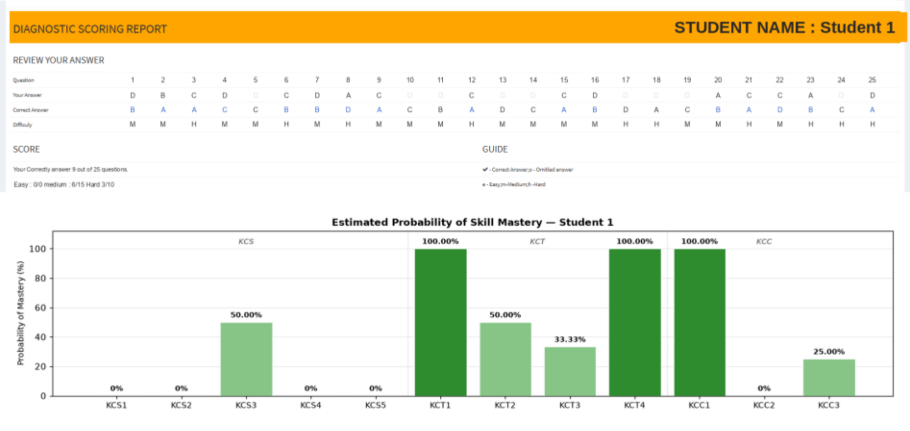


Figure 6. CBT Diagnostic Scorecard for the PCK Domain for Student 1 (system screenshot)

Figures 5 and 6 reveal diagnostic functions of the CBT that are more difficult to implement efficiently in conventional paper-based tests. In addition to recapping correct/incorrect answers per item along with their difficulty level—a combination of the Medium (M) and Hard (H) categories according to the item analysis of the instrument—the system presents an Estimated Probability of Skill Mastery graph that maps Student 1's mastery probability for each skill indicator (CCK1–HCK2 for CK and KCS1–KCC3 for PCK), complete with a mapping of sample items to each indicator. Thus, the CBT does not merely provide a final score but directly identifies which indicators have been mastered and which remain weak—in Student 1's case, several indicators show a mastery probability close to zero—so that lecturers can follow up with specific and immediate remediation. It is this automation of scoring, standardization of reporting, and diagnostic depth that affirms the added value of the CBT in assessing the professional knowledge of mathematics pre-service teachers.

At the indicator level, the same report pinpoints exactly which indicators a student has and has not mastered. For Student 1, for example, several higher-order indicators—advanced reasoning and proof within SCK, and knowledge of students within KCS—registered no mastery, whereas basic content indicators were fully mastered. This illustrates how indicator-level diagnosis can guide individualized remediation far more precisely than dimension averages alone, without implying that this individual profile generalizes to the cohort.

Summary of key findings. In sum, both CK (50.00%) and PCK (52.67%) fell below the 75% mastery threshold; 60.00% of the students were in the low category and none reached the high category; SCK (36.91%) and KCC (46.25%) were the weakest dimensions; CK and PCK were strongly and positively associated ($r = 0.72$; $p < 0.001$); and the CBT was able to decompose these results down to the indicator and individual-student levels, as illustrated by the profile of Student 1.

3.2. Discussion

Interpreted against the study's objectives, three points stand out. First, neither CK nor PCK reached the mastery threshold, and—contrary to the common assumption that content

Commented [AM2]: 1.Additional limitations should include:
•arbitrary mastery threshold;
•non-equated CK and PCK tests;
•very small numbers of items per indicator;
•possible unreliability of dimension scores;
•no classification-accuracy assessment;
•no usability or feasibility evaluation of the CBT;
•no comparison with paper-based testing;
•no evidence that reports led to intervention;
•possible semester-level confounding;
•common-method effects;
•multiple exploratory correlations.

mastery precedes and exceeds pedagogical knowledge—PCK was marginally higher than CK; this pattern resonates with Shulman's argument that transforming content into teachable form is a distinct challenge [7] and with the MKT framework's treatment of CK and PCK as related but non-identical constructs [6]. Second, the weakest areas lay in advanced specialized content within CK and in knowledge of curriculum within PCK, echoing the COACTIV model's claim that these facets critically shape the quality of instructional planning [2], [25]. Third, CK achievement was far more homogeneous than PCK, suggesting that pedagogical capacity develops more unevenly across pre-service teachers, consistent with reviews describing PCK as the most context-dependent and hardest-to-cultivate component [8], [9].

The pattern of correlations indicates that stronger mastery in particular content dimensions tends to co-occur with the corresponding pedagogical capacity, yet content mastery accounts for only part of the variation in PCK—consistent with evidence that pre-service teachers in developing contexts often show comparatively weaker PCK because curricula emphasize content coverage over pedagogical transformation [1], [10]. Rather than forming automatically from content knowledge, PCK appears to require explicit, dedicated learning experiences. The profile of Student 1 is consistent with this interpretation but is treated here only as an illustrative example, not as evidence in its own right.

These results are broadly consistent with prior Indonesian studies reporting that the professional knowledge of mathematics pre-service teachers is generally moderate and uneven, with PCK frequently emerging as the more fragile component [10], [11]. The finding that PCK was marginally higher than CK is noteworthy and can be interpreted cautiously: the PCK items, which draw on familiar classroom situations, may be more accessible to students than the CK items, which demand advanced and abstract mathematical reasoning; moreover, the far larger dispersion of PCK (SD = 23.96) indicates that this apparent advantage is unevenly distributed across the cohort. The particularly low SCK (36.91%) is plausibly linked to teacher-education curricula that emphasize procedural and elementary content over advanced mathematical reasoning—such as formal proof, mathematical modeling, and complex problem solving—so that the more cognitively demanding indicators (SCK4–SCK7) remain underdeveloped [24], [25].

Theoretically, the study offers a specific insight rather than a mere application of existing models: in a low-proficiency, underrepresented setting, CK and PCK remain strongly interdependent even though neither reaches mastery and PCK can slightly exceed CK. This qualifies the intuitive assumption that CK necessarily leads PCK and suggests that, at early proficiency levels, the two constructs may develop in tandem rather than strictly sequentially—extending the MKT and COACTIV frameworks to conditions from which they were not originally derived [6], [25], [27]. Methodologically, the study contributes a CBT model that jointly diagnoses CK and PCK down to the indicator level, addressing the scarcity of instruments that do so simultaneously [22], [28]. In policy terms, the finding that no student reached the high category provides empirical grounds for data-driven rather than assumption-driven curriculum reform [5], [23], a concern underscored by Indonesia's recent international assessment results [3].

Positioned this way, the CBT functions not merely as an examination tool but as a diagnostic-assessment ecosystem that yields rich, decomposable cognitive data actionable for program managers [23], [30]. Several limitations should nonetheless temper these conclusions and caution against generalizing beyond the single institution and the 30 participants studied here: (a) the small sample size, which limits statistical power and generalizability; (b) the total-purposive, single-institution sampling; (c) the bivariate scope, which does not model other determinants simultaneously; (d) the reliance solely on quantitative CBT data, without syllabus and document analysis, classroom observation, or interviews that could triangulate and explain why students underperformed; (e) the limited external evidence of the instrument's validity and reliability beyond the internal analyses reported here; and (f) the multiple-choice format, which cannot fully capture the reasoning that open-ended or performance-based tasks would reveal. In practical terms, these findings imply that teacher-education curricula should allocate explicit, assessed learning time to advanced specialized content (formal proof, mathematical modeling, and complex problem solving) and to knowledge of students and curriculum, and should embed diagnostic CBT within course cycles so that indicator-level weaknesses are identified early and addressed through targeted, evidence-based remediation. Future research should employ larger multi-institution and longitudinal samples, mixed-methods designs that add qualitative triangulation, and multivariate models to strengthen internal and external validity.

CONCLUSION AND RECOMMENDATIONS

This study set out to measure and profile the CK and PCK of mathematics pre-service teachers through a diagnostic CBT. On average, both constructs remained below the mastery threshold, with the majority of students in the low category and none in the high category; the weakest areas were advanced specialized content within CK and knowledge of curriculum within PCK, while CK and PCK were strongly and positively associated. Theoretically, the study clarifies that CK and PCK can remain interdependent—and PCK can slightly exceed CK—even at low overall proficiency; methodologically, it demonstrates a CBT capable of layered, indicator-level diagnosis of both constructs; and practically, it provides a standardized, efficient means of generating actionable competence profiles for teacher education. These conclusions should be read in light of the study's limitations, namely a small, single-institution, purposively selected sample analyzed with bivariate, multiple-choice-based methods. Overall, the central message is that diagnostic computer-based assessment can meaningfully advance the monitoring and development of mathematics pre-service teachers' professional knowledge, provided its profiles are used to inform evidence-based curriculum and instructional reform.

Recommendations. The following recommendations are offered as literature-informed, potentially effective alternatives suggested by—rather than directly proven by—the present findings, and they should be validated in future studies.

For lecturers. Remedial instruction may target not only the weakest dimensions (SCK within CK and KCC within PCK) but also the weakest indicators identified in Table 5. In the CK domain, priority may be given to strengthening mathematical proof—proving properties or theorems (SCK4) and applying proof strategies (SCK5)—through scaffolded proof practice, analysis of proof errors, and habituation to deductive argumentation, followed by mathematical modeling (SCK6) and complex problem solving (SCK7) via problem-based learning. In the PCK domain, the focus may fall on analyzing the strengths and weaknesses of representations (KCT4), applying standards and curricula (KCC2), and designing measurement instruments (KCC3) through case-based learning, analysis of teaching videos, and curriculum-document study.

For teacher education institution curriculum managers. The results encourage integrating a diagnostic CBT—capable of mapping achievement down to the indicator and individual levels—into the study program's quality-assurance cycle, and rebalancing the curriculum so that advanced mathematical reasoning and pedagogical-content competence receive as much emphasis as procedural content mastery.

For future researchers. Future studies may use larger, multi-institution, and longitudinal samples, mixed-methods designs that add qualitative triangulation (syllabus analysis, classroom observation, and interviews), and multivariate models involving additional determinants (such as academic background, learning motivation, and quality of teaching) to strengthen external validity and to explain the causes of the observed weaknesses. A further agenda is to develop and empirically test enhancements to the CBT itself—automatic item-level feedback with worked explanations, direct links to remedial materials for each weak indicator, and adaptive item delivery—so that, once validated, the system can function not only as a diagnostic tool but also as an integrated learning-support environment.

ACKNOWLEDGMENTS

The authors thank the Mathematics Education study program of the teacher education institution for facilitating the implementation of the research, as well as all the students who participated as respondents.

REFERENCES

[1] X. Yang and G. Kaiser, "The impact of mathematics teachers' professional competence on instructional quality and students' mathematics learning outcomes," *Curr. Opin. Behav. Sci.*, vol. 48, p. 101225, Dec. 2022, doi: 10.1016/j.cobeha.2022.101225.

[2] S. Krauss *et al.*, "Pedagogical content knowledge and content knowledge of secondary mathematics teachers.," *J. Educ. Psychol.*, vol. 100, no. 3, pp. 716–725, Aug. 2008, doi: 10.1037/0022-0663.100.3.716.

[3] OECD, *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*. in PISA. OECD Publishing, 2023. doi: 10.1787/53f23881-en.

[4] International Association for the Evaluation of Educational Achievement, *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. Amsterdam: IEA, 2012.

[5] H. Qian and P. Youngs, "The effect of teacher education programs on future elementary mathematics teachers' knowledge: a five-country analysis using TEDS-M data," *J. Math. Teach. Educ.*, vol. 19, no. 4, pp. 371–396, Aug. 2016, doi: 10.1007/s10857-014-9297-0.

[6] D. Loewenberg Ball, M. H. Thames, and G. Phelps, "Content Knowledge for Teaching: What Makes It Special?," *J. Teach. Educ.*, vol. 59, no. 5, pp. 389–407, Nov. 2008, doi: 10.1177/0022487108324554.

[7] L. S. Shulman, "Those Who Understand: Knowledge Growth in Teaching," *Educ. Res.*, vol. 15, no. 2, pp. 4–14, Feb. 1986, doi: 10.3102/0013189X015002004.

[8] M. Grigaliūnienė, E. Lehtinen, L. Verschaffel, and F. Depaepe, "Systematic Review of Research on Pedagogical Content Knowledge in Mathematics: Insights from a Topic-Specific Approach," *ZDM – Math. Educ.*, vol. 57, no. 4, pp. 777–794, Aug. 2025, doi: 10.1007/s11858-025-01684-1.

[9] F. Depaepe, L. Verschaffel, and G. Kelchtermans, "Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research," *Teach. Teach. Educ.*, vol. 34, pp. 12–25, Aug. 2013, doi: 10.1016/j.tate.2013.03.001.

[10] S. Y. Ningsih, Turmudi, and D. Juandi, "Pedagogical content knowledge (PCK) profile of prospective teachers in mathematics learning," *J. Phys. Conf. Ser.*, vol. 1521, no. 3, p. 032057, Apr. 2020, doi: 10.1088/1742-6596/1521/3/032057.

[11] S. Fiangga, S. Khabibah, S. M. Amin, and R. Ekawati, "A Learning Design Analysis of the Pre-service Teachers' Mathematics Pedagogical Content Knowledge," *J. Phys. Conf. Ser.*, vol. 1899, no. 1, p. 012153, May 2021, doi: 10.1088/1742-6596/1899/1/012153.

[12] X. Yang, B. Schwarz, and I. K. C. Leung, "Pre-service mathematics teachers' professional modeling competencies: a comparative study between Germany, Mainland China, and Hong Kong," *Educ. Stud. Math.*, vol. 109, no. 2, pp. 409–429, Feb. 2022, doi: 10.1007/s10649-021-10064-x.

[13] J. Njiku, "Pre-service teachers manifested mathematics pedagogical content knowledge: The role of the teaching practicum," *Pedagog. Res.*, vol. 10, no. 1, p. em0229, Jan. 2025, doi: 10.29333/pr/15647.

[14] A. J. Boevé, R. R. Meijer, C. J. Albers, Y. Beetsma, and R. J. Bosker, "Introducing Computer-Based Testing in High-Stakes Exams in Higher Education: Results of a Field Experiment," *PLOS ONE*, vol. 10, no. 12, p. e0143616, Dec. 2015, doi: 10.1371/journal.pone.0143616.

[15] C. Y. Piaw, "Replacing Paper-based Testing with Computer-based Testing in Assessment: Are we Doing Wrong?," *Procedia - Soc. Behav. Sci.*, vol. 64, pp. 655–664, Nov. 2012, doi: 10.1016/j.sbspro.2012.11.077.

[16] A. Berry, P. Friedrichsen, and J. Loughran, Eds., *Re-examining pedagogical content knowledge in science education*. in Teaching and learning in science series. London New York: Routledge, 2015. doi: 10.4324/9781315735665.

[17] M. Schmid, E. Brianza, and D. Petko, "Self-reported technological pedagogical content knowledge (TPACK) of pre-service teachers in relation to digital technology use in lesson plans," *Comput. Hum. Behav.*, vol. 115, p. 106586, Feb. 2021, doi: 10.1016/j.chb.2020.106586.

[18] L. Maas, M. J. S. Brinkhuis, L. Kester, and L. Wijngaards-de Meij, "Cognitive Diagnostic Assessment in University Statistics Education: Valid and Reliable Skill Measurement for Actionable Feedback Using Learning Dashboards," *Appl. Sci.*, vol. 12, no. 10, p. 4809, May 2022, doi: 10.3390/app12104809.

