**Riemann:** *Research of Mathematics and Mathematics Education*
Volume 8, Issue 2, 483-507, August 2026
e-ISSN: 2721-883X | p-ISSN: 2721-8848
https://journal.sanagustin.ac.id/index.php/reimann

# Analyzing Student Misconceptions in Ratio, Proportion, and Percentage in Pharmaceutical Mathematics Using a Modified Three-Tier Diagnostic Test

**Farida Tuahuns*[1], Alifa Sabrina[2], Fitri Savitri[3]**
[1]Undergraduate Pharmacy Study Program, IKIFA College of Health Sciences, Jakarta, Indonesia
[2,3]Diploma III Pharmacy Study Program, IKIFA College of Health Sciences, Jakarta, Indonesia

Email address:
faridatuahuns@gmail.com*, alifasabrina@ikifa.ac.id, fitrisavitri@ikifa.ac.id

*Corresponding author

Received: Mei, 2026 | Revised: June 2026 | Accepted: July 2026 | Published: August 2026

***Abstract*:**
Pharmaceutical calculations require a solid understanding of ratios, proportions, and percentages to ensure accurate drug preparation and dosage determination. However, pharmacy students often struggle to apply these concepts, which may lead to misconceptions. This study aimed to develop a modified three-tier diagnostic test (MTTDT) to identify pharmacy students' misconceptions in ratio, proportion, and percentage calculations. The instrument was developed using the ADDIE model. Unlike conventional three-tier tests, the second tier was modified from a multiple-choice to an open-ended essay format, enabling separate assessment of conceptual understanding, problem-solving process, and calculation accuracy. Content validity was evaluated by seven experts using Aiken's V, while empirical validity and reliability were assessed in a separate validation cohort using point-biserial correlation and Cronbach's alpha ($\alpha = 0.727$). Of the ten items initially designed, nine met the empirical validity criterion and were retained for the main study, which was administered to 97 first-year pharmacy students. On average, 20.44% of responses reflected misconceptions and 25.44% reflected a lack of knowledge. Analysis of the open-ended second-tier responses revealed a marked disparity between conceptual understanding (81.56%) and calculation accuracy (43.00%), although most students recognized relevant concepts; their performance declined substantially during the problem-solving and unit-conversion steps. These findings suggest that the modified three-tier diagnostic test, particularly its open-ended second tier, can effectively surface procedural and unit-conversion-related misconceptions that conventional multiple-choice formats may fail to detect, supporting its use for targeted instructional intervention in pharmaceutical mathematics education.

**Keywords:** misconceptions, pharmaceutical mathematics, three-tier diagnostic test, ratio and proportion, pharmacy education

## Introduction

Mathematics in pharmacy is applied to drug development and optimizing the dosage (Chamseddine & Rejniak, 2020; Woodcock & Woosley, 2008), concentration, dilution, and converting between measurement systems (Sharma & Dunham, 2025). These calculations are essential in clinical contexts such as determining individualized drug doses, preparing intravenous solutions, and adjusting medication regimens based on patient-specific parameters. Because of its importance, mathematics is a required subject in the pharmacy curriculum (Alrabiah et al., 2022). To guarantee patient safety, pharmacists must possess strong mathematical proficiency. Calculation competence is evaluated in the Mathematical Competence section of professional licensure examinations (National Association of Boards of Pharmacy, 2024). In the Indonesian pharmacy education system, mathematics is one of the basic science subjects studied. In the undergraduate pharmacy curriculum, mathematics, along with physics, general chemistry, organic chemistry, physical chemistry, and analytical chemistry, accounts for 10–15% of the total credits (Commission for Educational Development, 2013) to support the learning outcomes of pharmacy graduates (Association of Indonesian Pharmacy Higher Education (APTFI), 2024). At the IKIFA College of Health Science (STIKes IKIFA), the mathematics course is titled Pharmaceutical Mathematics.

For healthcare workers, pharmaceutical calculations are a fundamental requirement (Aronson et al., 2022; Mnatzaganian et al., 2020), such as for the proper dosage, preparation, and administration of pharmaceuticals (Fredrickson et al., 2025); yet some research has shown a lack of proficiency in this area among pharmacy (Crawford et al., 2024) and nursing (Eastwood et al., 2011; McMullan et al., 2010; Wennberg-Capellades et al., 2022). This lack of competence may compromise patient safety, increase the risk of medication errors, and reduce the efficiency of healthcare delivery systems (Jairoun et al., 2024). Most calculations in pharmacy involve only basic mathematics, but some students still experience difficulties (Mnatzaganian et al., 2020; Prescott et al., 2024), for example, with ratios (Hegener, 2021). Several studies show that the composition of ratio and proportion applications in pharmaceutical problems is quite large (Arkell & Rutter, 2012; Hegener et al., 2013). In addition, the ratio and proportion approach, together with dimensional analysis, is the most widely used by students in solving pharmaceutical problems (Klausner & Holt, 2022). Topics related to proportion, fractions, decimals, and percentages are divided in the material, with the largest composition for assessing students’ baseline mathematics assessment (Frankart et al., 2025). Therefore, a solid conceptual understanding of ratio and proportion is essential for pharmacy students, as weaknesses in these foundational topics may directly affect their ability to perform pharmaceutical calculations accurately.

Unfortunately, research on students' understanding of ratios and proportions has largely been limited to pre-university education. However, existing studies rarely focus on pharmacy students and do not specifically examine misconceptions in ratio and proportion within pharmaceutical contexts. Research results show that students' understanding of ratios and proportions is still lacking (Johar et al., 2024; Muttaqin et al., 2017; Nofriati et al., 2020; Sari et al., 2024, 2025b, 2025a; Sukasno et al., 2024; Wahyuningrum et al., 2022). This lack of understanding often stems from underlying misconceptions. Misconceptions can be defined as an individual's error in comprehending a notion that does not align with the opinions of a scientific expert (Wahyudi et al., 2021). When students improperly generalize a notion, it impairs their capacity to interpret newly created information, leading to misconceptions (Mohyuddin & Khalil, 2016). Identifying misconceptions early is essential for designing targeted instructional interventions that improve students' conceptual understanding and calculation competence. Various methods, such as diagnostic tests and interviews, can be used to identify misconceptions (Chen, 2009; Gurel et al., 2017; Novak, 1990).

Diagnostic tests can be used to identify each student's limitations, strengths, and weaknesses, as well as to uncover common misconceptions among students. Various forms of diagnostic tests with different levels have been developed, such as two-, three-, and four-level multiple-choice diagnostic tests. Even today, multi-level instruments have been developed for chemistry and science education (Habiddin et al., 2021). In this study, a three-level diagnostic test was used to evaluate students' misconceptions, where the second level did not use multiple-choice questions but rather short descriptions, so that students not only selected the answer they considered correct at the first level, but also briefly explained how they arrived at that answer. This makes the use of three tiers more effective because it does not require questions to measure confidence in the answer at the second tier (the fourth-tier question in a four-tier diagnostic test) (Rusilowati, 2015). Additionally, the more tiers used, the more choices students must make for a single question. This implies that the number of questions available to students will be limited.

The three-tier diagnostic test has previously been applied in the field of pharmacy to identify students' misconceptions regarding biochemistry concepts (Milenković et al., 2016). In addition, the three-tier diagnostic test, combined with the Certainty of Response Index, has been used to identify misconceptions among prospective elementary school teachers on the topics of integers and fractions (Nurwahida & Munir, 2022). However, to the best of the authors' knowledge, no study has yet applied a three-tier diagnostic test to identify pharmacy students' misconceptions specifically regarding ratios, proportions, and percentages despite their central role in pharmaceutical calculations.