[19] A. Barana, M. Marchisio, and M. Sacchet, "Interactive Feedback for Learning Mathematics in a Digital Learning Environment," *Educ. Sci.*, vol. 11, no. 6, p. 279, Jun. 2021, doi: 10.3390/educsci11060279.

[20] S. Rezat, "How automated feedback from a digital mathematics textbook affects primary students' conceptual development: two case studies," *ZDM – Math. Educ.*, vol. 53, no. 6, pp. 1433–1445, Nov. 2021, doi: 10.1007/s11858-021-01263-0.

[21] L. N. Yaqin, L. D. Prasojo, N. A. Haji-Othman, N. Yusof, and A. Habibi, "Addressing the Digital Divide in Indonesian Higher Education: Insights, Implications, and Potential Solutions," in *From Digital Divide to Digital Inclusion*, Ł. Tomczyk, F. D. Guillén-Gámez, J. Ruiz-Palmero, and A. Habibi, Eds., in Lecture Notes in Educational Technology. , Singapore: Springer Nature Singapore, 2023, pp. 291–307. doi: 10.1007/978-981-99-7645-4_13.

[22] Naila Rajiha and S. Rini, "COMPUTER-BASED STANDARDIZED TESTING (CBT) SUMMATIVE ASSESSMENT: ONE INSTITUTION EXPERIENCE," *JISAE J. Indones. Stud. Assess. Eval.*, vol. 9, no. 1, pp. 1–7, Jan. 2023, doi: 10.21009/jisae.v9i1.33233.

[23] G. Kaiser and J. König, "Competence Measurement in (Mathematics) Teacher Education and Beyond: Implications for Policy," *High. Educ. Policy*, vol. 32, no. 4, pp. 597–615, Dec. 2019, doi: 10.1057/s41307-019-00139-z.

[24] T. Kleickmann *et al.*, "Teachers' Content Knowledge and Pedagogical Content Knowledge: The Role of Structural Differences in Teacher Education," *J. Teach. Educ.*, vol. 64, no. 1, pp. 90–106, Jan. 2013, doi: 10.1177/0022487112460398.

[25] J. Baumert and M. Kunter, "The COACTIV Model of Teachers' Professional Competence," in *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers*, M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, and M. Neubrand, Eds., Boston, MA: Springer US, 2013, pp. 25–48. doi: 10.1007/978-1-4614-5149-5_2.

[26] B. Rowan, R. Correnti, and R. J. Miller, "What Large-Scale, Survey Research Tells Us about Teacher Effects on Student Achievement: Insights from the Prospects Study of Elementary Schools," *Teach. Coll. Rec. Voice Scholarsh. Educ.*, vol. 104, no. 8, pp. 1525–1567, Aug. 2002, doi: 10.1111/1467-9620.00212.

[27] S. Blömeke, J.-E. Gustafsson, and R. J. Shavelson, "Beyond Dichotomies: Competence Viewed as a Continuum," *Z. Für Psychol.*, vol. 223, no. 1, pp. 3–13, Jan. 2015, doi: 10.1027/2151-2604/a000194.

[28] E. Mohammadpour and Y. Maroofi, "A performance-based test to measure teachers' mathematics and science content and pedagogical knowledge," *Heliyon*, vol. 9, no. 3, p. e13932, Mar. 2023, doi: 10.1016/j.heliyon.2023.e13932.

[29] T. Clark and H. Endres, "Computer-based diagnostic assessment of high school students' grammar skills with automated feedback – an international trial," *Assess. Educ. Princ. Policy Pract.*, vol. 28, no. 5–6, pp. 602–632, Nov. 2021, doi: 10.1080/0969594X.2021.1970513.

[30] G. A. Nortvedt and N. Buchholtz, "Assessment in mathematics education: responding to issues regarding methodology, policy, and equity," *ZDM*, vol. 50, no. 4, pp. 555–570, Aug. 2018, doi: 10.1007/s11858-018-0963-z.

Bukti Revisi ke-3

Journal of Mathematics Instruction, Social Research and Opinion
Vol. xx, No. x, March/June/September/December 202x, pp. 1-1x, https://doi.org/
ISSN 2962-7842

Assessing Mathematics Pre-Service Teachers' Content Knowledge and Pedagogical Content Knowledge through a Diagnostic Computer-Based Test

Article Info

Article history:

Received mm dd, yyyy
Revised mm dd, yyyy
Accepted mm dd, yyyy

Keywords:

Computer-based test
Content knowledge
Diagnostic assessment
Mathematical knowledge for teaching
Mathematics pre-service teachers
Pedagogical content knowledge

ABSTRACT

Content Knowledge (CK) and Pedagogical Content Knowledge (PCK) are core components of mathematics teachers' competence, yet they are rarely measured simultaneously through technology-based diagnostic assessment. This study measures and profiles the CK and PCK of mathematics pre-service teachers using a diagnostic Computer-Based Test (CBT). A quantitative descriptive-correlational design was applied to 30 purposively selected pre-service teachers. Each construct was measured with 25 multiple-choice items whose content validity was established by expert judgment and whose reliability was confirmed before use (KR-20 = 0.81 for CK; 0.79 for PCK). Data were analyzed with descriptive statistics, Pearson correlation, and simple linear regression after assumption checks. Mean CK (50.00%) and PCK (52.67%) were both below the 75% mastery threshold, with 60% of students in the low category; specialized content knowledge and knowledge of content and curriculum were the weakest dimensions. CK and PCK correlated strongly and positively ($r = 0.72$; $p < 0.001$), with CK accounting for about 52% of the variance in PCK as a predictive, non-causal association. The CBT generated actionable diagnostic profiles at the dimension and indicator levels. These findings support using diagnostic CBT to guide data-driven curriculum reform integrating content mastery and pedagogical capacity.

Corresponding Author: Reza Oktiana Akbar

INTRODUCTION

The quality of a nation's mathematics education depends heavily on teachers' professional competence. Cross-national studies have shown that mathematics teachers' professional knowledge, comprising Content Knowledge (CK) and Pedagogical Content Knowledge (PCK) is a primary predictor of instructional effectiveness and students' academic achievement [1], [2]. The 2022 PISA results placed Indonesian students' mathematics literacy 68th out of 81 countries, below the OECD average, indicating a fundamental problem in the quality of mathematics instruction [3]. This condition is inseparable from the quality of professional knowledge formed in teacher education

institutions, so ensuring an adequate professional-knowledge profile of pre-service teachers becomes a strategic imperative [4], [5].

Theoretically, the Mathematical Knowledge for Teaching (MKT) framework distinguishes CK, comprising common, specialized, and horizon content knowledge from PCK, which encompasses knowledge of content and students, content and teaching, and content and curriculum [6], [7]. Systematic reviews affirm that PCK is the most frequently studied yet the most difficult component to measure consistently because of its contextual and multidimensional nature [8], [9]. Studies in Indonesia show that the professional knowledge of mathematics pre-service teachers varies widely across institutions and curricula, with a notable weakness in the PCK aspect [10], [11]. Similar findings have been reported in cross-national contexts [12], [13].

Alongside the rapid digitalization of education, competency-evaluation systems are now shifting massively from traditional approaches toward technology-based assessment. Various studies indicate that CBT is often more valid and reliable than conventional examination methods [14], [15]; however, its use to objectively measure both the content knowledge and the pedagogical content knowledge of pre-service teachers simultaneously remains very limited because of the complexity of these constructs [16], [17]. The potential of CBT to generate diagnostic feedback based on learning analytics has also not been optimally exploited [18], even though interactive digital feedback has been shown to improve mathematical understanding [19], [20].

The urgency of this study is heightened in the context of Indonesia's uneven higher-education digital transformation [21], which is not yet supported by adequate standardization of assessment systems [22]. Measuring teacher competence in the digital era is required to produce data that are actionable for study-program managers and policymakers [23]. A number of studies affirm that CK and PCK are relatively independent constructs that require different intervention pathways [24], [25], and that teacher effects on student achievement are substantial [26]. Taken together, however, this body of work remains fragmented: CK and PCK are rarely measured simultaneously with a single validated instrument, the diagnostic potential of CBT is seldom exploited at the indicator level, and the evidence is concentrated in a few well-studied national settings, leaving a clear gap for integrated, technology-based diagnosis of both constructs.

Building on this gap, the present study makes three specific contributions rather than a general claim to novelty. Theoretically, it provides empirical evidence on how CK and PCK co-vary at low overall proficiency, showing that their interdependence persists even when neither reaches mastery, an insight that refines, rather than merely applies, the MKT and COACTIV frameworks [6], [25], [27]. Methodologically, it operationalizes a CBT that produces layered, indicator-level diagnostic profiles for both constructs simultaneously, addressing the scarcity of validated instruments that jointly diagnose CK and PCK [28], [29]. Practically, it demonstrates how such diagnostic data can be translated into targeted curriculum and instructional decisions. By drawing on an underrepresented context, the study also contributes comparative evidence to the international literature on teacher knowledge rather than a purely local application of CBT [12], [30]. Accordingly, the study aims to: (1) measure and profile pre-service teachers' CK and PCK; (2) identify strengths and weaknesses at the dimension and indicator levels; and (3) demonstrate how diagnostic profiles can inform instructional intervention.

METHOD

This study employed a quantitative approach with a descriptive-survey and correlational design. The population comprised students of the Mathematics Education study program at a teacher education institution who were in their fifth and seventh semesters. A sample of 30 students was determined through purposive sampling with the inclusion criteria: (1) active students who had adequately completed mathematics-content and pedagogy courses; and (2) willing to take the entire series of CBT tests. Given the limited sample size, all data were analyzed as a single intact group (n = 30) and were not directed toward inter-semester comparison.

Research setting and participants. The research was conducted in the odd semester of the 2024/2025 academic year in the Mathematics Education study program of a state Islamic teacher education institution in Indonesia. Participants were selected using total (saturated) purposive sampling: all 30 students who met the inclusion criterion, having completed the core mathematics-content and pedagogy courses were included, so that the sample comprised the entire eligible cohort rather than a probability sample. The sample size (n = 30) therefore reflects the actual number of eligible students at the single participating institution and suits the study's diagnostic, exploratory aim of profiling CK and PCK in depth; it is nonetheless acknowledged as a constraint on statistical power and on the generalizability of the correlational and regression results, as elaborated in the Discussion.

The operationalization of the variables comprised two main constructs. CK was defined as mastery of the mathematics content to be taught, measured through 25 multiple-choice items; PCK was defined as the ability to integrate content knowledge with pedagogical strategies, also measured through 25 multiple-choice items. Scores were converted into percentages of achievement relative to the ideal maximum score. The CBT instrument was designed to generate layered diagnostic reports, namely the total score, achievement per dimension, achievement per indicator, and an individual achievement profile for each student. Each item was mapped to a specific indicator and dimension (see the notes below Table 1) so that the CBT system could produce achievement per dimension and per indicator for all students at once, as well as their individual diagnostic profiles. The composition of the instrument is presented in Table 1. The multiple-choice format was chosen to enable fully automated, standardized, and scalable scoring across all participants and indicators at once, a prerequisite for the layered diagnostic reporting; its rationale and limitations are elaborated under ‘Item format’.

Table 1. Composition of CBT Instrument Items by MKT Dimension

Domain	Dimension	Number of Items
CK	Common Content Knowledge (CCK)	7
CK	Specialized Content Knowledge (SCK)	14
CK	Horizon Content Knowledge (HCK)	4
PCK	Knowledge of Content and Students (KCS)	10
PCK	Knowledge of Content and Teaching (KCT)	9
PCK	Knowledge of Content and Curriculum (KCC)	6
Total		50

Notes on dimension codes and indicator coverage (based on the developed CK and PCK instruments). **CK – CCK** (Common Content Knowledge, items 1–7): understanding basic concepts (e.g., lines and angles), performing calculations and measurements accurately, and applying mathematical concepts in everyday life. **CK – SCK** (Specialized Content Knowledge, items 8–21): explaining concepts and principles, analyzing and connecting mathematical objects, proving properties, building models, and solving complex mathematical problems. **CK – HCK** (Horizon Content Knowledge, items 22–25): connecting school mathematics with advanced academic mathematics and with other disciplines. **PCK – KCS** (Knowledge of Content and Students, items 1–10): analyzing students' preconceptions and understanding as well as predicting and analyzing students' misconceptions, errors, and difficulties. **PCK – KCT** (Knowledge of Content and Teaching, items 11–19): selecting and organizing learning materials, selecting sources, strategies, and techniques, and selecting and analyzing representations for presenting material. **PCK – KCC** (Knowledge of Content and Curriculum, items 20–25): designing content and topic combinations, applying standards and curricula, and designing instruments to measure students' abilities.

Instrument development and content validity. The CBT instrument was developed through a systematic procedure. First, a test blueprint was constructed by mapping each item to a specific MKT dimension and indicator (Table 1). Second, the draft items were reviewed by three experts—two mathematics-education lecturers and one educational-assessment specialist—to establish content validity; the level of expert agreement was quantified using Aiken's V, and items with $V \geq 0.80$ were retained while the remaining items were revised or replaced. Third, the revised items were trialed and analyzed empirically before being administered in the present study. To illustrate the item format, each CK item presented a mathematical task with four to five options, whereas each PCK item was framed as a short classroom scenario or a sample of student work followed by options representing alternative pedagogical decisions. The complete test blueprint and sample items are available from the authors upon reasonable request.

Reliability and item analysis. Based on the item analysis, the instrument demonstrated adequate psychometric quality. Internal-consistency reliability computed with the Kuder–Richardson formula (KR-20) reached 0.81 for the CK test and 0.79 for the PCK test, both within the acceptable-to-good range. The item difficulty indices ranged from 0.20 to 0.87 (a balanced mix of easy, medium, and hard items), and the discrimination indices were largely satisfactory ($D \geq 0.30$); only items meeting the validity, difficulty, and discrimination criteria were retained in the final CBT.

Item format. Multiple-choice items were chosen because the CBT was designed for automated, standardized, and scalable diagnosis across many participants and indicators simultaneously. To capture pedagogical reasoning rather than mere recall, the PCK items were built around short classroom scenarios, samples of student work, and instructional-decision situations, with distractors constructed from typical student misconceptions and common teaching errors. It is nevertheless acknowledged that the multiple-choice format cannot fully capture the depth of pedagogical reasoning that open-ended or performance-based tasks would reveal, a limitation revisited in the Discussion.

The analysis techniques comprised descriptive and inferential analyses. Descriptive analysis produced the mean (M), standard deviation (SD), minimum, maximum, median, and percentage of achievement at the total, dimension, indicator, and individual levels, as well as a three-level categorization (High, Medium, Low). The categorization used the following cut-off points: High (≥ 75%), Medium (60–74.99%), and Low (< 60%), where the 75% mark represents the ideal mastery criterion adopted in this study. Inferential analysis was conducted through two techniques: (1) Pearson correlation to test the CK–PCK relationship and the corresponding inter-dimension relationships (CCK–KCS, SCK–KCT, HCK–KCC) at $\alpha = 0.05$; and (2) simple linear regression with CK as the predictor and PCK as the criterion. Prior to the inferential tests, the underlying assumptions were examined, namely the normality of the CK and PCK distributions (Shapiro–Wilk), the linearity of their relationship, the absence of influential outliers, and homoscedasticity; all assumptions were adequately met, so that the parametric analyses were appropriate.

All scoring, achievement computation, correlation, and linear regression were performed using Microsoft Excel through standardized formulas (including CORREL, SLOPE, INTERCEPT, and RSQ) so that every figure can be traced back to the raw data. The study also observed standard research-ethics principles: participation was entirely voluntary, informed consent was obtained from all respondents before data collection, and the participants' anonymity and confidentiality were maintained throughout the recording, processing, and reporting of data by using codes instead of names, with the individual-profile example presented under the pseudonym Student 1.

RESULTS AND DISCUSSION

3.1. Results

This section reports the findings of the data analysis in the following order: respondent characteristics, descriptive statistics, the distribution of achievement categories, dimension- and indicator-level achievement, the correlation and regression results, and an individual diagnostic profile. The 30 respondents were fifth- and seventh-semester students of the Mathematics Education study program who had completed the core mathematics-content and pedagogy courses; all of them completed the full series of CK and PCK tests through the CBT system.

Overall, students' professional-knowledge achievement had not yet reached the ideal competency threshold (≥ 75%). The mean CK score was 50.00% (SD = 10.16), PCK 52.67% (SD = 23.96), and the combined Professional Knowledge (PK) score 51.33% (SD = 16.04). The complete descriptive statistics are presented in Table 2.

Table 2. Descriptive Statistics of CK, PCK, and Professional Knowledge Scores (n = 30)

Variable	Mean (%)	SD	Min	Max	Median
Content Knowledge (CK)	50.00	10.16	24	68	52.00
Pedagogical Content Knowledge (PCK)	52.67	23.96	8	96	54.00
Professional Knowledge (PK)	51.33	16.04	26	78	50.00

Table 2 shows that the mean PCK is slightly higher than CK (a difference of 2.67 points), an interesting pattern indicating that pedagogical-transformation capacity slightly exceeds content mastery. The large standard deviation of PCK (SD = 23.96) and the much smaller one for CK (SD = 10.16) reflect substantial heterogeneity in knowledge profiles, with score ranges of 24–68% for CK and 8–96% for PCK. The distribution of Professional Knowledge categories is presented in Table 3.

Table 3. Distribution of Professional Knowledge Categories (n = 30)

Category	Frequency (f)	Percentage (%)
High	0	0.00

Medium	12	40.00
Low	18	60.00
Total	30	100.00

All students (100%) were in the Low or Medium category, with 60.00% in the Low category and 40.00% in the Medium category, and no student reaching the High category. This indicates that the majority of pre-service teachers have not attained an optimal level of professional-knowledge mastery. Achievement per dimension is presented in Table 4 and Figure 1.

Table 4. Average Achievement per CK and PCK Dimension

Domain	Dimension	Items	Mean (%)	SD
CK	CCK	7	73.33	20.47
CK	SCK	14	36.91	24.16
CK	HCK	4	55.00	35.60
PCK	KCS	10	54.07	33.18
PCK	KCT	9	57.50	26.14
PCK	KCC	6	46.25	32.30

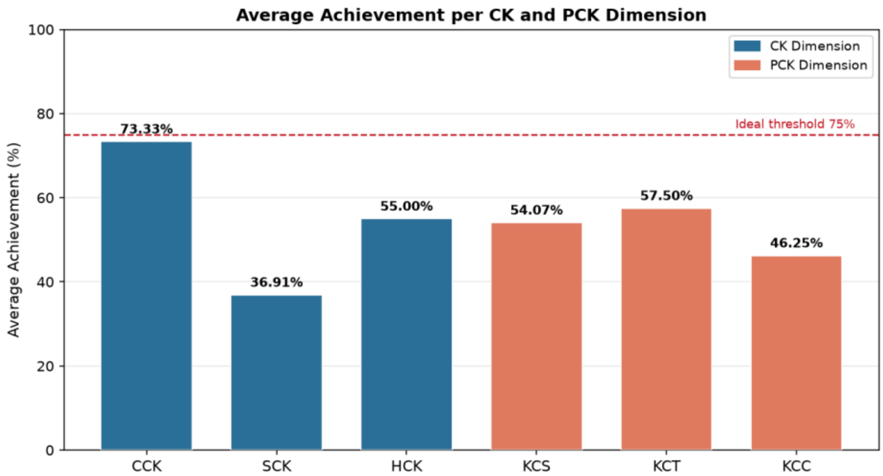


Figure 1. Average Achievement per CK and PCK Dimension

In the CK domain, the CCK dimension was the highest (73.33%) and SCK the lowest (36.91%); in the PCK domain, KCT was the highest (57.50%) while KCC was the lowest of all six dimensions (46.25%). Diagnostically, the low SCK and KCC signal a weak ability to relate content to the structure and organization of the curriculum—an area that needs to be prioritized in intervention. Achievement at the indicator level is presented in Table 5 and visualized in Figure 2.

Table 5. Average Achievement per CK and PCK Indicator

Domain	Indicator	Mean (%)
CK	CCK1	86.67
CK	CCK2	80.00
CK	CCK3	46.67
CK	SCK1	63.33
CK	SCK2	46.67
CK	SCK3	43.33
CK	SCK4	20.00
CK	SCK5	23.33
CK	SCK6	30.00

CK	SCK7	33.33
CK	HCK1	53.33
CK	HCK2	56.67
PCK	KCS1	58.33
PCK	KCS2	48.33
PCK	KCS3	58.33
PCK	KCS4	51.67
PCK	KCS5	53.33
PCK	KCT1	55.00
PCK	KCT2	66.67
PCK	KCT3	60.00
PCK	KCT4	36.67
PCK	KCC1	51.67
PCK	KCC2	43.33
PCK	KCC3	45.00

Notes on indicator codes (based on the CK and PCK instruments). CK indicators — **CCK1**: understanding basic school-mathematics concepts, such as lines, angles, sequences, and basic statistics (items 1–3); **CCK2**: performing calculations and measurements accurately (items 4–5); **CCK3**: applying mathematical concepts in everyday life (items 6–7); **SCK1**: explaining academic mathematical concepts and principles (items 8–9); **SCK2**: analyzing mathematical objects (items 10–11); **SCK3**: connecting mathematical properties and objects (items 12–13); **SCK4**: proving mathematical properties or theorems (items 14–16); **SCK5**: applying proof strategies (item 17); **SCK6**: building and constructing mathematical models (items 18–19); **SCK7**: solving complex mathematical problems (items 20–21); **HCK1**: connecting school mathematics with advanced academic mathematics (items 22–23); **HCK2**: connecting mathematics with other disciplines (items 24–25). PCK indicators — **KCS1**: analyzing students' preconceptions (items 1–2); **KCS2**: identifying students' level of understanding (items 3–4); **KCS3**: predicting students' misconceptions (items 5–6); **KCS4**: identifying sources of students' learning difficulties (items 7–8); **KCS5**: analyzing students' literacy issues and errors (items 9–10); **KCT1**: organizing and sequencing learning materials (items 11–12); **KCT2**: selecting learning sources, strategies, and techniques (items 13–14); **KCT3**: selecting representations and illustrations for presenting material (items 15–17); **KCT4**: analyzing the weaknesses and strengths of representations (items 18–19); **KCC1**: composing content and topic combinations (items 20–21); **KCC2**: applying standards and curricula (items 22–23); **KCC3**: designing instruments to measure students' abilities (items 24–25).

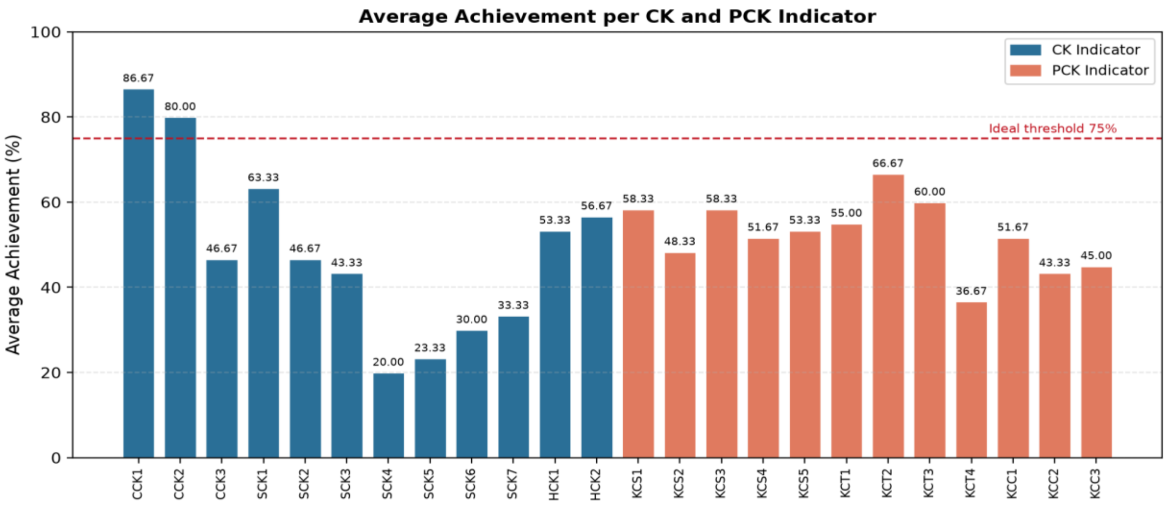


Figure 2. Average Achievement per CK and PCK Indicator

Figure 2 visualizes the average achievement of each indicator exactly as listed in Table 5. The range of achievement across indicators spans 20.00% to 86.67%. The highest achievement is on the CK indicator CCK1 (86.67%; understanding basic school-mathematics concepts), which is fundamental and therefore the most widely mastered by students. Conversely, the lowest overall achievement is on indicator SCK4 (20.00%; proving mathematical properties or theorems), the skill with the highest cognitive demand in the CK domain, followed by SCK5 (23.33%; applying proof strategies) and SCK6 (30.00%; building and constructing mathematical models). The pattern in the SCK dimension declines gradually with complexity, from explaining concepts and principles (SCK1; 63.33%), analyzing and connecting mathematical objects (SCK2 46.67%; SCK3 43.33%), solving complex problems (SCK7; 33.33%), to the most difficult modeling and proving (SCK6 30.00%; SCK5 23.33%; SCK4 20.00%), a pattern consistent with cognitive load theory. Because each dimension's achievement remains the mean of its constituent indicators, for example, KCS (54.07%) is the mean of KCS1–KCS5 and CCK (73.33%) the mean of CCK1–CCK3—this pattern maps precisely which items and indicators are the sources of weakness, so that lecturers can design remediation targeted at concrete rather than generic weaknesses. The results of the Pearson correlation test are presented in Table 6.

Table 6. Pearson Correlation between Domains and between Parallel Dimensions

Relationship	r	p	Interpretation
CK – PCK	0.72	< 0.001	Strong
CCK – KCS	0.42	< 0.05	Moderate
SCK – KCT	0.54	< 0.01	Moderate
HCK – KCC	0.86	< 0.001	Very strong

The strong positive correlation between CK and PCK ($r = 0.72$; $p < 0.001$) confirms that the two domains are constructively interdependent, reinforced by the correlations between parallel dimensions, especially HCK–KCC ($r = 0.86$) and SCK–KCT ($r = 0.54$). The predictive relationship of CK on PCK was tested through simple linear regression (Table 7 and Figure 3).

Table 7. Results of the Simple Linear Regression of CK on PCK

Parameter	Value
Slope (b)	1.708
Intercept (a)	−32.734
Coefficient of determination (R^2)	0.5245
Pearson r	0.72
Equation	PCK = 1.708 × CK − 32.734

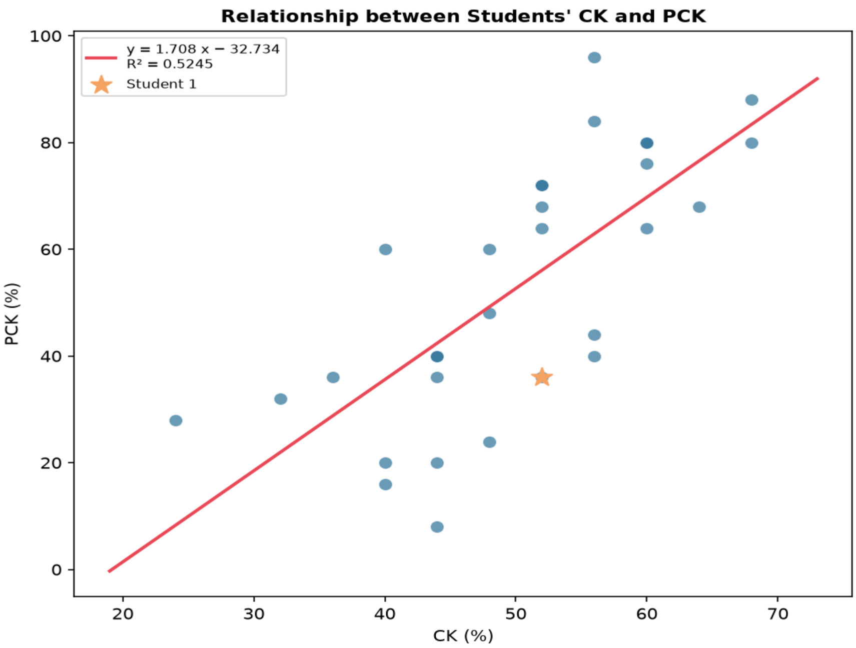


Figure 3. Scatter Plot and Regression Line of the CK–PCK Relationship

The equation PCK = 1.708 × CK − 32.734 with $R^2 = 0.52$ indicates that about 52% of the variance in PCK is statistically associated with CK. Because the design is correlational, this relationship should be interpreted as a predictive association rather than a causal effect; the remaining 48% of the variance signals that PCK does not form automatically from content mastery but requires an explicit pedagogical learning pathway.

The CBT also breaks aggregate achievement down into individual diagnostic profiles automatically. To illustrate this capability, rather than to draw general conclusions from a single case, Table 8 and Figure 4 summarize the profile of one participant (pseudonym Student 1), and Figures 5 and 6 show anonymized screenshots of the corresponding diagnostic report generated by the CBT. Such reports are produced automatically for every participant.

Table 8. Individual Achievement Profile of Student 1 Compared with the Class Average

Domain	Student 1 Score (%)	Class Average (%)	Difference	Category
Content Knowledge (CK)	52.00	50.00	+2.00	Low
Pedagogical Content Knowledge (PCK)	36.00	52.67	−16.67	Low

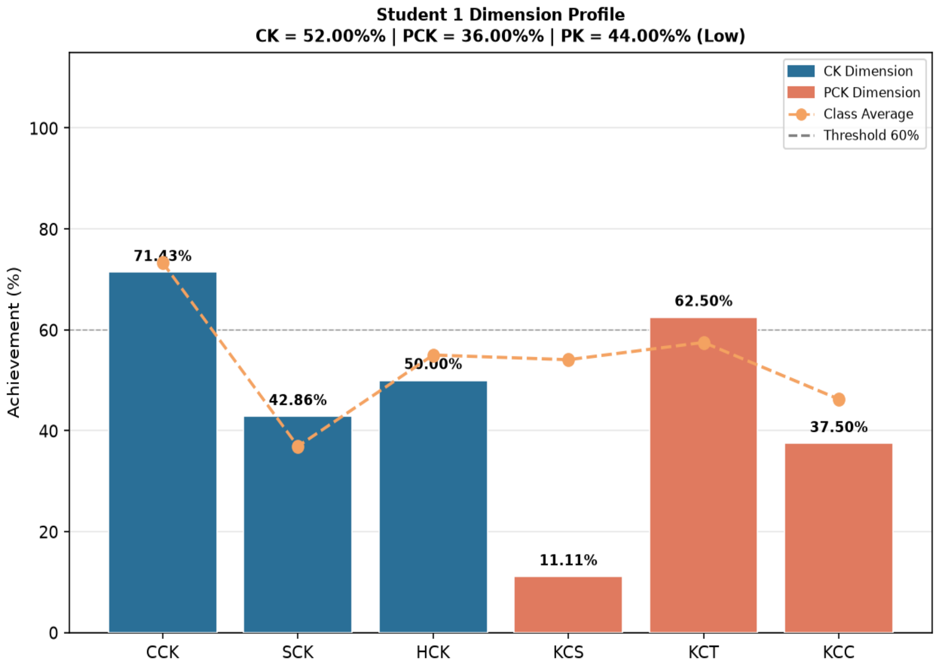


Figure 4. CK and PCK Achievement Profile of Student 1 Compared with the Class Average

As an illustrative example, Student 1 answered 13 of 25 CK items and 9 of 25 PCK items correctly, falling into the Low category for both domains, with CK slightly above and PCK well below the class average. Anonymized screenshots of the report are shown in Figures 5 and 6, with no identifiable student information displayed. This single case is used only to demonstrate the diagnostic output of the CBT and is not intended as a basis for general interpretation.