Misconceptions in mathematics can stem from two different sources: an incomplete understanding of the underlying mathematical concepts (conceptual knowledge), or a correct conceptual understanding accompanied by errors in the procedural application of those concepts (procedural knowledge) (Hiebert & Lefevre, 1986). Basic mathematical calculations applied in pharmaceutical calculations (such as ratio, proportion, and particularly unit conversion in dosage calculations) are prone to the same types of misunderstandings, and errors of this kind may plausibly affect medication dosing accuracy. The conventional multiple-choice format used in second-level diagnostic tests cannot distinguish between these two issues, as students may select the correct answer without being able to perform the related calculations accurately, or conversely, may perform the calculation correctly while still holding an incorrect conceptual understanding of the underlying concept. Therefore, it is necessary to modify the second-level test to an open-ended essay format, which provides a more accurate assessment of conceptual understanding, problem-solving process, and calculation accuracy as separate components.

The three-tier diagnostic test in this study was developed using the ADDIE model, which consists of Analysis, Design, Development, Implementation, and Evaluation phases (Branch, 2009). The ADDIE model was used because of its simplicity and worldwide use in education development (Siahaan, 2025), such as in textbooks (Savitri et al., 2025) or instruments (Maxnun et al., 2024; Ramdhani et al., 2025). An analytical rubric is used to reduce subjectivity in the evaluation of the second tier (Cahyadi et al., 2023; Jonan, 2020; Nurhaifa et al., 2020). This study contributes to the limited research on misconceptions among pharmacy students. Therefore, this study aims to develop a modified three-tier diagnostic test to identify pharmacy students' misconceptions regarding ratios, proportions, and percentages.

## Research Methods

This research was conducted at IKIFA College of Health Sciences. The population comprised all active first-year Diploma III (D3) and undergraduate (S1) pharmacy students who had received material on ratio, proportion, and percentage. A total of 97 respondents were sampled using a total sampling technique. Data were collected through a modified three-tier diagnostic test administered directly to students at the end of the course within a specified timeframe. The instrument was adapted and developed from previous studies and validated before use. Ethical approval for this study was granted by the Ethics Committee of IKIFA College of Health Sciences (STIKes IKIFA) (No:005452/Komite Etik Penelitian STIKes IKIFA/2025).

### *Instrument Development: ADDIE Model*

This study used an educational Research and Development (R&D) design, combining systematic instrument development with validation and descriptive quantitative analysis. Instrument development followed the ADDIE model (Analysis, Design, Development, Implementation, Evaluation) (Branch, 2009).

1. **Analyze Stage**

   At this stage, an analysis was conducted of the multiple-choice misconception assessment instruments for two-level, three-level, and four-level diagnostic tests. A review of the curriculum from the APTFI and the Commission for Educational Development was also conducted, along with an analysis of common errors made by students in pharmaceutical calculations, to serve as a guide in the development of test items. The outputs from this stage are a series of priority topics (ratios, proportions, and percentages in dosages, dilutions, concentrations, and unit conversions) and common error patterns.

2. **Design Stage**

At this stage, a blueprint for assessment instruments was developed for the topics of ratios, proportions, and percentages in the context of pharmaceutical calculations, such as dose and concentration calculations, solution dilution, and unit conversion. Ten questions were designed, with each question consisting of three stages: first, second, and third tiers.

a. **The first tier**

Consists of multiple-choice questions related to the application of mathematical concepts in the pharmaceutical field, with five answer options, four of which are distractors.

b. **The second tier**

This section contains the answers to the questions in the first tier, which in the previous diagnostic test were in multiple-choice format (closed-ended reasoning). However, in this study, the format was modified to include an answer that describes the method for solving the question in the first tier (open-ended reasoning).

c. **The third tier**

The third part of each item is a confidence level, which investigates whether or not they are confident about their responses.

Additionally, a rubric has been designed for the second level of assessment; this rubric is used solely to determine whether students' answers are correct or incorrect, not as a basis for assigning scores. The second tier involves open-ended reasoning and is divided into three interconnected stages:

a. **Conceptual Understanding**

At this stage, students can understand the problem being asked by rewriting all available information as known and requested information. If a student writes only partial information or makes an error in writing it, they are categorized as incorrect at this stage.

b. **Problem-solving process**

At this stage, students can write the problem-solving process correctly and can also use the correct formula. If a student does not write the information in the form of known and requested information (conceptual understanding stage), but can solve the problem using the correct formula, they are categorized as correct for conceptual understanding and the problem-solving process.

c. **Calculation results**

Students can write the calculation results correctly, including the correct units. If a student does not complete the calculation or makes an error in converting units, they are categorized as incorrect at this stage.

There is a distinction between omitting known or requested information entirely and writing it down but with errors. Some students immediately move on to the problem-solving stage without formally listing the known and requested information; this is considered a matter of writing style (the student's preferred approach to answering), not evidence of a conceptual misunderstanding, as long as the subsequent problem-solving process is correct. Conversely, students' attempts to write down this information, even if accompanied by errors, are considered an indication of a genuine gap in conceptual understanding. A student's answer is categorized as correct at the second tier if they can answer all three stages correctly. The output of this stage was a ten-item instrument blueprint accompanied by a three-component scoring rubric for the second tier.

3. **Development Stage**

Content validity was evaluated by a panel of seven experts: one lecturer with a background in physics education who teaches Pharmaceutical Physics (co-author), one lecturer with a background in mathematics education who teaches Pharmaceutical Mathematics (co-author), one lecturer with a background in microbiology who teaches Microbiology, and four pharmacist (apoteker) lecturers whose teaching responsibilities cover chemistry and pharmacy coursework. The corresponding author, who developed the test items, was not part of this validation panel, thereby avoiding direct self-assessment of item content. Content validity was assessed using Aiken's V, interpreted according to established criteria ($V \geq 0.80$ = strong; 0.70–0.79 = acceptable with minor revision required; $<0.70$ = poor) (Tajuddin et al., 2025). In addition to assessing the relevance of test items using Aiken's V, the panel of experts also independently verified the accuracy of the answer key by recalculating each test item using their own solution methods; all of these recalculation results matched the established answer key, thereby confirming the accuracy of the answer key prior to empirical testing.

Following the content validity assessment, empirical validity and reliability testing were conducted on a separate group consisting of 60 upper-level pharmacy students at STIKes IKIFA who had already completed the relevant coursework. This group was distinct from both the practicality pilot group and the main study sample. This separate group was necessary because the main study used a total sampling technique, in which all first-year pharmacy students ($n = 97$) were designated as the target sample. Point-biserial correlation was used to assess empirical validity, while Cronbach's alpha was used to assess reliability, with a minimum acceptable value of 0.70. The output of this stage was a content and empirically validated instrument, informed jointly by expert judgment and item level statistical testing.

4. **Implementation Stage**

   The instrument was piloted on two lecturers and ten students. Feedback from the pilot test was used to revise the items. The output of this stage was a revised instrument informed by pilot-test feedback.

5. **Evaluation Stage**

   Item analysis included difficulty index, discrimination index, and distractor effectiveness. Item difficulty was classified into five levels, ranging from very difficult (0.00–0.19) to very easy (0.80–1.00). Discrimination indices were categorized from poor (≤0.19) to very good (≥0.40), with intermediate levels of fair (0.20–0.29) and good (0.30–0.39) (Istyastono, 2016). Distractor effectiveness was evaluated using the distractor index, which ranged from extremely poor (>200%) to very good (76–125%) (Suryadin, 2024). The output of this stage was the item difficulty, discrimination, and distractor-effectiveness indices used to finalize item selection for the main study, along with the analysis of students' misconceptions.

***Misconception Analysis Framework***

Student responses were categorized based on correctness, reasoning, and confidence levels using the framework proposed by Arslan et al. (2012), as shown in Table 1. The second tier was modified into an open-ended format to capture students' reasoning processes more comprehensively.