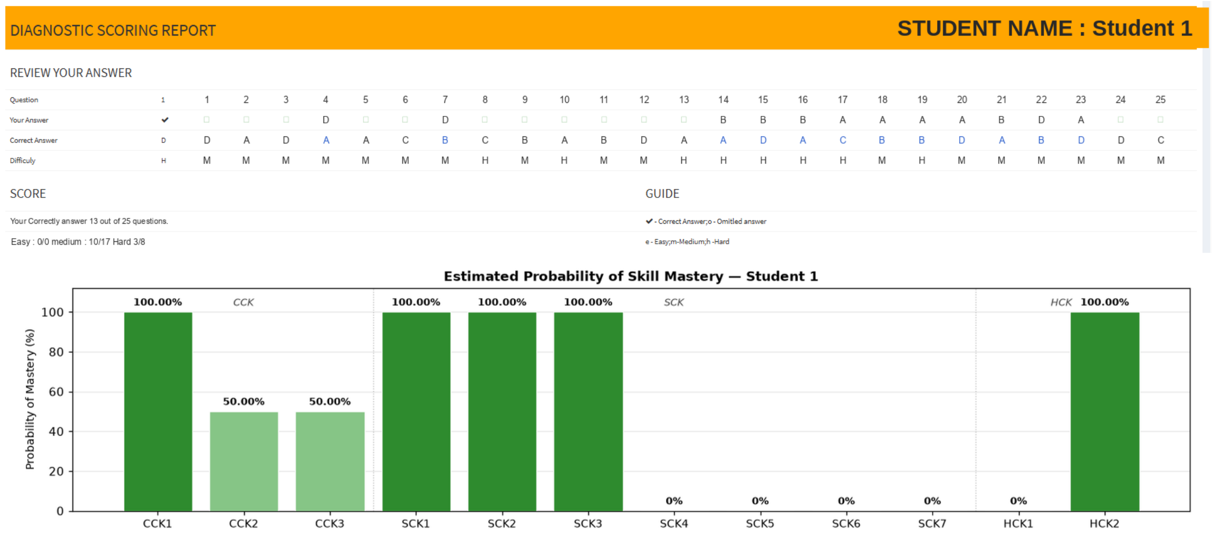


Figure 5. CBT Diagnostic Scorecard for the CK Domain for Student 1 (original system screenshot)

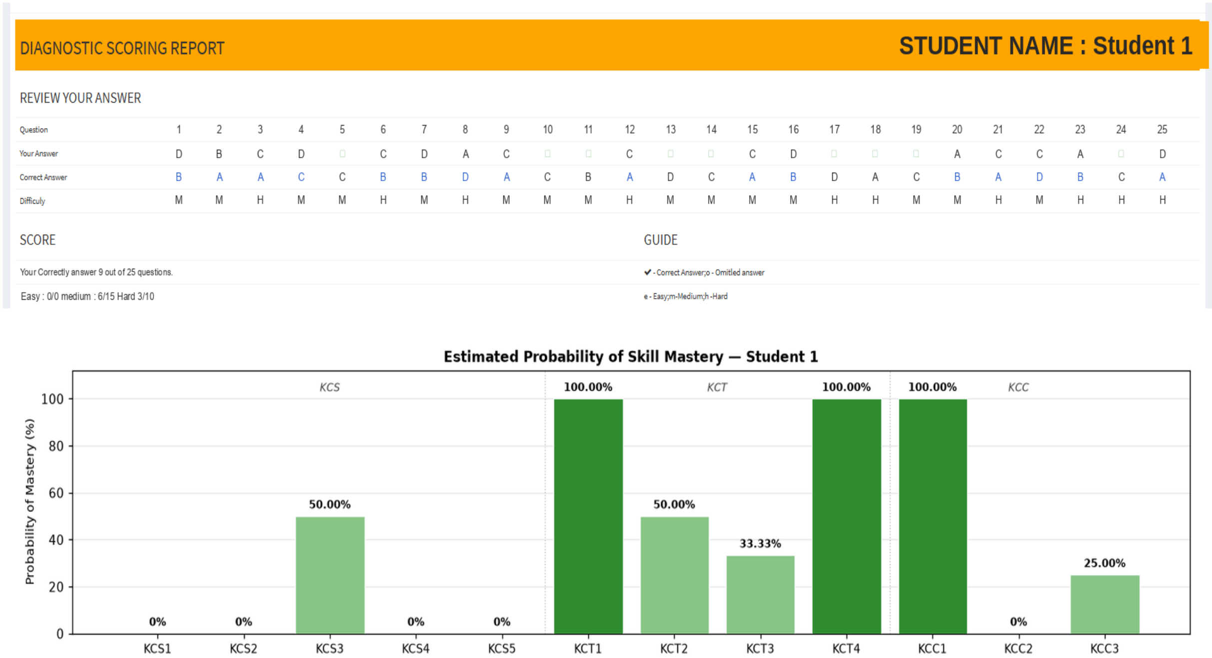


Figure 6. CBT Diagnostic Scorecard for the PCK Domain for Student 1 (system screenshot)

Figures 5 and 6 reveal diagnostic functions of the CBT that are more difficult to implement efficiently in conventional paper-based tests. In addition to recapping correct/incorrect answers per item along with their difficulty level, a combination of the Medium (M) and Hard (H) categories according to the item analysis of the instrument, the system presents an Estimated Probability of Skill Mastery graph that maps Student 1's mastery probability for each skill indicator (CCK1–HCK2 for CK and KCS1–KCC3 for PCK), complete with a mapping of sample items to each indicator. Thus, the CBT does not merely provide a final score but directly identifies which indicators have been mastered and which remain weak, in Student 1's case, several indicators show a mastery probability close to zero, so that lecturers can follow up with specific and immediate remediation. It is this automation of scoring, standardization of reporting, and diagnostic depth that affirms the added value of the CBT in assessing the professional knowledge of mathematics pre-service teachers.

At the indicator level, the same report pinpoints exactly which indicators a student has and has not mastered. For Student 1, for example, several higher-order indicators, advanced reasoning and proof within SCK, and knowledge of students within KCS registered no mastery, whereas basic content indicators were fully mastered. This illustrates how indicator-level diagnosis can guide individualized remediation far more precisely than dimension averages alone, without implying that this individual profile generalizes to the cohort.

Summary of key findings. In sum, both CK (50.00%) and PCK (52.67%) fell below the 75% mastery threshold; 60.00% of the students were in the low category and none reached the high category; SCK (36.91%) and KCC (46.25%) were the weakest dimensions; CK and PCK were strongly and positively associated ($r = 0.72$; $p < 0.001$); and the CBT was able to decompose these results down to the indicator and individual-student levels, as illustrated by the profile of Student 1.

3.2. Discussion

Interpreted against the study's objectives, three points stand out. First, neither CK nor PCK reached the mastery threshold, and contrary to the common assumption that content mastery precedes and exceeds pedagogical knowledge, PCK was marginally higher than CK; this pattern resonates with Shulman's argument that transforming content into teachable form is a distinct challenge [7] and with the MKT framework's treatment of CK and PCK as related but non-identical constructs [6]. Second, the weakest areas lay in advanced specialized content within CK and in knowledge of curriculum within PCK, echoing the COACTIV model's claim that these facets critically shape the quality of instructional planning [2], [25]. Third, CK achievement was far more homogeneous than PCK, suggesting that pedagogical capacity develops more unevenly across pre-service teachers, consistent with reviews describing PCK as the most context-dependent and hardest-to-cultivate component [8], [9].

The pattern of correlations indicates that stronger mastery in particular content dimensions tends to co-occur with the corresponding pedagogical capacity, yet content mastery accounts for only part of the variation in PCK, consistent with evidence that pre-service teachers in developing contexts often show comparatively weaker PCK because curricula emphasize content coverage over pedagogical transformation [1], [10]. Rather than forming automatically from content knowledge, PCK appears to require explicit, dedicated learning experiences. The profile of Student 1 is consistent with this interpretation but is treated here only as an illustrative example, not as evidence in its own right.

These results are broadly consistent with prior Indonesian studies reporting that the professional knowledge of mathematics pre-service teachers is generally moderate and uneven, with PCK frequently emerging as the more fragile component [10], [11]. The finding that PCK was marginally higher than CK is noteworthy and can be interpreted cautiously: the PCK items, which draw on familiar classroom situations, may be more accessible to students than the CK items, which demand advanced and abstract mathematical reasoning; moreover, the far larger dispersion of PCK ($SD = 23.96$) indicates that this apparent advantage is unevenly distributed across the cohort. The particularly low

SCK (36.91%) is plausibly linked to teacher-education curricula that emphasize procedural and elementary content over advanced mathematical reasoning—such as formal proof, mathematical modeling, and complex problem solving—so that the more cognitively demanding indicators (SCK4–SCK7) remain underdeveloped [24], [25].

Theoretically, the study offers a specific insight rather than a mere application of existing models: in a low-proficiency, underrepresented setting, CK and PCK remain strongly interdependent even though neither reaches mastery and PCK can slightly exceed CK. This qualifies the intuitive assumption that CK necessarily leads PCK and suggests that, at early proficiency levels, the two constructs may develop in tandem rather than strictly sequentially—extending the MKT and COACTIV frameworks to conditions from which they were not originally derived [6], [25], [27]. Methodologically, the study contributes a CBT model that jointly diagnoses CK and PCK down to the indicator level, addressing the scarcity of instruments that do so simultaneously [22], [28]. In policy terms, the finding that no student reached the high category provides empirical grounds for data-driven rather than assumption-driven curriculum reform [5], [23], a concern underscored by Indonesia's recent international assessment results [3].

Positioned this way, the CBT functions not merely as an examination tool but as a diagnostic-assessment ecosystem that yields rich, decomposable cognitive data actionable for program managers [23], [30]. Several limitations should nonetheless temper these conclusions and caution against generalizing beyond the single institution and the 30 participants studied here: (a) the small sample size, which limits statistical power and generalizability; (b) the total-purposive, single-institution sampling; (c) the bivariate scope, which does not model other determinants simultaneously; (d) the reliance solely on quantitative CBT data, without syllabus and document analysis, classroom observation, or interviews that could triangulate and explain why students underperformed; (e) the limited external evidence of the instrument's validity and reliability beyond the internal analyses reported here; and (f) the multiple-choice format, which cannot fully capture the reasoning that open-ended or performance-based tasks would reveal. In practical terms, these findings imply that teacher-education curricula should allocate explicit, assessed learning time to advanced specialized content (formal proof, mathematical modeling, and complex problem solving) and to knowledge of students and curriculum, and should embed diagnostic CBT within course cycles so that indicator-level weaknesses are identified early and addressed through targeted, evidence-based remediation. Future research should employ larger multi-institution and longitudinal samples, mixed-methods designs that add qualitative triangulation, and multivariate models to strengthen internal and external validity.

CONCLUSION AND RECOMMENDATIONS

This study set out to measure and profile the CK and PCK of mathematics pre-service teachers through a diagnostic CBT. On average, both constructs remained below the mastery threshold, with the majority of students in the low category and none in the high category; the weakest areas were advanced specialized content within CK and knowledge of curriculum within PCK, while CK and PCK were strongly and positively associated. Theoretically, the study clarifies that CK and PCK can remain interdependent and PCK can slightly exceed CK, even at low overall proficiency; methodologically, it demonstrates a

CBT capable of layered, indicator-level diagnosis of both constructs; and practically, it provides a standardized, efficient means of generating actionable competence profiles for teacher education. These conclusions should be read in light of the study's limitations, namely a small, single-institution, purposively selected sample analyzed with bivariate, multiple-choice-based methods. Overall, the central message is that diagnostic computer-based assessment can meaningfully advance the monitoring and development of mathematics pre-service teachers' professional knowledge, provided its profiles are used to inform evidence-based curriculum and instructional reform.

Recommendations. The following recommendations are offered as literature-informed, potentially effective alternatives suggested by, rather than directly proven by the present findings, and they should be validated in future studies.

For lecturers. Remedial instruction may target not only the weakest dimensions (SCK within CK and KCC within PCK) but also the weakest indicators identified in Table 5. In the CK domain, priority may be given to strengthening mathematical proof, proving properties or theorems (SCK4) and applying proof strategies (SCK5) through scaffolded proof practice, analysis of proof errors, and habituation to deductive argumentation, followed by mathematical modeling (SCK6) and complex problem solving (SCK7) via problem-based learning. In the PCK domain, the focus may fall on analyzing the strengths and weaknesses of representations (KCT4), applying standards and curricula (KCC2), and designing measurement instruments (KCC3) through case-based learning, analysis of teaching videos, and curriculum-document study.

For teacher education institution curriculum managers. The results encourage integrating a diagnostic CBT, capable of mapping achievement down to the indicator and individual levels into the study program's quality-assurance cycle, and rebalancing the curriculum so that advanced mathematical reasoning and pedagogical-content competence receive as much emphasis as procedural content mastery.

For future researchers. Future studies may use larger, multi-institution, and longitudinal samples, mixed-methods designs that add qualitative triangulation (syllabus analysis, classroom observation, and interviews), and multivariate models involving additional determinants (such as academic background, learning motivation, and quality of teaching) to strengthen external validity and to explain the causes of the observed weaknesses. A further agenda is to develop and empirically test enhancements to the CBT itself, automatic item-level feedback with worked explanations, direct links to remedial materials for each weak indicator, and adaptive item delivery. So that, once validated, the system can function not only as a diagnostic tool but also as an integrated learning-support environment.

ACKNOWLEDGMENTS

The authors thank the Mathematics Education study program of the teacher education institution for facilitating the implementation of the research, as well as all the students who participated as respondents.

REFERENCES

[1] X. Yang and G. Kaiser, “The impact of mathematics teachers’ professional competence on instructional quality and students’ mathematics learning outcomes,” *Curr. Opin. Behav. Sci.*, vol. 48, p. 101225, Dec. 2022, doi: 10.1016/j.cobeha.2022.101225.

[2] S. Krauss *et al.*, “Pedagogical content knowledge and content knowledge of secondary mathematics teachers.,” *J. Educ. Psychol.*, vol. 100, no. 3, pp. 716–725, Aug. 2008, doi: 10.1037/0022-0663.100.3.716.

[3] OECD, *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*. in PISA. OECD Publishing, 2023. doi: 10.1787/53f23881-en.

[4] International Association for the Evaluation of Educational Achievement, *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. Amsterdam: IEA, 2012.