**Table 1.** Three-Tier Diagnostic Test Category

| First tier | Second tier | Third tier | Categories |
|---|---|---|---|
| Correct | Correct | Certain | Scientific knowledge |
| Correct | Incorrect | Certain | Misconception (false positive) |
| Incorrect | Correct | Certain | Misconception (false negative) |
| Incorrect | Incorrect | Certain | Misconception |
| Correct | Correct | Uncertain | Lucky guess, lack of confidence |
| Correct | Incorrect | Uncertain | Lack of knowledge |
| Incorrect | Correct | Uncertain | Lack of knowledge |
| Incorrect | Incorrect | Uncertain | Lack of knowledge |

The percentage of misconceptions is calculated for each category in each question to determine the most dominant concept experiencing misconceptions. The answers for Level 1 (multiple-choice accuracy) and Level 2 (essay accuracy) were graded separately by two graders (the corresponding author and a co-author) based on the answer key, respectively, before reconciliation took place. Inter-rater agreement was quantified using Cohen's kappa (McHugh, 2012). Any differences in classification (conceptual understanding, problem-solving process, and calculation results) at the second tier were discussed until a consensus was reached based on the rubric. The classification of essay responses into three subcomponents was performed by the corresponding author. When the classification was uncertain, the co-author was consulted before a final decision was made by the corresponding author. This singlerater procedure with consultation for subcomponent classification is acknowledged as a limitation.

## Results and Discussions

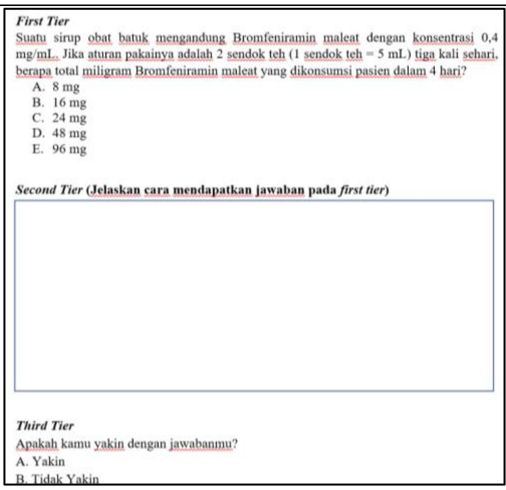
### *Development of the Modified Three-Tier Diagnostic Test (MTTDT)*

The MTTDT instrument includes changes to the measurement objectives, response format, reasoning construction, assessment techniques, validation approaches, and cognitive levels as shown in Table 2.

**Table 2.** Comparison of the Standard and Modified Three-Tier Diagnostic Test Instruments

| Dimension | Standard TTDT | Modified TTDT (This study) | Type of modification |
|---|---|---|---|
| **Measurement Objective** | Identifying misconceptions | Identifying misconceptions and locating where misconceptions occur in the problem-solving process | Modification of measurement objective |
| **Second-tier format** | Multiple-choice reasoning (closed-ended) | Open-ended reasoning in essay form | Modification of response format |
| **Reasoning Construct** | Understanding mathematical concepts | Applying mathematical concepts in the pharmaceutical calculations context | Conceptual modification |
| **Scoring Technique in the second tier** | Dichotomous scoring (correct/incorrect) | Analytical rubric for evaluating conceptual understanding, problem-solving process, and calculation results | Methodological modification |
| **Validation Technique** | Content validity via expert review | Content validity through expert review of both answer and open-ended reasoning tiers, with empirical validity analysis conducted on first-tier responses | Modification of the validation approach |
| **Cognitive Level** | Reasoning options in Bloom's C2–C3 | Reasoning tasks in Bloom's C3–C5 require procedural and quantitative reasoning | Cognitive-level modification |

These modifications indicate that the instrument not only identifies students' misconceptions but also captures their reasoning processes and locates errors within the problem-solving stages. In particular, the use of open-ended responses in the second tier and an analytical rubric enables a more comprehensive assessment of students' conceptual understanding, problem-solving process, and calculation accuracy. Following these modifications, the instrument was constructed in the form of ten items representing various pharmaceutical contexts. These items were designed to reflect real-world calculation scenarios encountered in pharmacy practice. An example of the modified three-tier diagnostic test item is presented in Figure 1.



***First Tier***

Suatu sirup obat batuk mengandung Bromfeniramin maleat dengan konsentrasi 0,4 mg/mL. Jika aturan pakainya adalah 2 sendok teh (1 sendok teh = 5 mL) tiga kali sehari, berapa total miligram Bromfeniramin maleat yang dikonsumsi pasien dalam 4 hari?

A. 8 mg
B. 16 mg
C. 24 mg
D. 48 mg
E. 96 mg

***Second Tier* (Jelaskan cara mendapatkan jawaban pada *first tier*)**

***Third Tier***

Apakah kamu yakin dengan jawabanmu?
A. Yakin
B. Tidak Yakin

**First tier**

A cough medicine contains brompheniramine maleate with a concentration of 0.4 mg/mL. If the dosage instruction is 2 teaspoons (1 teaspoon = 5 mL) three times a day, how many milligrams of brompheniramine maleate are consumed by the patient in 4 days?

**Second tier**

Explain how you obtained your answer in the first tier.

**Third tier**

Are you confident in your answer?

A. Yes
B. No

**Figure 1.** Example of the modified three-tier diagnostic test item.

***Content Validity (Aiken's V)***

After the instrument was constructed, its quality was evaluated through content validity assessment by expert reviewers. The results of the content validity analysis using Aiken's V coefficient are presented in Table 3.

**Table 3.** Content Validity Results (Aiken's V)

| Items | Aiken's V | Interpretation |
|---|---|---|
| **1** | 0.905 | Strong content validity |
| **2** | 0.857 | Strong content validity |
| **3** | 0.810 | Strong content validity |
| **4** | 0.810 | Strong content validity |
| **5** | 0.905 | Strong content validity |
| **6** | 0.762 | Acceptable, minor revision required |
| **7** | 0.857 | Strong content validity |
| **8** | 0.810 | Strong content validity |

| Items | Aiken's V | Interpretation |
|---|---|---|
| **9** | 0.857 | Strong content validity |
| **10** | 0.762 | Acceptable, minor revision required |

As shown in Table 3, eight of the ten items achieved V ≥ 0.80 (strong content validity), while the remaining two items (Items 6 and 10) achieved V = 0.762 (acceptable, minor revision required). Expert feedback and the corresponding revisions are summarised in Table 4.

**Table 4.** Expert Feedback

| Subject area | Item numbers | Expert suggestions | Revision |
|---|---|---|---|
| **Basic pharmaceutics** | 1,2,6,10 | a. Incorrect writing of "S3dd"<br>b. Incorrect abbreviation of units<br>c. The type of antibiotic used in the item was replaced with a more relevant one<br>d. Options should be arranged from largest to smallest or vice versa<br>e. Image needs clarification | a. "S3dd" revised to "S.3dd"<br>b. Unit abbreviation revised from "gr" to "g"<br>c. Antibiotics changed from Penicillin to Cefixime trihydrate<br>d. Options arranged from smallest to largest<br>e. Image re-edited for clarity |
| **Semi-solid and Liquid Pharmaceutical Technology** | 3 | The written dosage volume needs clarification regarding the dosage form container | The dosage volume "60 ml" revised to "60 ml/bottle" |
| **Chemistry** | 4,7,8 | a. Percentage writing should not be in fractional form<br>b. Options should be arranged from largest to smallest or vice versa | a. Writing of 2/5% revised to 0.4%<br>b. Options arranged from smallest to largest |

| Subject area | Item numbers | Expert suggestions | Revision |
|---|---|---|---|
| **Microbiology** | 9 | Image needs clarification | Image re-edited for clarity |
| **Phytochemicals** | 5 | a. Image needs clarification<br>b. Additional explanation is needed during the lecture | a. Data in the image presented in table form for clarity<br>b. Explanation of the yield formula provided during the lecture. |