[5] H. Qian and P. Youngs, “The effect of teacher education programs on future elementary mathematics teachers’ knowledge: a five-country analysis using TEDS-M data,” *J. Math. Teach. Educ.*, vol. 19, no. 4, pp. 371–396, Aug. 2016, doi: 10.1007/s10857-014-9297-0.

[6] D. Loewenberg Ball, M. H. Thames, and G. Phelps, “Content Knowledge for Teaching: What Makes It Special?,” *J. Teach. Educ.*, vol. 59, no. 5, pp. 389–407, Nov. 2008, doi: 10.1177/0022487108324554.

[7] L. S. Shulman, “Those Who Understand: Knowledge Growth in Teaching,” *Educ. Res.*, vol. 15, no. 2, pp. 4–14, Feb. 1986, doi: 10.3102/0013189X015002004.

[8] M. Grigaliūnienė, E. Lehtinen, L. Verschaffel, and F. Depaepe, “Systematic Review of Research on Pedagogical Content Knowledge in Mathematics: Insights from a Topic-Specific Approach,” *ZDM – Math. Educ.*, vol. 57, no. 4, pp. 777–794, Aug. 2025, doi: 10.1007/s11858-025-01684-1.

[9] F. Depaepe, L. Verschaffel, and G. Kelchtermans, “Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research,” *Teach. Teach. Educ.*, vol. 34, pp. 12–25, Aug. 2013, doi: 10.1016/j.tate.2013.03.001.

[10] S. Y. Ningsih, Turmudi, and D. Juandi, “Pedagogical content knowledge (PCK) profile of prospective teachers in mathematics learning,” *J. Phys. Conf. Ser.*, vol. 1521, no. 3, p. 032057, Apr. 2020, doi: 10.1088/1742-6596/1521/3/032057.

[11] S. Fiangga, S. Khabibah, S. M. Amin, and R. Ekawati, “A Learning Design Analysis of the Pre-service Teachers’ Mathematics Pedagogical Content Knowledge,” *J. Phys. Conf. Ser.*, vol. 1899, no. 1, p. 012153, May 2021, doi: 10.1088/1742-6596/1899/1/012153.

[12] X. Yang, B. Schwarz, and I. K. C. Leung, “Pre-service mathematics teachers’ professional modeling competencies: a comparative study between Germany, Mainland China, and Hong Kong,” *Educ. Stud. Math.*, vol. 109, no. 2, pp. 409–429, Feb. 2022, doi: 10.1007/s10649-021-10064-x.

[13] J. Njiku, “Pre-service teachers manifested mathematics pedagogical content knowledge: The role of the teaching practicum,” *Pedagog. Res.*, vol. 10, no. 1, p. em0229, Jan. 2025, doi: 10.29333/pr/15647.

[14] A. J. Boevé, R. R. Meijer, C. J. Albers, Y. Beetsma, and R. J. Bosker, “Introducing Computer-Based Testing in High-Stakes Exams in Higher Education: Results of a Field Experiment,” *PLOS ONE*, vol. 10, no. 12, p. e0143616, Dec. 2015, doi: 10.1371/journal.pone.0143616.

[15] C. Y. Piaw, “Replacing Paper-based Testing with Computer-based Testing in Assessment: Are we Doing Wrong?,” *Procedia - Soc. Behav. Sci.*, vol. 64, pp. 655–664, Nov. 2012, doi: 10.1016/j.sbspro.2012.11.077.

[16] A. Berry, P. Friedrichsen, and J. Loughran, Eds., *Re-examining pedagogical content knowledge in science education*. in Teaching and learning in science series. London New York: Routledge, 2015. doi: 10.4324/9781315735665.

[17] M. Schmid, E. Brianza, and D. Petko, “Self-reported technological pedagogical content knowledge (TPACK) of pre-service teachers in relation to digital technology use in lesson plans,” *Comput. Hum. Behav.*, vol. 115, p. 106586, Feb. 2021, doi: 10.1016/j.chb.2020.106586.

[18] L. Maas, M. J. S. Brinkhuis, L. Kester, and L. Wijngaards-de Meij, “Cognitive Diagnostic Assessment in University Statistics Education: Valid and Reliable Skill Measurement for Actionable Feedback Using Learning Dashboards,” *Appl. Sci.*, vol. 12, no. 10, p. 4809, May 2022, doi: 10.3390/app12104809.

[19] A. Barana, M. Marchisio, and M. Sacchet, “Interactive Feedback for Learning Mathematics in a Digital Learning Environment,” *Educ. Sci.*, vol. 11, no. 6, p. 279, Jun. 2021, doi: 10.3390/educsci11060279.

[20] S. Rezat, “How automated feedback from a digital mathematics textbook affects primary students’ conceptual development: two case studies,” *ZDM – Math. Educ.*, vol. 53, no. 6, pp. 1433–1445, Nov. 2021, doi: 10.1007/s11858-021-01263-0.

[21] L. N. Yaqin, L. D. Prasojo, N. A. Haji-Othman, N. Yusof, and A. Habibi, “Addressing the Digital Divide in Indonesian Higher Education: Insights, Implications, and Potential Solutions,” in *From Digital Divide to Digital Inclusion*, Ł. Tomczyk, F. D. Guillén-Gámez, J. Ruiz-Palmero, and A. Habibi, Eds., in Lecture Notes in Educational Technology. , Singapore: Springer Nature Singapore, 2023, pp. 291–307. doi: 10.1007/978-981-99-7645-4_13.

[22] Naila Rajiha and S. Rini, “COMPUTER-BASED STANDARDIZED TESTING (CBT) SUMMATIVE ASSESSMENT: ONE INSTITUTION EXPERIENCE,” *JISAE J. Indones. Stud. Assess. Eval.*, vol. 9, no. 1, pp. 1–7, Jan. 2023, doi: 10.21009/jisae.v9i1.33233.
[23] G. Kaiser and J. König, “Competence Measurement in (Mathematics) Teacher Education and Beyond: Implications for Policy,” *High. Educ. Policy*, vol. 32, no. 4, pp. 597–615, Dec. 2019, doi: 10.1057/s41307-019-00139-z.
[24] T. Kleickmann *et al.*, “Teachers’ Content Knowledge and Pedagogical Content Knowledge: The Role of Structural Differences in Teacher Education,” *J. Teach. Educ.*, vol. 64, no. 1, pp. 90–106, Jan. 2013, doi: 10.1177/0022487112460398.
[25] J. Baumert and M. Kunter, “The COACTIV Model of Teachers’ Professional Competence,” in *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers*, M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, and M. Neubrand, Eds., Boston, MA: Springer US, 2013, pp. 25–48. doi: 10.1007/978-1-4614-5149-5_2.
[26] B. Rowan, R. Correnti, and R. J. Miller, “What Large-Scale, Survey Research Tells Us about Teacher Effects on Student Achievement: Insights from the Prospects Study of Elementary Schools,” *Teach. Coll. Rec. Voice Scholarsh. Educ.*, vol. 104, no. 8, pp. 1525–1567, Aug. 2002, doi: 10.1111/1467-9620.00212.
[27] S. Blömeke, J.-E. Gustafsson, and R. J. Shavelson, “Beyond Dichotomies: Competence Viewed as a Continuum,” *Z. Für Psychol.*, vol. 223, no. 1, pp. 3–13, Jan. 2015, doi: 10.1027/2151-2604/a000194.
[28] E. Mohammadpour and Y. Maroofi, “A performance-based test to measure teachers’ mathematics and science content and pedagogical knowledge,” *Heliyon*, vol. 9, no. 3, p. e13932, Mar. 2023, doi: 10.1016/j.heliyon.2023.e13932.
[29] T. Clark and H. Endres, “Computer-based diagnostic assessment of high school students’ grammar skills with automated feedback – an international trial,” *Assess. Educ. Princ. Policy Pract.*, vol. 28, no. 5–6, pp. 602–632, Nov. 2021, doi: 10.1080/0969594X.2021.1970513.
[30] G. A. Nortvedt and N. Buchholtz, “Assessment in mathematics education: responding to issues regarding methodology, policy, and equity,” *ZDM*, vol. 50, no. 4, pp. 555–570, Aug. 2018, doi: 10.1007/s11858-018-0963-z.

Bukti Cek Similaritas/Plagiasi

10% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography

Match Groups

47 Not Cited or Quoted 8%
Matches with neither in-text citation nor quotation marks

10 Missing Quotations 2%
Matches that are still very similar to source material

0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

7% Internet sources
7% Publications
2% Submitted works (Student Papers)

Match Groups

- **47** Not Cited or Quoted 8%
  Matches with neither in-text citation nor quotation marks
- **10** Missing Quotations 2%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

7% Internet sources
7% Publications
2% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Publication	Henry Obiora Chukwudi, Babatope Osagbemiro. "Risk Awareness about Tetracycl...	1%
2	Internet	e-journal.iainptk.ac.id	<1%
3	Internet	www.frontiersin.org	<1%
4	Internet	journal-gehu.com	<1%
5	Publication	Aisling Leavy, Mairead Hourigan. "Using Lesson Study to Support the Teaching of ...	<1%
6	Internet	cahaya-ic.com	<1%
7	Publication	Nurmawati Nurmawati, Sri Kadarwati, Eko Andy Purnomo, Venissa Dian Mawarsa...	<1%
8	Student papers	Utah Education Network	<1%
9	Internet	ijmra.in	<1%
10	Internet	jals.hmu.ac.ir	<1%

11 Internet

journal.globresco.com <1%

12 Student papers

University of Hong Kong <1%

13 Student papers

IAIN Syekh Nurjati Cirebon <1%

14 Internet

journal.akademimerdeka.com <1%

15 Publication

"International Perspectives on Teacher Knowledge, Beliefs and Opportunities to ... <1%

16 Internet

doczz.net <1%

17 Internet

perinataljournal.com <1%

18 Internet

www.coursehero.com <1%

19 Student papers

Universiti Sains Malaysia <1%

20 Internet

archive.org <1%

21 Internet

download.atlantis-press.com <1%

22 Internet

erepo.uef.fi <1%

23 Internet

www.researchgate.net <1%

24 Internet

www.science.gov <1%

25 Publication

Prospective Mathematics Teachers' Knowledge of Algebra, 2014. <1%

26 Publication

Sasmayunita Sasmayunita, Sultan Sultan, Mantasiah Mantasiah, Sakaria Sakaria, ... <1%

27 Internet

abjournals.org <1%

28 Internet

f1000research.com <1%

29 Internet

hal.science <1%

30 Internet

ndl.ethernet.edu.et <1%

31 Internet

siducat.org <1%

32 Internet

www.jurnal.stit-buntetpesantren.ac.id <1%

33 Publication

"Generative AI in Higher Education Assessment", Springer Science and Business ... <1%

34 Publication

Depaepe, Fien, Joke Torbeyns, Nathalie Vermeersch, Dirk Janssens, Rianne Jansse... <1%

35 Publication

Vinodhini D., Vinoth Raman, Stephan Antony Raj A., Vanitha V.. "chapter 9 Enhanc... <1%

Journal of Mathematics Instruction, Social Research and Opinion
Vol. 5, No. 3, September 2026, pp. 2271 – 2285, https://doi.org/10.58421/misro.v5i3.1770
ISSN 2962-7842

Assessing Mathematics Pre-Service Teachers' Content Knowledge and Pedagogical Content Knowledge through a Diagnostic Computer-Based Test

Reza Oktiana Akbar[1], Muhamad Ali Misri[2]
[1,2]Universitas Islam Negeri Siber Syekh Nurjati Cirebon, Indonesia

Article Info

Article history:

Received 2026-06-20
Revised 2026-09-01
Accepted 2026-09-02

Keywords:

Computer-based test
Content knowledge
Diagnostic assessment
Mathematical knowledge for teaching
Mathematics pre-service teachers
Pedagogical content knowledge

ABSTRACT

Content Knowledge (CK) and Pedagogical Content Knowledge (PCK) are core components of mathematics teachers' competence, yet they are rarely measured simultaneously through technology-based diagnostic assessment. This study measures and profiles the CK and PCK of mathematics pre-service teachers using a diagnostic Computer-Based Test (CBT). A quantitative descriptive-correlational design was applied to 30 purposively selected pre-service teachers. Each construct was measured with 25 multiple-choice items whose content validity was established by expert judgment and whose reliability was confirmed before use (KR-20 = 0.81 for CK; 0.79 for PCK). Data were analyzed with descriptive statistics, Pearson correlation, and simple linear regression after assumption checks. Mean CK (50.00%) and PCK (52.67%) were both below the 75% mastery threshold, with 60% of students in the low category; specialized content knowledge and knowledge of content and curriculum were the weakest dimensions. CK and PCK correlated strongly and positively ($r = 0.72$; $p < 0.001$), with CK accounting for about 52% of the variance in PCK as a predictive, non-causal association. The CBT generated actionable diagnostic profiles at the dimension and indicator levels. These findings support using diagnostic CBT to guide data-driven curriculum reform integrating content mastery and pedagogical capacity.

Corresponding Author:

Reza Oktiana Akbar
Mathematics Education Study Program, UIN Siber Syekh Nurjati Cirebon, Indonesia
Email: rezaoktiana@uinssc.ac.id

1. INTRODUCTION

The quality of a nation's mathematics education depends heavily on teachers' professional competence. Cross-national studies have shown that mathematics teachers' professional knowledge, comprising Content Knowledge (CK) and Pedagogical Content Knowledge (PCK), is a primary predictor of instructional effectiveness and students' academic achievement [1], [2]. The 2022 PISA results placed Indonesian students' mathematics literacy 68th out of 81 countries, below the OECD average, indicating a fundamental problem in the quality of mathematics instruction [3]. This condition is inseparable from the quality of professional knowledge formed in teacher education institutions, so ensuring an adequate professional-knowledge profile of pre-service teachers becomes a strategic imperative [4], [5].

Theoretically, the Mathematical Knowledge for Teaching (MKT) framework distinguishes CK, comprising common, specialized, and horizon content knowledge, from PCK, which encompasses knowledge of content and students, content and teaching, and content and curriculum [6], [7]. Systematic reviews affirm that PCK is the most frequently studied yet the most difficult component to measure consistently because of its contextual and multidimensional nature [8], [9]. Studies in Indonesia show that the professional knowledge of mathematics pre-service teachers varies widely across institutions and curricula, with a notable weakness in the PCK aspect [10], [11]. Similar findings have been reported in cross-national contexts [12], [13].

Alongside the rapid digitalization of education, competency-evaluation systems are now shifting massively from traditional approaches toward technology-based assessment. Various studies indicate that CBT is often more valid and reliable than conventional examination methods [14], [15]; however, its use to objectively measure both the content knowledge and the pedagogical content knowledge of pre-service teachers simultaneously remains very limited because of the complexity of these constructs [16], [17]. The potential of CBT to generate diagnostic feedback based on learning analytics has also not been optimally exploited [18], even though interactive digital feedback has been shown to improve mathematical understanding [19], [20].

The urgency of this study is heightened in the context of Indonesia's uneven higher-education digital transformation [21], which is not yet supported by adequate standardization of assessment systems [22]. Measuring teacher competence in the digital era is required to produce data that are actionable for study-program managers and policymakers [23]. A number of studies affirm that CK and PCK are relatively independent constructs that require different intervention pathways [24], [25], and that teacher effects on student achievement are substantial [26]. Taken together, however, this body of work remains fragmented: CK and PCK are rarely measured simultaneously with a single validated instrument, the diagnostic potential of CBT is seldom exploited at the indicator level, and the evidence is concentrated in a few well-studied national settings, leaving a clear gap for integrated, technology-based diagnosis of both constructs.