As previously described, empirical validity and reliability testing (using SPSS version 26) yielded the following results. Nine of the ten items achieved statistically significant point-biserial correlations ($r = 0.390–0.786$, $p < 0.05$), while one item (Item 5) did not ($r = 0.141$, $p = 0.283$) and was therefore excluded from the final instrument. Although Item 5 had strong content validity according to the expert panel (Aiken's V = 0.905, see Table 3), its exclusion following empirical testing was likely due to a ceiling effect (the validation group consisted of upper-level pharmacy students who had already mastered this phytochemistry topic through previous coursework), resulting in a high proportion of correct answers (Mir et al., 2025). This finding highlights the limitations of using upper-level cohorts for empirical item validation when the instrument is intended for a different target population (first-year students), as prior mastery of the material can obscure an item's discriminative power for the intended audience. Reliability testing yielded a Cronbach's alpha value of 0.727 for the remaining nine items.

After conducting empirical validity and reliability testing, the misconception instrument was then tested on a pilot study with 2 lecturers and 10 students, where one of the improvements was related to writing $3/4$ to $\frac{3}{4}$.

***Item Analysis***

Following the finalization of the instrument, an evaluation was conducted to assess its quality. This evaluation included item analysis, such as item difficulty, discrimination index, and distractor effectiveness. The results of the item analysis are presented in Table 5 (there was a change in the order of the questions during the validity testing and data collection phases).

**Table 5.** Summary of Item Analysis

| Item | Difficulty Index | | Discrimination Index | | Distractor Quality |
|---|---|---|---|---|---|
| | Value | Category | Value | Category | |
| **1** | 0.57 | Moderate | 0.620 | Very good | Ineffective |
| **2** | 0.27 | Difficult | 0.603 | Very good | Functional |

| Item | Difficulty Index | | Discrimination Index | | Distractor Quality |
|---|---|---|---|---|---|
| | Value | Category | Value | Category | |
| **3** | 0.49 | Moderate | 0.614 | Very good | Functional |
| **4** | 0.33 | Difficult | 0.757 | Very good | Moderately effective |
| **5** | 0.53 | Moderate | 0.636 | Very good | Moderately effective |
| **6** | 0.43 | Moderate | 0.719 | Very good | Functional |
| **7** | 0.35 | Difficult | 0.620 | Very good | Ineffective |
| **8** | 0.54 | Moderate | 0.701 | Very good | Functional |
| **9** | 0.48 | Moderate | 0.699 | Very good | Ineffective |

As shown in Table 5, several items (Items 2, 3, 6, and 8) contain functional distractors, indicating that the incorrect options were sufficiently plausible to capture students with misconceptions. In contrast, Items 1, 7, and 9 contain ineffective distractors, suggesting that some incorrect options were not sufficiently attractive to serve as diagnostic indicators. Meanwhile, Items 4 and 5 demonstrate moderately effective distractors, reflecting partial effectiveness in identifying students' conceptual errors. This classification is used to determine the extent to which incorrect answer options reveal learners' contextual error patterns (Suprihanto & Sahrul, 2026).

The analysis results show that most are in the difficult category. This distribution indicates the instrument levels of challenge. The discrimination index for all items is in the very good category, indicating that each item discriminates between students with high and low levels of conceptual understanding. Furthermore, the presence of distractors that function in several items indicates that the incorrect answer options are plausible enough to attract students with less precise or incomplete conceptual understanding. Thus, these distractors serve as diagnostic indicators that help reveal patterns of students' conceptual errors. Overall, the results of the item analysis indicate that the developed instrument is appropriate for use in diagnosing misconceptions in pharmaceutical mathematics.

***Analyzing the Misconceptions of Pharmacy Students***

Prior to the categorization process, inter-rater reliability was calculated using Cohen's kappa, with values of Cohen's $\kappa = 0.995$ (99.77% agreement) for the first tier and $\kappa = 0.991$ (99.54% agreement) for the second tier, across all 873 assessment pairs (9 items × 97 students). Table 6 presents the distribution of students' responses for the nine items (n = 97 per item). Scientific Knowledge (SK) accounted for 40.22% of responses on average, which implies that fewer than half of all students showed correct understanding paired with appropriate reasoning and confidence. However, on average, 20.44% of respondents held misconceptions; specifically, 44 respondents (45.4% of the sample) held misconceptions regarding question 4. This phenomenon can be interpreted using the perspective of confidence-based assessment, which suggests that incorrect answers accompanied by high confidence are often associated with misconceptions rather than a simple lack of knowledge (Hasan et al., 1999).

Additionally, false positives and false negatives accounted for 7.67% and 0.22%, respectively. The relatively high proportion of false positives suggests that some students selected correct answers without providing appropriate reasoning, indicating the possibility of guessing.

In addition to false positives, the lucky guess category (correct answer, correct reasoning, but low confidence) accounted for an average of 3.00% of responses. Although this proportion is comparatively small, its presence alongside false positives indicates that correctness alone is an unreliable indicator of genuine understanding: some students arrived at correct answers without full conviction in their reasoning, while others expressed confidence in answers unsupported by valid reasoning. Together, these two categories underscore why a conventional multiple-choice format, which captures only the final answer, would have misclassified a combined 10.67% of responses as unproblematic, when in fact they reflect either guesswork or unwarranted confidence.

In the context of the three-tier diagnostic test used in this study, students' confidence responses in the third tier indicate how strongly they believe in their answers. This finding highlights the importance of incorporating an essay format in the second tier of the MTTDT, as it enables the evaluation of the consistency between answers and reasoning.

Variations in the level of misunderstanding among the items were found to be related to differences in the content of the items and the structure of their calculations. Question 4, which recorded the highest number of answers containing misunderstandings (44 out of 97 students, see Table 6), required only a single unit conversion step (converting the given concentration per 1,000 mL to a quantity of 120 mL, with a final conversion from grams to milligrams). Although its structure is simple, this question was often answered incorrectly by students due to decimal errors or unit conversion mistakes, as illustrated in Figure 2. This suggests that even a simple question involving a single unit conversion step can lead students to become overconfident and careless in their calculations. Conversely, questions requiring several sequential steps, such as Question 9 (batch-scale production calculation), recorded the lowest rate of misunderstanding (9 out of 97 students), tied with Question 5. Unlike Question 9, however, Question 5's low misconception rate cannot be attributed to procedural complexity, as it involved only a straightforward two-step calculation (division followed by percentage conversion); its low error rate may instead reflect the familiarity of percentage-based reasoning to pharmacy students, who encounter this format frequently across coursework.

In contrast, item 2 (see Table 6), which recorded the lowest number of "Scientific Knowledge" responses (26 out of 97 students), required students to work backward from a known result (a measured mass of 1.2 mg) to an unknown initial quantity, rather than following the more common forward approach of calculating a result from a given value. This inverse ratio structure appears to be conceptually more challenging than the forward ratio format used in most other test items. As illustrated in Figure 4, this difficulty is often compounded by incomplete unit conversion: students perform the inverse calculation correctly ($x = 300$mg) but fail to convert the result to the grams specified by the question, indicating that the demand for reverse reasoning and the unit conversion step interact to produce a higher error rate on this item. Item 1, which recorded the highest number of Scientific Knowledge responses (49 out of 97 students), involved a single, direct percentage-to-fraction calculation without reverse reasoning or multi-step sequencing, consistent with its relatively lower difficulty level.