Building on this gap, the present study makes three specific contributions rather than a general claim to novelty. Theoretically, it provides empirical evidence on how CK and PCK co-vary at low overall proficiency, showing that their interdependence persists even when neither reaches mastery, an insight that refines, rather than merely applies, the MKT and COACTIV frameworks [6], [25], [27]. Methodologically, it operationalizes a CBT that produces layered, indicator-level diagnostic profiles for both constructs simultaneously, addressing the scarcity of validated instruments that jointly diagnose CK and PCK [28], [29]. Practically, it demonstrates how such diagnostic data can be translated into targeted curriculum and instructional decisions. By drawing on an underrepresented context, the study also contributes comparative evidence to the international literature on teacher knowledge rather than a purely local application of CBT [12], [30]. Accordingly, the study aims to: (1) measure and profile pre-service teachers' CK and PCK; (2) identify strengths and weaknesses at the dimension and indicator levels; and (3) demonstrate how diagnostic profiles can inform instructional intervention.

2. METHOD

This study employed a quantitative approach with a descriptive-survey and correlational design. The population comprised students of the Mathematics Education study program at a teacher education institution who were in their fifth and seventh semesters. A sample of 30 students was determined through purposive sampling with the inclusion criteria: (1) active students who had adequately completed mathematics-content and pedagogy courses; and (2) willing to take the entire series of CBT tests. Given the limited sample size, all data were analyzed as a single intact group (n = 30) and were not directed toward inter-semester comparison.

Research setting and participants. The research was conducted in the odd semester of the 2024/2025 academic year in the Mathematics Education study program of a state Islamic teacher education institution in Indonesia. Participants were selected using total (saturated) purposive sampling: all 30 students who met the inclusion criterion, having completed the core mathematics-content and pedagogy courses, were included, so that the sample comprised the entire eligible cohort rather than a probability sample. The sample size (n = 30) therefore reflects the actual number of eligible students at the single participating institution and suits the study's diagnostic, exploratory aim of profiling CK and PCK in depth; it is nonetheless acknowledged as a constraint on statistical power and on the generalizability of the correlational and regression results, as elaborated in the Discussion.

The operationalization of the variables comprised two main constructs. CK was defined as mastery of the mathematics content to be taught, measured through 25 multiple-choice items; PCK was defined as the ability to integrate content knowledge with pedagogical strategies, also measured through 25 multiple-choice items. Scores were converted into percentages of achievement relative to the ideal maximum score. The CBT instrument was designed to generate layered diagnostic reports, namely the total score, achievement per dimension, achievement per indicator, and an individual achievement profile for each student. Each item was mapped to a specific indicator and dimension (see the notes below Table 1) so that the CBT system could produce achievement per dimension and per indicator for all students at once, as well as their individual diagnostic profiles. The composition of the instrument is presented in Table 1. The multiple-choice format was chosen to enable fully automated, standardized, and scalable scoring across all participants and indicators at once, a prerequisite for the layered diagnostic reporting; its rationale and limitations are elaborated under 'Item format'.

Table 1. Composition of CBT Instrument Items by MKT Dimension

Domain	Dimension	Number of Items
CK	Common Content Knowledge CCK)	7
CK	Specialized Content Knowledge (SCK)	14
CK	Horizon Content Knowledge (HCK)	4
PCK	Knowledge of Content and Students (KCS)	10
PCK	Knowledge of Content and Teaching (KCT)	9
PCK	Knowledge of Content and Curriculum (KCC)	6
Total		50

 https://doi.org/10.58421/misro.v5i3.1770

Notes on dimension codes and indicator coverage (based on the developed CK and PCK instruments). **CK – CCK** (Common Content Knowledge, items 1–7): understanding basic concepts (e.g., lines and angles), performing calculations and measurements accurately, and applying mathematical concepts in everyday life. **CK – SCK** (Specialized Content Knowledge, items 8–21): explaining concepts and principles, analyzing and connecting mathematical objects, proving properties, building models, and solving complex mathematical problems. **CK – HCK** (Horizon Content Knowledge, items 22–25): connecting school mathematics with advanced academic mathematics and with other disciplines. **PCK – KCS** (Knowledge of Content and Students, items 1–10): analyzing students' preconceptions and understanding as well as predicting and analyzing students' misconceptions, errors, and difficulties. **PCK – KCT** (Knowledge of Content and Teaching, items 11–19): selecting and organizing learning materials, selecting sources, strategies, and techniques, and selecting and analyzing representations for presenting material. **PCK – KCC** (Knowledge of Content and Curriculum, items 20–25): designing content and topic combinations, applying standards and curricula, and designing instruments to measure students' abilities.

Instrument development and content validity. The CBT instrument was developed through a systematic procedure. First, a test blueprint was constructed by mapping each item to a specific MKT dimension and indicator (Table 1). Second, the draft items were reviewed by three experts—two mathematics-education lecturers and one educational-assessment specialist—to establish content validity; the level of expert agreement was quantified using Aiken's V, and items with $V \geq 0.80$ were retained while the remaining items were revised or replaced. Third, the revised items were trialed and analyzed empirically before being administered in the present study. To illustrate the item format, each CK item presented a mathematical task with four to five options, whereas each PCK item was framed as a short classroom scenario or a sample of student work followed by options representing alternative pedagogical decisions. The complete test blueprint and sample items are available from the authors upon reasonable request.

Reliability and item analysis. Based on the item analysis, the instrument demonstrated adequate psychometric quality. Internal-consistency reliability computed with the Kuder–Richardson formula (KR-20) reached 0.81 for the CK test and 0.79 for the PCK test, both within the acceptable-to-good range. The item difficulty indices ranged from 0.20 to 0.87 (a balanced mix of easy, medium, and hard items), and the discrimination indices were largely satisfactory ($D \geq 0.30$); only items meeting the validity, difficulty, and discrimination criteria were retained in the final CBT.

Item format. Multiple-choice items were chosen because the CBT was designed for automated, standardized, and scalable diagnosis across many participants and indicators simultaneously. To capture pedagogical reasoning rather than mere recall, the PCK items were built around short classroom scenarios, samples of student work, and instructional-decision situations, with distractors constructed from typical student misconceptions and common teaching errors. It is nevertheless acknowledged that the multiple-choice format cannot fully capture the depth of pedagogical reasoning that open-ended or performance-based tasks would reveal, a limitation revisited in the Discussion.

The analysis techniques comprised descriptive and inferential analyses. Descriptive analysis produced the mean (M), standard deviation (SD), minimum, maximum, median, and percentage of achievement at the total, dimension, indicator, and individual levels, as well as a three-level categorization (High, Medium, Low). The categorization used the following cut-off points: High (≥ 75%), Medium (60–74.99%), and Low (< 60%), where the 75% mark represents the ideal mastery criterion adopted in this study. Inferential analysis was conducted through two techniques: (1) Pearson correlation to test the CK–PCK relationship and the corresponding inter-dimension relationships (CCK–KCS, SCK–KCT, HCK–KCC) at $\alpha = 0.05$; and (2) simple linear regression with CK as the predictor and PCK as the criterion. Prior to the inferential tests, the underlying assumptions were examined, namely the normality of the CK and PCK distributions (Shapiro–Wilk), the linearity of their relationship, the absence of influential outliers, and homoscedasticity; all assumptions were adequately met, so that the parametric analyses were appropriate.

All scoring, achievement computation, correlation, and linear regression were performed using Microsoft Excel through standardized formulas (including CORREL, SLOPE, INTERCEPT, and RSQ) so that every figure can be traced back to the raw data. The study also observed standard research-ethics principles: participation was entirely voluntary, informed consent was obtained from all respondents before data collection, and the participants' anonymity and confidentiality were maintained throughout the recording, processing, and reporting of data by using codes instead of names, with the individual-profile example presented under the pseudonym Student 1.

3. RESULTS AND DISCUSSION

3.1. Results

This section reports the findings of the data analysis in the following order: respondent characteristics, descriptive statistics, the distribution of achievement categories, dimension- and indicator-level achievement, the correlation and regression results, and an individual diagnostic profile. The 30 respondents were fifth- and seventh-semester students of the Mathematics Education study program who had completed the core mathematics-content and pedagogy courses; all of them completed the full series of CK and PCK tests through the CBT system.

Overall, students' professional-knowledge achievement had not yet reached the ideal competency threshold (≥ 75%). The mean CK score was 50.00% (SD = 10.16), PCK 52.67% (SD = 23.96), and the combined Professional Knowledge (PK) score 51.33% (SD = 16.04). The complete descriptive statistics are presented in Table 2.

Table 2. Descriptive Statistics of CK, PCK, and Professional Knowledge Scores (n = 30)

Variable	Mean (%)	SD	Min	Max	Median
Content Knowledge (CK)	50.00	10.16	24	68	52.00
Pedagogical Content Knowledge (PCK)	52.67	23.96	8	96	54.00
Professional Knowledge (PK)	51.33	16.04	26	78	50.00

Table 2 shows that the mean PCK is slightly higher than CK (a difference of 2.67 points), an interesting pattern indicating that pedagogical-transformation capacity slightly exceeds content mastery. The large standard deviation of PCK (SD = 23.96) and the much smaller one for CK (SD = 10.16) reflect substantial heterogeneity in knowledge profiles, with score ranges of 24–68% for CK and 8–96% for PCK. The distribution of Professional Knowledge categories is presented in Table 3.

Table 3. Distribution of Professional Knowledge Categories (n = 30)

Category	Frequency (f)	Percentage (%)
High	0	0.00
Medium	12	40.00
Low	18	60.00
Total	30	100.00

https://doi.org/10.58421/misro.v5i3.1770

All students (100%) were in the Low or Medium category, with 60.00% in the Low category and 40.00% in the Medium category, and no student reached the High category. This indicates that the majority of pre-service teachers have not attained an optimal level of professional-knowledge mastery. Achievement per dimension is presented in Table 4 and Figure 1.

Table 4. Average Achievement per CK and PCK Dimension

Domain	Dimension	Items	Mean (%)	SD
CK	CCK	7	73.33	20.47
CK	SCK	14	36.91	24.16
CK	HCK	4	55.00	35.60
PCK	KCS	10	54.07	33.18
PCK	KCT	9	57.50	26.14
PCK	KCC	6	46.25	32.30

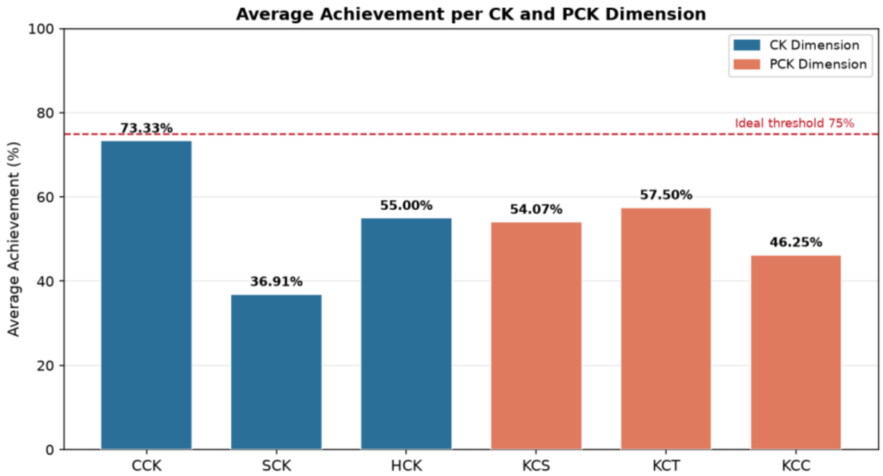


Figure 1. Average Achievement per CK and PCK Dimension

In the CK domain, the CCK dimension was the highest (73.33%) and SCK the lowest (36.91%); in the PCK domain, KCT was the highest (57.50%) while KCC was the lowest of all six dimensions (46.25%). Diagnostically, the low SCK and KCC signal a weak ability to relate content to the structure and organization of the curriculum—an area that needs to be prioritized in intervention. Achievement at the indicator level is presented in Table 5 and visualized in Figure 2.

Table 5. Average Achievement per CK and PCK Indicator

Domain	Indicator	Mean (%)
CK	CCK1	86.67
CK	CCK2	80.00
CK	CCK3	46.67
CK	SCK1	63.33
CK	SCK2	46.67
CK	SCK3	43.33
CK	SCK4	20.00
CK	SCK5	23.33
CK	SCK6	30.00
CK	SCK7	33.33
CK	HCK1	53.33
CK	HCK2	56.67
PCK	KCS1	58.33
PCK	KCS2	48.33

https://doi.org/10.58421/misro.v5i3.1770

Domain	Indicator	Mean (%)
PCK	KCS3	58.33
PCK	KCS4	51.67
PCK	KCS5	53.33
PCK	KCT1	55.00
PCK	KCT2	66.67
PCK	KCT3	60.00
PCK	KCT4	36.67
PCK	KCC1	51.67
PCK	KCC2	43.33
PCK	KCC3	45.00

Notes on indicator codes (based on the CK and PCK instruments). CK indicators — **CCK1**: understanding basic school-mathematics concepts, such as lines, angles, sequences, and basic statistics (items 1–3); **CCK2**: performing calculations and measurements accurately (items 4–5); **CCK3**: applying mathematical concepts in everyday life (items 6–7); **SCK1**: explaining academic mathematical concepts and principles (items 8–9); **SCK2**: analyzing mathematical objects (items 10–11); **SCK3**: connecting mathematical properties and objects (items 12–13); **SCK4**: proving mathematical properties or theorems (items 14–16); **SCK5**: applying proof strategies (item 17); **SCK6**: building and constructing mathematical models (items 18–19); **SCK7**: solving complex mathematical problems (items 20–21); **HCK1**: connecting school mathematics with advanced academic mathematics (items 22–23); **HCK2**: connecting mathematics with other disciplines (items 24–25). PCK indicators — **KCS1**: analyzing students' preconceptions (items 1–2); **KCS2**: identifying students' level of understanding (items 3–4); **KCS3**: predicting students' misconceptions (items 5–6); **KCS4**: identifying sources of students' learning difficulties (items 7–8); **KCS5**: analyzing students' literacy issues and errors (items 9–10); **KCT1**: organizing and sequencing learning materials (items 11–12); **KCT2**: selecting learning sources, strategies, and techniques (items 13–14); **KCT3**: selecting representations and illustrations for presenting material (items 15–17); **KCT4**: analyzing the weaknesses and strengths of representations (items 18–19); **KCC1**: composing content and topic combinations (items 20–21); **KCC2**: applying standards and curricula (items 22–23); **KCC3**: designing instruments to measure students' abilities (items 24–25).