**Table 6.** Analysis of the modified three-tier diagnostic test

| **Items** | **Categories** | | | | | |
|---|---|---|---|---|---|---|
| | **SK** | **FP** | **FN** | **M** | **LG** | **LK** |
| **1** | 49 | 5 | 0 | 10 | 7 | 26 |
| **2** | 26 | 6 | 0 | 29 | 2 | 34 |
| **3** | 44 | 8 | 0 | 15 | 4 | 26 |
| **4** | 30 | 0 | 2 | 44 | 2 | 19 |
| **5** | 47 | 8 | 0 | 9 | 4 | 29 |
| **6** | 39 | 2 | 0 | 30 | 2 | 24 |
| **7** | 31 | 13 | 0 | 20 | 2 | 31 |
| **8** | 48 | 9 | 0 | 18 | 4 | 18 |
| **9** | 48 | 18 | 0 | 9 | 0 | 22 |
| **Average** | 40.22 | 7.67 | 0.22 | 20.44 | 3.00 | 25.44 |

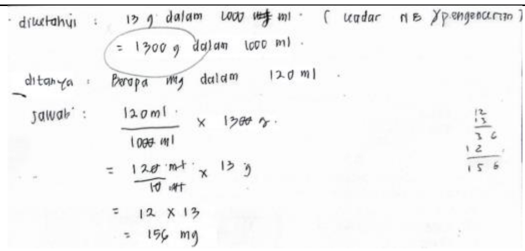
SK = Scientific Knowledge, FP = Misconception (False Positive), FN = Misconception (False Negative), M = Misconception, LG = Lucky Guess, LK = Lack of Knowledge

To provide a clearer understanding of how these misconceptions manifest in students' reasoning (second-tier), particularly for Item 4 as discussed above, an example of a student's response is presented in Figure 2. Figure 2 presents an example of a student's answer to Item 4, where the student used the principle of proportion to determine the amount of nutrient broth needed for 120 mL of solution. In a proportional calculation, all quantities must be expressed using the same units of measurement. The student's answer in the final calculation yielded 156 mg. Using a proportional calculation would have yielded a value of 1.56 grams first, which would then have to be converted to milligrams to match the units required in the problem. In fact, by solving this problem from the beginning, students can convert 13 grams to 13,000 milligrams (mg) to match the final units, so there is no need for further conversion at the end. However, the result of the conversion made by the students is wrong, namely 1,300 milligrams instead of 13,000 milligrams.

Berikut ini adalah etiket di kemasan media pertumbuhan *Nutrient Broth*.


Hitunglah berapa milligram yang dibutuhkan untuk membuat media sebanyak 120mL jika dilabel tertulis suspensikan 13,0g ke dalam 1000mL!

A. 1,56 mg
B. 15,6 mg
C. 156 mg
D. 1.560 mg
E. 15.600 mg




**First tier (Problem)**

How many milligrams of Nutrient Broth are needed to prepare 120 mL of medium if 13.0 g dissolves in 1000 mL?

**Second tier (Student's written reasoning)**

Known: 13 gram in 1000 mL
Asked: Amount of Nutrient Broth in 120 mL
Answer:

$= \frac{120\ mL}{1000\ mL} \times 13g$ ⟶ Use of non-equivalent units

$= \frac{120\ mL}{10\ mL} \times 13g$

$= 12 \times 13$

$= 156\ mg$

**Third tier (Students' confidence)**

Certain

**Figure 2.** Example of uncategorized response in Item 4

The second-tier responses were evaluated using a rubric consisting of three components: conceptual understanding, problem-solving process, and calculation results. A student's answer is categorized as correct at the second tier if they can answer all three stages correctly. Table 7 presents the distribution of correct and incorrect responses for each rubric component in the second tier.

**Table 7.** Distribution of Student Responses in the Second Tier Based on the Analytic Rubric

| No | Conceptual Understanding | | Problem-solving Process | | Calculation Results | |
|---|---|---|---|---|---|---|
| | Correct | Incorrect | Correct | Incorrect | Correct | Incorrect |
| **1** | 90 | 7 | 62 | 35 | 55 | 42 |
| **2** | 77 | 20 | 42 | 55 | 26 | 71 |
| **3** | 80 | 17 | 51 | 46 | 48 | 49 |
| **4** | 73 | 24 | 64 | 33 | 32 | 65 |
| **5** | 84 | 13 | 60 | 37 | 51 | 46 |
| **6** | 84 | 13 | 48 | 49 | 42 | 55 |
| **7** | 75 | 22 | 48 | 49 | 34 | 63 |

| No | Conceptual Understanding | | Problem-solving Process | | Calculation Results | |
|---|---|---|---|---|---|---|
| | Correct | Incorrect | Correct | Incorrect | Correct | Incorrect |
| **8** | 91 | 6 | 70 | 27 | 52 | 45 |
| **9** | 80 | 17 | 54 | 43 | 47 | 50 |
| **Average** | 81.56 | 15.44 | 55.44 | 41.56 | 43.00 | 54.00 |

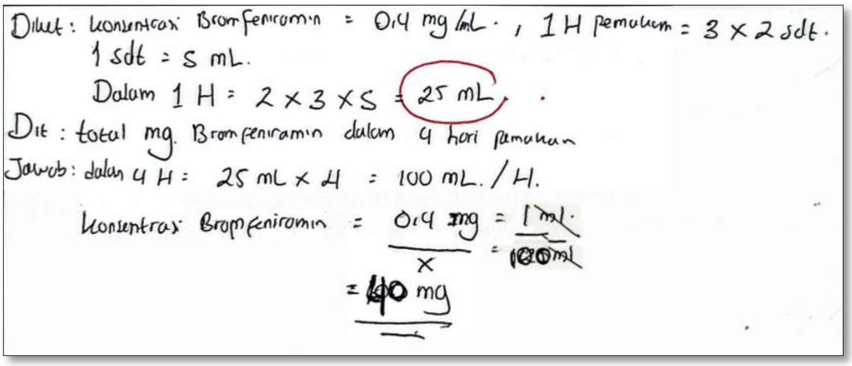
The results show that conceptual understanding is generally higher than procedural accuracy and calculation correctness. The second-tier assessment results reveal a significant disparity between conceptual understanding, problem-solving processes, and the accuracy of calculation outcomes. Although most students were able to recognize relevant concepts and select appropriate formulas, their performance declined at the problem-solving stage and further declined at the calculation stage, particularly when systematic procedures and unit conversions were required.

Key findings show that the average conceptual understanding score (81.56%) was considerably higher than the average calculation accuracy score (43.00%), indicating that students were able to identify the correct formula by rote memorization without fully understanding it, particularly in terms of unit conversion (Winarni et al., 2025). This pattern aligns with the theory proposed by Sweller (1988) regarding cognitive load, where tasks requiring the simultaneous retention and processing of multiple pieces of information in working memory impose a heavier intrinsic cognitive load than tasks requiring only formula recognition.

This finding is also consistent with the study by Wang et al. (2022), which found that as the number of calculation steps required to solve algebraic problems increases, the accuracy rate drops dramatically. Similarly, in this study, pharmaceutical calculations involving dosages, dilutions, and unit conversions require students to maintain and process several interrelated elements (such as known values, units, conversion factors, and the order of operations) that may exceed working memory capacity even when the underlying concepts are understood. This gap may also reflect a lack of procedural automation: Wang et al. (2022) demonstrated that repeated practice significantly reduces the cognitive load on multi-step calculation tasks. Although their finding was based on repeated exposure within a single testing session rather than curriculum-level training, it plausibly extends to suggest that more frequent, applied practice with dosage and dilution calculations, beyond didactic instruction alone, may help reduce procedural errors of this kind.

This pattern is consistent with the high prevalence of misconceptions and false-positive responses reported in Table 6, in which apparently correct answers were not supported by coherent reasoning or consistent results. Overall, the findings demonstrate that students' difficulties in pharmaceutical mathematics are primarily associated with challenges in the integrated application of concepts, procedures, and computational precision, rather than deficiencies in conceptual knowledge, as identified through the modified three-tier diagnostic assessment.

Several findings from the second-tier responses indicate that some students made errors in unit conversion at the final stage, while others misinterpreted the problem, leading to incorrect multiplication (arithmetic errors) during the calculation process (Figure 3). In addition, some students approached the problem using different conceptual strategies compared with their peers but failed to complete the required unit conversion to grams, reflecting incomplete quantitative reasoning (Figure 4).



Dik: konsentrasi Bromfeniramin = 0,4 mg/mL, 1 H pemakaian = 3 × 2 sdt.
1 sdt = 5 mL.
Dalam 1 H = 2 × 3 × 5 = 25 mL.
Dit: total mg Bromfeniramin dalam 4 hari pemakaian
Jawab: dalam 4 H = 25 mL × 4 = 100 mL/H.
Konsentrasi Bromfeniramin = 0,4 mg / x = 1 mL / 100 mL
= 40 mg

**Second tier**
Known: Brompheniramine concentration = 0.4 mg/mL,
1 day of consumption = 3 × 2 teaspoons
1 teaspoon = 5 mL
In 1 day: 2 × 3 × 5 = 25 mL → arithmetic errors
Asked: total mg of brompheniramine in 4 days of consumption
Answer:
In 4 days: $25\ ml \times 4 = 100\ mL$
Brompheniramine concentration: $\frac{0{,}4\ mg}{x} = \frac{1\ mL}{100\ mL} = 40\ mg$

**Figure 3.** Example of a student response showing an arithmetic error in dosage calculation.

The answer in Figure 3 indicates that the student incorrectly calculated the total volume as $2 \times 3 \times 5 = 25\ mL$, resulting in 25 mL instead of the correct value. These findings indicate that the error was not caused by a misconception about the concentration concept (mg/mL), but rather by an arithmetic error in calculating $2 \times 3 \times 5$ as 25 mL. This initial error in representing the problem information propagated throughout the entire problem-solving process. This finding suggests that in pharmaceutical mathematics, mastery of ratio and concentration concepts must be accompanied by fluency in basic arithmetic to prevent dosage errors that may have clinical implications (Aydin & Dinç, 2017; Hijji, 2025).

The response in Figure 4 indicates that the student did not complete the solution because the final result was not converted into grams. The response in Figure 4(a) applied basic multiplication, whereas the response in Figure 4(b) used a proportional reasoning approach. Both responses suggest that students focused primarily on obtaining a numerical answer without adequately considering the measurement units required in the problem. This phenomenon can be explained from the perspective of quantitative reasoning (Thompson, 2022) which emphasizes that quantities represent the coordination between numerical values and meaningful units. Therefore, although students demonstrate procedural competence, their integrated quantitative understanding has not yet been fully developed.

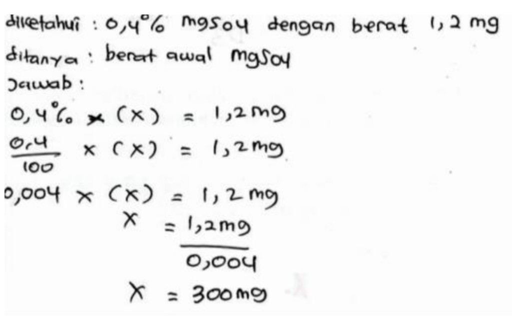
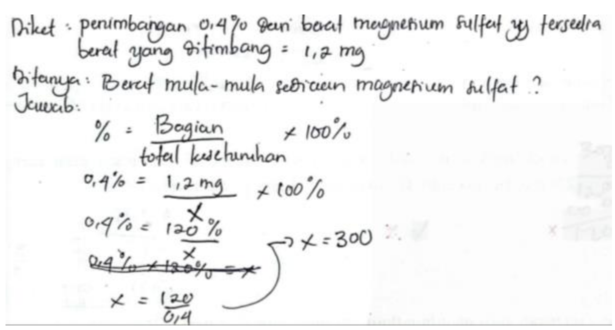
| | |
|---|---|
| <br><br>(a) | <br><br>(b) |
| **Second tier**<br>Known: 0.4% $MgSO_4$ with weight 1.2 mg<br>Asked: 0.4% $MgSO_4$ with weight 1.2 mg<br>Answer:<br>$0{,}4\% \times x = 1{,}2\ mg$<br>$\frac{0{,}4}{100} \times x = 1{,}2\ mg$<br>$0{,}004 \times x = 1{,}2\ mg$<br>$x = \frac{1{,}2\ mg}{0{,}004}$<br>$x = 300\ mg$<br><br>The process isn't finished yet; the units haven't been converted as asked in the question. | **Second tier**<br>Known: weighing of 0.4% from magnesium sulfate available<br>weight that was weighed = 1.2 mg<br>Asked:<br>initial weight of magnesium sulfate?<br>Answer:<br>$\% = \frac{part}{total} \times 100\%$<br>$0{,}4\% = \frac{1{,}2\ mg}{x} \times 100\%$<br>$0{,}4\% = \frac{120\%}{x}$<br>$x = \frac{120}{0{,}4}$<br>$x = 300$<br><br>The process isn't finished yet; the units haven't been converted as asked in the question. |

**Figure 4.** Examples of student responses showing errors in unit interpretation during pharmaceutical calculations: (a) using basic multiplication; (b) using proportional reasoning

Overall, these findings suggest that instructional strategies specifically targeting the procedural execution of unit conversion, rather than the introduction of additional conceptual material, may yield the greatest improvement in pharmaceutical calculation accuracy. Repeated, low stakes practice opportunities, such as brief in class calculation drills or laboratory based exercises specifically targeting the unit-conversion steps identified as high error points in this study (e.g., Items 2 and 4), may be more effective than relying solely on summative examinations to develop this skill.

## Conclusions and Suggestions

This study developed and empirically validated a modified three-level diagnostic test to identify misconceptions among pharmacy students regarding ratios, proportions, and percentages, which are fundamental concepts in calculating drug doses and concentrations. The main findings indicate that many of the difficulties experienced by students stem from errors in procedural execution (e.g., unit conversion), rather than a lack of understanding of basic concepts; many students were able to identify the correct concepts and formulas but still produced incorrect final answers due to unit conversion errors. This discrepancy has direct implications for patient safety, as calculations that are conceptually correct but procedurally flawed can result in drug dosages that are off by an order of magnitude (a pattern illustrated by the students' answers in this study) without the students realizing their errors. The open-ended essay format in the second stage proved capable of revealing these differences, something that the conventional multiple-choice format could not demonstrate.

For pharmacy faculty, these findings imply that the focus of instruction should not be limited to conceptual understanding but must also include deliberate and repeated practice focused on unit conversion steps, delivered through low stakes formative exercises rather than relying solely on summative exams. The diagnostic instrument developed in this study can serve as a formative tool for this purpose, enabling instructors to identify specific points of procedural error before those errors surface in high stakes assessments or clinical practice. This study was limited to students from a single institution, and future research should examine whether this conceptual procedural disparity replicates across other pharmacy programs and curricular structures.

## Acknowledgements

The authors would like to express their sincere gratitude to the pharmacy students of STIKes IKIFA who participated as respondents in this study. The authors also extend their appreciation to the lecturers who served as expert validators and provided valuable feedback during the instrument development process. Also, for Mr Collins Bekoe and Mr Benjamin Atta Owusu for their valuable discussion and suggestions.