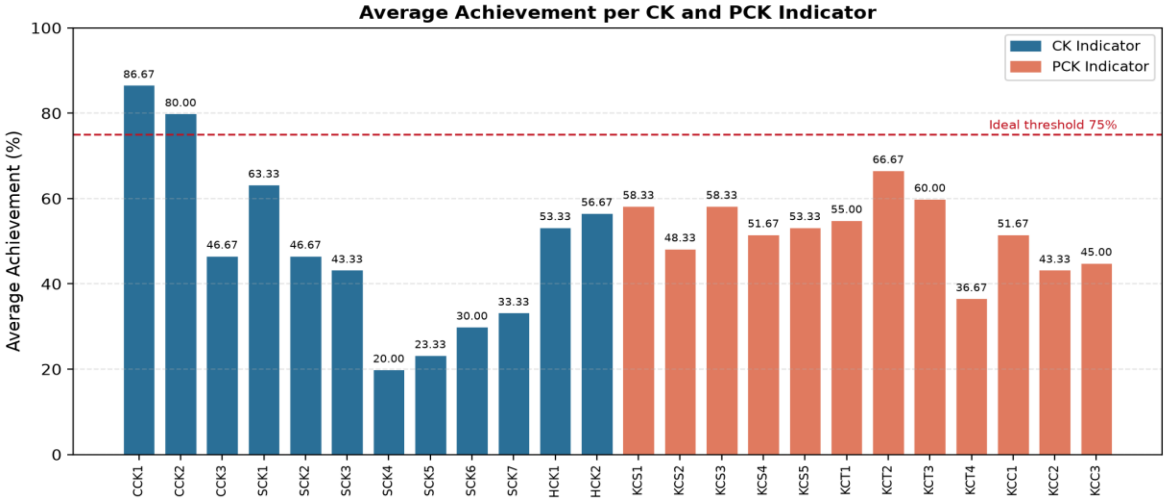


Figure 2. Average Achievement per CK and PCK Indicator

Figure 2 visualizes the average achievement of each indicator exactly as listed in Table 5. The range of achievement across indicators spans 20.00% to 86.67%. The highest achievement is on the CK indicator CCK1 (86.67%; understanding basic school-mathematics concepts), which is fundamental and therefore the most widely mastered by students. Conversely, the lowest overall achievement is on indicator SCK4 (20.00%; proving mathematical properties or theorems), the skill with the highest cognitive demand in the CK domain, followed by SCK5 (23.33%; applying proof strategies) and SCK6 (30.00%; building and constructing mathematical models). The pattern in the SCK dimension declines gradually with complexity, from explaining concepts and principles (SCK1; 63.33%), analyzing and connecting mathematical objects (SCK2 46.67%; SCK3 43.33%), solving complex problems (SCK7; 33.33%), to the most difficult modeling and proving (SCK6 30.00%; SCK5 23.33%; SCK4 20.00%), a pattern consistent with cognitive load theory. Because each dimension's achievement remains the mean of its constituent indicators, for example, KCS (54.07%) is the mean of KCS1–KCS5 and CCK (73.33%) is the mean of CCK1–CCK3—this pattern maps precisely which items and indicators are the sources of weakness, so that lecturers can design remediation targeted at concrete rather than generic weaknesses. The results of the Pearson correlation test are presented in Table 6.

Table 6. Pearson Correlation between Domains and between Parallel Dimensions

Relationship	r	p	Interpretation
CK – PCK	0.72	< 0.001	Strong
CCK – KCS	0.42	< 0.05	Moderate
SCK – KCT	0.54	< 0.01	Moderate
HCK – KCC	0.86	< 0.001	Very strong

The strong positive correlation between CK and PCK ($r = 0.72$; $p < 0.001$) confirms that the two domains are constructively interdependent, reinforced by the correlations between parallel dimensions, especially HCK–KCC ($r = 0.86$) and SCK–KCT ($r = 0.54$). The predictive relationship of CK on PCK was tested through simple linear regression (Table 7 and Figure 3).

Table 7. Results of the Simple Linear Regression of CK on PCK

Parameter	Value
Slope (b)	1.708
Intercept (a)	−32.734
Coefficient of determination (R^2)	0.5245
Pearson r	0.72
Equation	$PCK = 1.708 \times CK - 32.734$

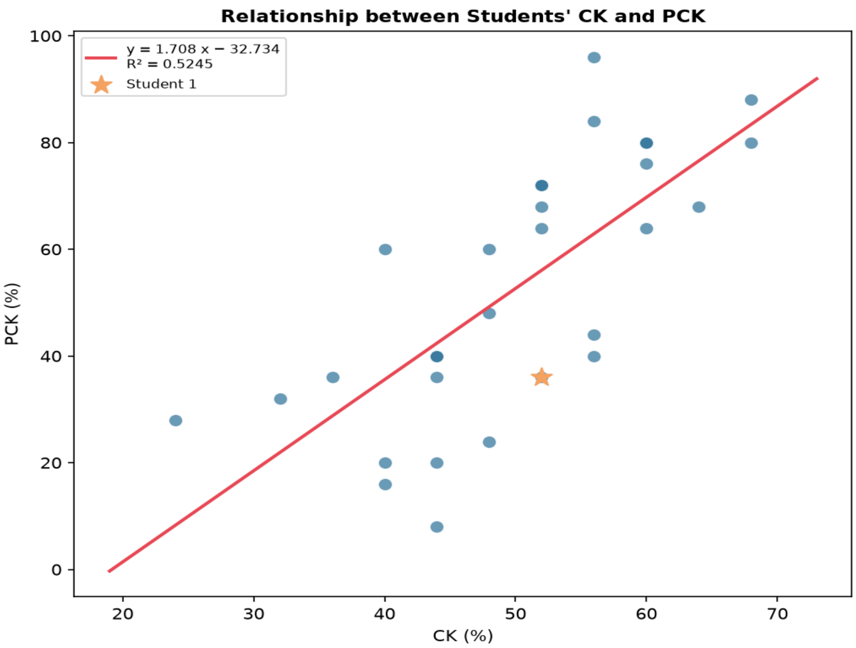


Figure 3. Scatter Plot and Regression Line of the CK–PCK Relationship

The equation $PCK = 1.708 \times CK - 32.734$ with $R^2 = 0.52$ indicates that about 52% of the variance in PCK is statistically associated with CK. Because the design is correlational, this relationship should be interpreted as a predictive association rather than a causal effect; the remaining 48% of the variance signals that PCK does not form automatically from content mastery but requires an explicit pedagogical learning pathway.

The CBT also breaks aggregate achievement down into individual diagnostic profiles automatically. To illustrate this capability, rather than to draw general conclusions from a single case, Table 8 and Figure 4 summarize the profile of one participant (pseudonym Student 1), and Figures 5 and 6 show anonymized screenshots of the corresponding diagnostic report generated by the CBT. Such reports are produced automatically for every participant.

Table 8. Individual Achievement Profile of Student 1 Compared with the Class Average

Domain	Student 1 Score (%)	Class Average (%)	Difference	Category
Content Knowledge (CK)	52.00	50.00	+2.00	Low
Pedagogical Content Knowledge (PCK)	36.00	52.67	−16.67	Low

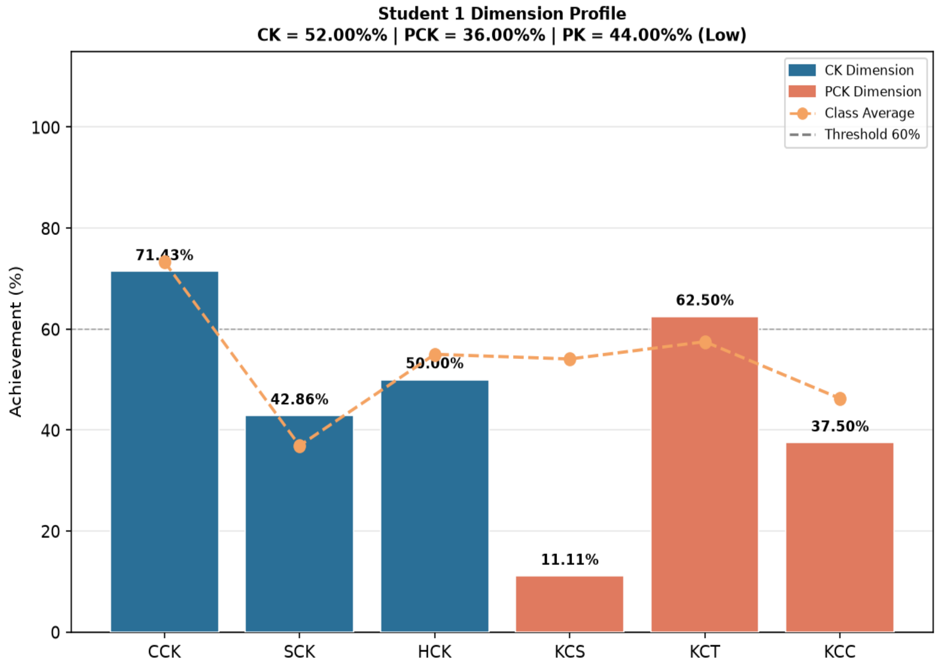


Figure 4. CK and PCK Achievement Profile of Student 1 Compared with the Class Average

As an illustrative example, Student 1 answered 13 of 25 CK items and 9 of 25 PCK items correctly, falling into the Low category for both domains, with CK slightly above and PCK well below the class average. Anonymized screenshots of the report are shown in Figures 5 and 6, with no identifiable student information displayed. This single case is used only to demonstrate the diagnostic output of the CBT and is not intended as a basis for general interpretation.

 https://doi.org/10.58421/misro.v5i3.1770

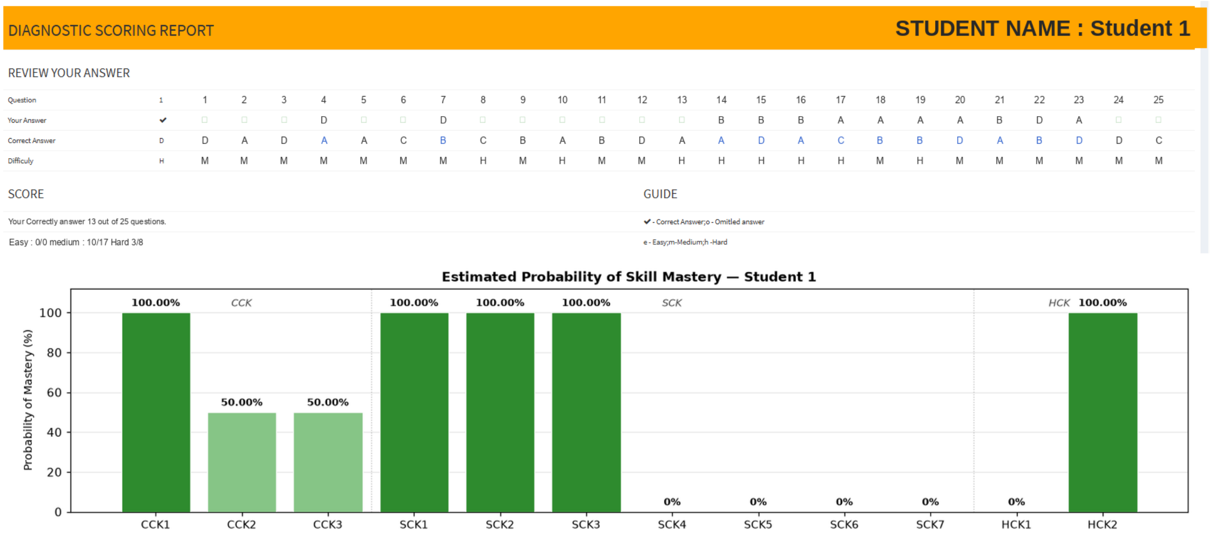


Figure 5. CBT Diagnostic Scorecard for the CK Domain for Student 1 (original system screenshot)

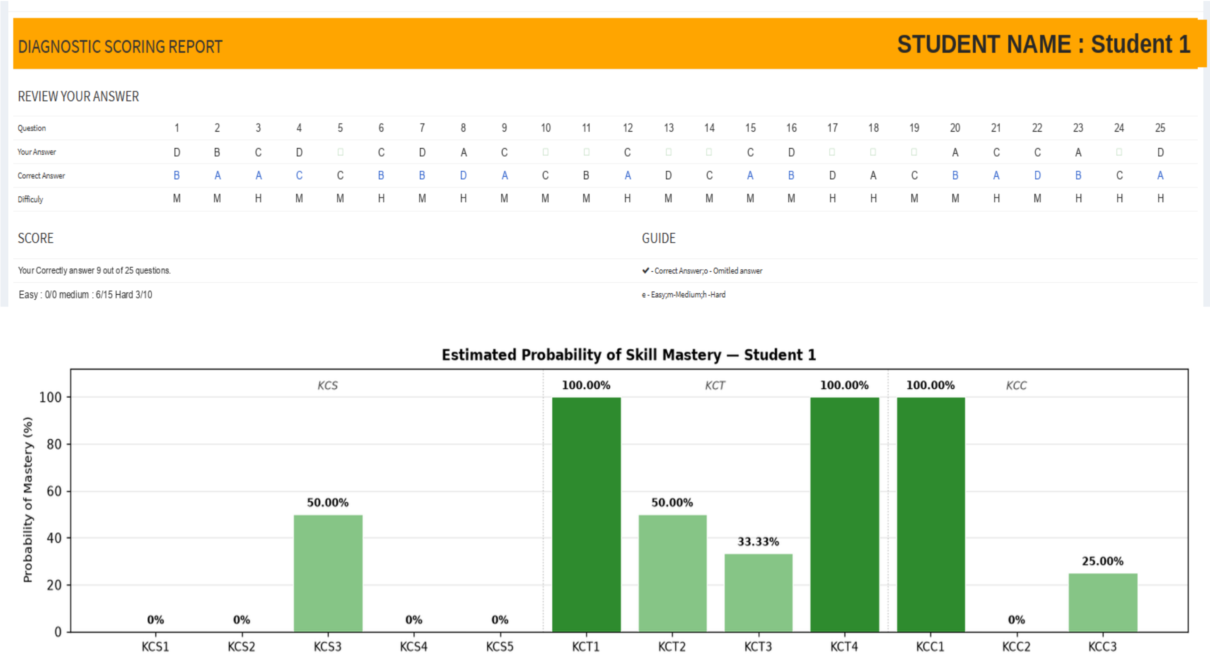


Figure 6. CBT Diagnostic Scorecard for the PCK Domain for Student 1 (system screenshot)

Figures 5 and 6 reveal diagnostic functions of the CBT that are more difficult to implement efficiently in conventional paper-based tests. In addition to recapping correct/incorrect answers per item along with their difficulty level, a combination of the Medium (M) and Hard (H) categories according to the item analysis of the instrument, the system presents an Estimated Probability of Skill Mastery graph that maps Student 1's mastery probability for each skill indicator (CCK1–HCK2 for CK and KCS1–KCC3 for PCK), complete with a mapping of sample items to each indicator. Thus, the CBT does not merely provide a final score but directly identifies which indicators have been mastered and which remain weak, in Student 1's case, several indicators show a mastery probability close to zero, so that lecturers can follow up with specific and immediate remediation. It is this automation of scoring, standardization of reporting, and diagnostic depth that affirms the added value of the CBT in assessing the professional knowledge of mathematics pre-service teachers.

At the indicator level, the same report pinpoints exactly which indicators a student has and has not mastered. For Student 1, for example, several higher-order indicators, advanced reasoning and proof within SCK, and knowledge of students within KCS registered no mastery, whereas basic content indicators were fully mastered. This illustrates how indicator-level diagnosis can guide individualized remediation far more precisely than dimension averages alone, without implying that this individual profile generalizes to the cohort.

Summary of key findings. In sum, both CK (50.00%) and PCK (52.67%) fell below the 75% mastery threshold; 60.00% of the students were in the low category and none reached the high category; SCK (36.91%) and KCC (46.25%) were the weakest dimensions; CK and PCK were strongly and positively associated ($r = 0.72$; $p < 0.001$); and the CBT was able to decompose these results down to the indicator and individual-student levels, as illustrated by the profile of Student 1.