**References**

Alrabiah, Z., Arafah, A., Rehman, M. U., Syed, W., Babelghaith, S., Alwhaibi, A., Alghadeer, S., Alhossan, A., & Al Arifi, M. N. (2022). Perception of pharmacy students toward numeracy: An observational study from King Saud University, Riyadh Saudi Arabia. *Frontiers in Public Health*, *10*, 1014328. https://doi.org/10.3389/fpubh.2022.1014328

Arkell, S., & Rutter, P. M. (2012). Numeracy skills of undergraduate entry level nurse, midwife and pharmacy students. *Nurse Education in Practice*, *12*(4), 198–203. https://doi.org/10.1016/j.nepr.2012.01.004

Aronson, B. D., Eddy, E., Long, B., Welch, O. K., Grundey, J., & Hinson, J. L. (2022). Identifying Low Pharmaceutical Calculation Performers Using an Algebra-Based Pretest. *American Journal of Pharmaceutical Education*, *86*(1), 8473. https://doi.org/10.5688/ajpe8473

Arslan, H. O., Cigdemoglu, C., & Moseley, C. (2012). A Three-Tier Diagnostic Test to Assess Pre-Service Teachers' Misconceptions about Global Warming, Greenhouse Effect, Ozone Layer Depletion, and Acid Rain. *International Journal of Science Education*, *34*(11), 1667–1686. https://doi.org/10.1080/09500693.2012.680618

Association of Indonesian Pharmacy Higher Education (APTFI). (2024). *Decree of the Chairperson of the Association of Indonesian Pharmacy Higher Education (APTFI) Number: 222/X/SK/APTFI/2024 concerning Learning Outcomes of Graduates of the Undergraduate (S1), Master's (S2), and Doctoral (S3) Pharmacy Programs, the Pharmacist Professional Program, and the Specialist Pharmacist Professional Program 2024*.

Aydin, A. K., & Dinç, L. (2017). Effects of Web-Based Instruction on Nursing Students' Arithmetical and Drug Dosage Calculation Skills. *CIN: Computers, Informatics, Nursing*, *35*(5), 262–269. https://doi.org/10.1097/CIN.0000000000000317

Branch, R. M. (2009). *Instructional Design: The ADDIE Approach*. Springer US. https://doi.org/10.1007/978-0-387-09506-6

Cahyadi, M. R., Maryanto, B. P. A., Syaifuddin, M., & Darmayanti, R. (2023). Development of Essay Test Assessment Rubric for Polya Theory-Based Mathematical Problem-Solving. *JNPM (Jurnal Nasional Pendidikan Matematika)*, *7*(1), 167–178. https://doi.org/10.33603/jnpm.v7i1.7724

Chamseddine, I. M., & Rejniak, K. A. (2020). Hybrid modeling frameworks of tumor development and treatment. *WIREs Systems Biology and Medicine*, *12*(1), e1461. https://doi.org/10.1002/wsbm.1461

Chen, S. (2009). Shadows: Young Taiwanese children's views and understanding. *International Journal of Science Education*, *31*(1), 59–79. https://doi.org/10.1080/09500690701633145

Commission for Educational Development. (2013). *Naskah Akademik Standar Kompetensi Lulusan Dan Standar Kurikulum Pendidikan Farmasi*.

Crawford, A. N., Anksorus, H. N., Clark Dula, C. A., Fredrickson, M. E., Peddi, A. N., Riley, B. L., Wisniewski, J. N., & Wu, S. (2024). Counting on Change: Conquering Challenges in Teaching Pharmaceutical Calculations. *American Journal of Pharmaceutical Education*, *88*(9), 101259. https://doi.org/10.1016/j.ajpe.2024.101259

Eastwood, K. J., Boyle, M. J., Williams, B., & Fairhall, R. (2011). Numeracy skills of nursing students. *Nurse Education Today*, *31*(8), 815–818. https://doi.org/10.1016/j.nedt.2010.12.014

Frankart, L. M., Dell, K. A., Matulewicz, A. T., Caldas, L. M., & Wantuch, G. A. (2025). Deficiencies in Mathematical Skills Identified in First-Year Students at Two Pharmacy Institutions. *American Journal of Pharmaceutical Education*, *89*(8), 101448. https://doi.org/10.1016/j.ajpe.2025.101448

Fredrickson, M. E., Nogid, A., Wu, S., Peddi, A. N., Whitney, R., Gratz, M., & Wisniewski, J. N. (2025). A Scoping Review of Teaching and Assessment Strategies for Pharmaceutical Calculations in Health Professions Education. *American Journal of Pharmaceutical Education*, *89*(8), 101430. https://doi.org/10.1016/j.ajpe.2025.101430

Gurel, D. K., Eryilmaz, A., & McDermott, L. C. (2017). Development and application of a four-tier test to assess pre-service physics teachers' misconceptions about geometrical optics. *Research in Science & Technological Education*, *35*(2), 238–260. https://doi.org/10.1080/02635143.2017.1310094

Habiddin, H., Akbar, D. F. K., Aziz, A. N., Hasan, H., & Mustapa, K. (2021). Developing a multi-tier instrument for chemistry teaching: A challenging exercise. • *The 4th International Conference On Mathematics And Science Education (ICoMSE) 2020: Innovative Research in Science and Mathematics Education in The Disruptive Era*, 020001. https://doi.org/10.1063/5.0043285

Hasan, S., Bagayoko, D., & Kelley, E. L. (1999). Misconceptions and the Certainty of Response Index (CRI). *Physics Education*, *34*(5), 294–299. https://doi.org/10.1088/0031-9120/34/5/304

Hegener, M. A. (2021). Design and evaluation of an application-based pharmaceutical calculations review module. *Currents in Pharmacy Teaching and Learning*, *13*(8), 1018–1023. https://doi.org/10.1016/j.cptl.2021.06.016

Hegener, M. A., Buring, S. M., & Papas, E. (2013). Impact of a Required Pharmaceutical Calculations Course on Mathematics Ability and Knowledge Retention. *American Journal of Pharmaceutical Education*, *77*(6), 124. https://doi.org/10.5688/ajpe776124

Hiebert, J., & Lefevre, P. (1986). *CONCEPTUAL AND PROCEDURAL KNOWLEDGE: THE CASE OF MATHEMATICS*.

Hijji, B. M. (2025). An Indispensable Requirement for Medical Dosage Calculation: Basic Mathematical Skills of Baccalaureate Nursing Students. *Nursing Reports*, *15*(5), 150. https://doi.org/10.3390/nursrep15050150

Istyastono, E. P. (2016). *Uji Statistik di Ilmu Farmasi dengan Program Statistika Komputasional R*. Sanata Dharma University Press.

Jairoun, A. A., Al-Hemyari, S. S., Shahwan, M., Al-Ani, M., Habeb, M., & El-Dahiyat, F. (2024). Bridging generations: The central role of career-stage mentorship in social and administrative pharmacy. *Exploratory Research in Clinical and Social Pharmacy*, *16*, 100506. https://doi.org/10.1016/j.rcsop.2024.100506

Johar, R., Moulina, A. R., Mailizar, Lestari, M., & Away, Y. (2024). Development of E-Learning based Remedial Videos on Fractions in Middle School. *Jurnal Pendidikan Matematika*, *18*(1), 97–112. https://doi.org/10.22342/jpm.v18i1.pp97-112

Jonan, Y. R. (2020). Pengembangan Rubrik Penskoran pada Asesmen Otentik untuk Materi Volume dan Luas Balok. *Journal of Medives : Journal of Mathematics Education IKIP Veteran Semarang*, *4*(2), 275–284. https://doi.org/10.31331/medivesveteran.v4i2.1174

Klausner, E. A., & Holt, M. J. (2022). The use of ratio and proportion, dimensional analysis, and equations to solve pharmacy calculations problems. *Currents in Pharmacy Teaching and Learning*, *14*(4), 492–498. https://doi.org/10.1016/j.cptl.2022.03.005

Maxnun, L., Kristiani, K., & Sulistyaningrum, C. D. (2024). Development of hots-based cognitive assessment instruments: ADDIE model. *Journal of Education and Learning (EduLearn)*, *18*(2), 489–498. https://doi.org/10.11591/edulearn.v18i2.21079

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, *22*(3), 276–282.