3.2. Discussion

Interpreted against the study's objectives, three points stand out. First, neither CK nor PCK reached the mastery threshold, and contrary to the common assumption that content mastery precedes and exceeds pedagogical knowledge, PCK was marginally higher than CK; this pattern resonates with Shulman's argument that transforming content into teachable form is a distinct challenge [7] and with the MKT framework's treatment of CK and PCK as related but non-identical constructs [6]. Second, the weakest areas lay in advanced specialized content within CK and in knowledge of curriculum within PCK, echoing the COACTIV model's claim that these facets critically shape the quality of instructional planning [2], [25]. Third, CK achievement was far more homogeneous than PCK, suggesting that pedagogical capacity develops more unevenly across pre-service teachers, consistent with reviews describing PCK as the most context-dependent and hardest-to-cultivate component [8], [9].

The pattern of correlations indicates that stronger mastery in particular content dimensions tends to co-occur with the corresponding pedagogical capacity, yet content mastery accounts for only part of the variation in PCK, consistent with evidence that pre-service teachers in developing contexts often show comparatively weaker PCK because curricula emphasize content coverage over pedagogical transformation [1], [10]. Rather than forming automatically from content knowledge, PCK appears to require explicit, dedicated learning experiences. The profile of Student 1 is consistent with this interpretation but is treated here only as an illustrative example, not as evidence in its own right.

These results are broadly consistent with prior Indonesian studies reporting that the professional knowledge of mathematics pre-service teachers is generally moderate and uneven, with PCK frequently emerging as the more fragile component [10], [11]. The finding that PCK was marginally higher than CK is noteworthy and can be interpreted cautiously: the PCK items, which draw on familiar classroom situations, may be more accessible to students than the CK items, which demand advanced and abstract mathematical reasoning; moreover, the far larger dispersion of PCK (SD = 23.96) indicates that this apparent advantage is unevenly distributed across the cohort. The particularly low

1

SCK (36.91%) is plausibly linked to teacher-education curricula that emphasize procedural and elementary content over advanced mathematical reasoning—such as formal proof, mathematical modeling, and complex problem solving—so that the more cognitively demanding indicators (SCK4–SCK7) remain underdeveloped [24], [25].

Theoretically, the study offers a specific insight rather than a mere application of existing models: in a low-proficiency, underrepresented setting, CK and PCK remain strongly interdependent even though neither reaches mastery and PCK can slightly exceed CK. This qualifies the intuitive assumption that CK necessarily leads PCK and suggests that, at early proficiency levels, the two constructs may develop in tandem rather than strictly sequentially—extending the MKT and COACTIV frameworks to conditions from which they were not originally derived [6], [25], [27]. Methodologically, the study contributes a CBT model that jointly diagnoses CK and PCK down to the indicator level, addressing the scarcity of instruments that do so simultaneously [22], [28]. In policy terms, the finding that no student reached the high category provides empirical grounds for data-driven rather than assumption-driven curriculum reform [5], [23], a concern underscored by Indonesia's recent international assessment results [3].

Positioned this way, the CBT functions not merely as an examination tool but as a diagnostic-assessment ecosystem that yields rich, decomposable cognitive data actionable for program managers [23], [30]. Several limitations should nonetheless temper these conclusions and caution against generalizing beyond the single institution and the 30 participants studied here: (a) the small sample size, which limits statistical power and generalizability; (b) the total-purposive, single-institution sampling; (c) the bivariate scope, which does not model other determinants simultaneously; (d) the reliance solely on quantitative CBT data, without syllabus and document analysis, classroom observation, or interviews that could triangulate and explain why students underperformed; (e) the limited external evidence of the instrument's validity and reliability beyond the internal analyses reported here; and (f) the multiple-choice format, which cannot fully capture the reasoning that open-ended or performance-based tasks would reveal. In practical terms, these findings imply that teacher-education curricula should allocate explicit, assessed learning time to advanced specialized content (formal proof, mathematical modeling, and complex problem solving) and to knowledge of students and curriculum. They should embed diagnostic CBT within course cycles so that indicator-level weaknesses are identified early and addressed through targeted, evidence-based remediation. Future research should employ larger multi-institution and longitudinal samples, mixed-methods designs that add qualitative triangulation, and multivariate models to strengthen internal and external validity.

CONCLUSION AND RECOMMENDATIONS

23 2

This study set out to measure and profile the CK and PCK of mathematics pre-service teachers through a diagnostic CBT. On average, both constructs remained below the mastery threshold, with the majority of students in the low category and none in the high category; the weakest areas were advanced specialized content within CK and knowledge of curriculum within PCK, while CK and PCK were strongly and positively associated. Theoretically, the study clarifies that CK and PCK can remain interdependent and PCK can slightly exceed CK, even at low overall proficiency; methodologically, it demonstrates a

CBT capable of layered, indicator-level diagnosis of both constructs; and practically, it provides a standardized, efficient means of generating actionable competence profiles for teacher education. These conclusions should be read in light of the study's limitations, namely a small, single-institution, purposively selected sample analyzed with bivariate, multiple-choice-based methods. Overall, the central message is that diagnostic computer-based assessment can meaningfully advance the monitoring and development of mathematics pre-service teachers' professional knowledge, provided its profiles are used to inform evidence-based curriculum and instructional reform.

Recommendations. The following recommendations are offered as literature-informed, potentially effective alternatives suggested by, rather than directly proven by, the present findings, and they should be validated in future studies.

For lecturers. Remedial instruction may target not only the weakest dimensions (SCK within CK and KCC within PCK) but also the weakest indicators identified in Table 5. In the CK domain, priority may be given to strengthening mathematical proof, proving properties or theorems (SCK4) and applying proof strategies (SCK5) through scaffolded proof practice, analysis of proof errors, and habituation to deductive argumentation, followed by mathematical modeling (SCK6) and complex problem solving (SCK7) via problem-based learning. In the PCK domain, the focus may fall on analyzing the strengths and weaknesses of representations (KCT4), applying standards and curricula (KCC2), and designing measurement instruments (KCC3) through case-based learning, analysis of teaching videos, and curriculum-document study.

For teacher education institution curriculum managers. The results encourage integrating a diagnostic CBT, capable of mapping achievement down to the indicator and individual levels, into the study program's quality-assurance cycle, and rebalancing the curriculum so that advanced mathematical reasoning and pedagogical-content competence receive as much emphasis as procedural content mastery.

For future researchers. Future studies may use larger, multi-institution, and longitudinal samples, mixed-methods designs that add qualitative triangulation (syllabus analysis, classroom observation, and interviews), and multivariate models involving additional determinants (such as academic background, learning motivation, and quality of teaching) to strengthen external validity and to explain the causes of the observed weaknesses. A further agenda is to develop and empirically test enhancements to the CBT itself, automatic item-level feedback with worked explanations, direct links to remedial materials for each weak indicator, and adaptive item delivery. So that, once validated, the system can function not only as a diagnostic tool but also as an integrated learning-support environment.

ACKNOWLEDGMENTS

The authors thank the Mathematics Education study program of the teacher education institution for facilitating the implementation of the research, as well as all the students who participated as respondents.

REFERENCES

[1] X. Yang and G. Kaiser, "The impact of mathematics teachers' professional competence on instructional quality and students' mathematics learning outcomes," *Curr. Opin. Behav. Sci.*, vol. 48, p. 101225, Dec. 2022, doi: 10.1016/j.cobeha.2022.101225.

[2] S. Krauss *et al.*, “Pedagogical content knowledge and content knowledge of secondary mathematics teachers,” *J. Educ. Psychol.*, vol. 100, no. 3, pp. 716–725, Aug. 2008, doi: 10.1037/0022-0663.100.3.716.

[3] OECD, *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*. In PISA. OECD Publishing, 2023. doi: 10.1787/53f23881-en.

[4] International Association for the Evaluation of Educational Achievement, *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M).* Amsterdam: IEA, 2012.

[5] H. Qian and P. Youngs, “The effect of teacher education programs on future elementary mathematics teachers’ knowledge: a five-country analysis using TEDS-M data,” *J. Math. Teach. Educ.*, vol. 19, no. 4, pp. 371–396, Aug. 2016, doi: 10.1007/s10857-014-9297-0.

[6] D. Loewenberg Ball, M. H. Thames, and G. Phelps, “Content Knowledge for Teaching: What Makes It Special?,” *J. Teach. Educ.*, vol. 59, no. 5, pp. 389–407, Nov. 2008, doi: 10.1177/0022487108324554.

[7] L. S. Shulman, “Those Who Understand: Knowledge Growth in Teaching,” *Educ. Res.*, vol. 15, no. 2, pp. 4–14, Feb. 1986, doi: 10.3102/0013189X015002004.

[8] M. Grigaliūnienė, E. Lehtinen, L. Verschaffel, and F. Depaepe, “Systematic Review of Research on Pedagogical Content Knowledge in Mathematics: Insights from a Topic-Specific Approach,” *ZDM – Math. Educ.*, vol. 57, no. 4, pp. 777–794, Aug. 2025, doi: 10.1007/s11858-025-01684-1.

[9] F. Depaepe, L. Verschaffel, and G. Kelchtermans, “Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research,” *Teach. Teach. Educ.*, vol. 34, pp. 12–25, Aug. 2013, doi: 10.1016/j.tate.2013.03.001.

[10] S. Y. Ningsih, Turmudi, and D. Juandi, “Pedagogical content knowledge (PCK) profile of prospective teachers in mathematics learning,” *J. Phys. Conf. Ser.*, vol. 1521, no. 3, p. 032057, Apr. 2020, doi: 10.1088/1742-6596/1521/3/032057.

[11] S. Fiangga, S. Khabibah, S. M. Amin, and R. Ekawati, “A Learning Design Analysis of the Pre-service Teachers’ Mathematics Pedagogical Content Knowledge,” *J. Phys. Conf. Ser.*, vol. 1899, no. 1, p. 012153, May 2021, doi: 10.1088/1742-6596/1899/1/012153.

[12] X. Yang, B. Schwarz, and I. K. C. Leung, “Pre-service mathematics teachers’ professional modeling competencies: a comparative study between Germany, Mainland China, and Hong Kong,” *Educ. Stud. Math.*, vol. 109, no. 2, pp. 409–429, Feb. 2022, doi: 10.1007/s10649-021-10064-x.

[13] J. Njiku, “Pre-service teachers manifested mathematics pedagogical content knowledge: The role of the teaching practicum,” *Pedagog. Res.*, vol. 10, no. 1, p. em0229, Jan. 2025, doi: 10.29333/pr/15647.

[14] A. J. Boevé, R. R. Meijer, C. J. Albers, Y. Beetsma, and R. J. Bosker, “Introducing Computer-Based Testing in High-Stakes Exams in Higher Education: Results of a Field Experiment,” *PLOS ONE*, vol. 10, no. 12, p. e0143616, Dec. 2015, doi: 10.1371/journal.pone.0143616.

[15] C. Y. Piaw, “Replacing Paper-based Testing with Computer-based Testing in Assessment: Are we Doing Wrong?,” *Procedia - Soc. Behav. Sci.*, vol. 64, pp. 655–664, Nov. 2012, doi: 10.1016/j.sbspro.2012.11.077.

[16] A. Berry, P. Friedrichsen, and J. Loughran, Eds., *Re-examining pedagogical content knowledge in science education*. In teaching and learning in science series. London, New York: Routledge, 2015. doi: 10.4324/9781315735665.

[17] M. Schmid, E. Brianza, and D. Petko, “Self-reported technological pedagogical content knowledge (TPACK) of pre-service teachers in relation to digital technology use in lesson plans,” *Comput. Hum. Behav.*, vol. 115, p. 106586, Feb. 2021, doi: 10.1016/j.chb.2020.106586.

[18] L. Maas, M. J. S. Brinkhuis, L. Kester, and L. Wijngaards-de Meij, “Cognitive Diagnostic Assessment in University Statistics Education: Valid and Reliable Skill Measurement for Actionable Feedback Using Learning Dashboards,” *Appl. Sci.*, vol. 12, no. 10, p. 4809, May 2022, doi: 10.3390/app12104809.

[19] A. Barana, M. Marchisio, and M. Sacchet, “Interactive Feedback for Learning Mathematics in a Digital Learning Environment,” *Educ. Sci.*, vol. 11, no. 6, p. 279, Jun. 2021, doi: 10.3390/educsci11060279.

[20] S. Rezat, “How automated feedback from a digital mathematics textbook affects primary students’ conceptual development: two case studies,” *ZDM – Math. Educ.*, vol. 53, no. 6, pp. 1433–1445, Nov. 2021, doi: 10.1007/s11858-021-01263-0.

[21] L. N. Yaqin, L. D. Prasojo, N. A. Haji-Othman, N. Yusof, and A. Habibi, “Addressing the Digital Divide in Indonesian Higher Education: Insights, Implications, and Potential Solutions,” in *From Digital Divide to Digital Inclusion*, Ł. Tomczyk, F. D. Guillén-Gámez, J. Ruiz-Palmero, and A. Habibi, Eds., in Lecture Notes in Educational Technology. Singapore: Springer Nature Singapore, 2023, pp. 291–307. doi: 10.1007/978-981-99-7645-4_13.

[22] Naila Rajiha and S. Rini, “Computer-Based Standardized Testing (CBT) Summative Assessment: One Institution Experience,” *JISAE J. Indones. Stud. Assess. Eval.*, vol. 9, no. 1, pp. 1–7, Jan. 2023, doi: 10.21009/jisae.v9i1.33233.

[23] G. Kaiser and J. König, “Competence Measurement in (Mathematics) Teacher Education and Beyond: Implications for Policy,” *High. Educ. Policy*, vol. 32, no. 4, pp. 597–615, Dec. 2019, doi: 10.1057/s41307-019-00139-z.

[24] T. Kleickmann *et al.*, “Teachers’ Content Knowledge and Pedagogical Content Knowledge: The Role of Structural Differences in Teacher Education,” *J. Teach. Educ.*, vol. 64, no. 1, pp. 90–106, Jan. 2013, doi: 10.1177/0022487112460398.

[25] J. Baumert and M. Kunter, “The COACTIV Model of Teachers’ Professional Competence,” in *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers*, M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, and M. Neubrand, Eds., Boston, MA: Springer US, 2013, pp. 25–48. doi: 10.1007/978-1-4614-5149-5_2.

[26] B. Rowan, R. Correnti, and R. J. Miller, “What Large-Scale, Survey Research Tells Us about Teacher Effects on Student Achievement: Insights from the Prospects Study of Elementary Schools,” *Teach. Coll. Rec. Voice Scholarsh. Educ.*, vol. 104, no. 8, pp. 1525–1567, Aug. 2002, doi: 10.1111/1467-9620.00212.

[27] S. Blömeke, J.-E. Gustafsson, and R. J. Shavelson, “Beyond Dichotomies: Competence Viewed as a Continuum,” *Z. Für Psychol.*, vol. 223, no. 1, pp. 3–13, Jan. 2015, doi: 10.1027/2151-2604/a000194.

[28] E. Mohammadpour and Y. Maroofi, “A performance-based test to measure teachers’ mathematics and science content and pedagogical knowledge,” *Heliyon*, vol. 9, no. 3, p. e13932, Mar. 2023, doi: 10.1016/j.heliyon.2023.e13932.

[29] T. Clark and H. Endres, “Computer-based diagnostic assessment of high school students’ grammar skills with automated feedback – an international trial,” *Assess. Educ. Princ. Policy Pract.*, vol. 28, no. 5–6, pp. 602–632, Nov. 2021, doi: 10.1080/0969594X.2021.1970513.

[30] G. A. Nortvedt and N. Buchholtz, “Assessment in mathematics education: responding to issues regarding methodology, policy, and equity,” *ZDM*, vol. 50, no. 4, pp. 555–570, Aug. 2018, doi: 10.1007/s11858-018-0963-z.