McMullan, M., Jones, R., & Lea, S. (2010). Patient safety: numerical skills and drug calculation abilities of nursing students and Registered Nurses. *Journal of Advanced Nursing*, *66*(4), 891–899. https://doi.org/10.1111/j.1365-2648.2010.05258.x

Milenković, D. D., Hrin, T. N., Segedinac, M. D., & Horvat, S. (2016). Development of a Three-Tier Test as a Valid Diagnostic Tool for Identification of Misconceptions

Related to Carbohydrates. *Journal of Chemical Education*, *93*(9), 1514–1520. https://doi.org/10.1021/acs.jchemed.6b00261

Mir, U. A., Joldan, S., Kumar, V., & Saini, S. (2025). *International Handbook of Research Methods And Statistics* (Vol. 1). The Readers Paradise.

Mnatzaganian, C. L., Singh, R. F., Best, B. M., & Morello, C. M. (2020). Effectiveness of Providing Video Podcasts to Pharmacy Students in a Self-Study Pharmaceutical Calculations Module. *American Journal of Pharmaceutical Education*, *84*(12), ajpe7977. https://doi.org/10.5688/ajpe7977

Mohyuddin, R. G., & Khalil, U. (2016). Misconceptions of Students in Learning Mathematics at Primary Level. *Bulletin of Education and Research*, *38*(1), 133–162.

Muttaqin, H., Putri, R. I. I., & Somakim, S. (2017). Design Research On Ratio And Proportion Learning By Using Ratio Table And Graph With Oku Timur Context At 7th Grade. *Journal on Mathematics Education*, *8*(2), 211–222. https://doi.org/10.22342/jme.8.2.3969.211-222

National Association of Boards of Pharmacy. (2024). *NAPLEX Content Outline*.

Nofriati, N. F., Hartono, Y., & Somakim, S. (2020). Learning Direct and Inverse Proportion Using Musi Tour. *International Journal on Emerging Mathematics Education*, *3*(2), 139–151. https://doi.org/10.12928/ijeme.v3i2.13578

Novak, J. D. (1990). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching*, *27*(10), 937–949. https://doi.org/10.1002/tea.3660271003

Nurhaifa, I., Hamdu, G., & Suryana, Y. (2020). Rubrik Penilaian Kinerja pada Pembelajaran STEM Berbasis Keterampilan 4C. *Indonesian Journal of Primary Education*, *4*(1), 101–110. https://doi.org/10.17509/ijpe.v4i1.24742

Nurwahida, N., & Munir, N. P. (2022). Analisis Miskonsepsi Calon Guru Sekolah Dasar pada Matakuliah Konsep Dasar Matematika dengan Menggunakan Three-Tier Diagnostic Test Dilengkapi Certainty Of Response Index. *JUDIKDAS: Jurnal Ilmu Pendidikan Dasar Indonesia*, *1*(3), 153–164. https://doi.org/10.51574/judikdas.v1i3.455

Prescott, W. A., Maerten-Rivera, J., Anadi, I. S., Woodruff, A. E., & Fusco, N. M. (2024). Impact of Collaborative Testing on Academic Performance in Pharmacy Education. *American Journal of Pharmaceutical Education*, *88*(8), 100738. https://doi.org/10.1016/j.ajpe.2024.100738

Ramdhani, L. I., Arianda, Y. D., Tjalla, A., & Wahyuni, L. D. (2025). Development Of Instrument For Understanding The Concept Of Writing Scientific Articles For Lecturers Using The Addie Model: Empirical Validity And Reliability. *Indonesian*

*Journal of Educational Development (IJED)*, *5*(4), 429–443. https://doi.org/10.59672/ijed.v5i4.4324

Rusilowati, A. (2015). Pengembangan Tes Diagnostik Sebagai Alat Evaluasi Kesulitan Belajar Fisika. *Prosiding Seminar Nasional Fisika Dan Pendidikan Fisika (SNFPF) Ke-6*, 1–10.

Sari, F. Y., Zulkardi, Putri, R. I. I., & Susanti, E. (2025a). How do junior high school students solve proportional reasoning problems? *Jurnal Elemen*, *11*(2), 363–375. https://doi.org/10.29408/jel.v11i2.27922

Sari, F. Y., Zulkardi, Putri, R. I. I., & Susanti, E. (2025b). Ratios in Agriculture: A Rice Fertilization Context for Ratio Learning Trajectory. *Jurnal Pendidikan Matematika*, *19*(3), 509–526. https://doi.org/10.22342/mej.v19i3.pp509-526

Sari, F. Y., Zulkardi, Putri, R. I. I., Susanti, E., & Nusantara, D. S. (2024). Ratio Material Learning Design Using the Context of Cooking Rice to Help Elementary Students Understand Concepts. *INOMATIKA*, *6*(1), 54–66. https://doi.org/10.35438/inomatika.v6i1.443

Savitri, F., Sabrina, A., & Tuahuns, F. (2025). Developing a textbook to support pharmacy students' mathematical problem-solving skills. *Indonesian Journal of Educational Development (IJED)*, *6*(3), 909–927. https://doi.org/10.59672/ijed.v6i3.5150

Sharma, P., & Dunham, A. (2025). *Pharmacy Calculations* (6th ed.). StatPearls Publishing.

Siahaan, L. H. (2025). *R&D dalam Pendidikan. Implementasi Model ADDIE dan 4D pada Pendidikan Bahasa Inggris dan PG PAUD* (1st ed.). Penerbit KBM Indonesia.

Sukasno, Zulkardi, Putri, R. I. I., & Somakim. (2024). Students' Cognitive Processes in Understanding Fractions Through the Tourist Context. *Jurnal Pendidikan Matematika*, *18*(1), 27–38. https://doi.org/10.22342/jpm.v18i1.pp27-38

Suprihanto, E., & Sahrul, S. (2026). *Teknik Penyusunan Instrumen Tes dan Non tes dalam Penelitian*. Yayasan Putra Adi Dharma.

Suryadin, A. (2024). *Assessment Prestasi Belajar (Untuk Guru dan Mahasiswa Keguruan)*. Eureka Media Aksara.

Tajuddin, N. I. I., Abas, U.-H., Aziz, K. A., Nor, R. N. H., Izni, N. A., Sudin, M. N., Anim, N. A. H. M., & Noor, N. M. (2025). Content Validity Assessment Using Aiken's V: Knowledge Integration Model for Blockchain in Higher Learning Institutions. *International Journal of Advanced Computer Science and Applications*, *16*(6). https://doi.org/10.14569/IJACSA.2025.0160659

Thompson, P. W. (2022). Quantitative Reasoning as an Educational Lens. In G. K. Akar, İ. Ö. Zembat, S. Arslan, & P. W. Thompson (Eds.), *Quantitative Reasoning in Mathematics and Science Education* (pp. 1–16). Springer. https://doi.org/10.1007/978-3-031-14553-7_1

Wahyudi, F., Didik, L. A., & Bahtiar, B. (2021). Pengembangan Instrumen Three Tier Test Diagnostik Untuk Menganalisis Tingkat Pemahaman Dan Miskonsepsi Siswa Materi Elastisitas. *Relativitas: Jurnal Riset Inovasi Pembelajaran Fisika*, *4*(2), 48–58. https://doi.org/10.29103/relativitas.v4i2.5184

Wahyuningrum, A. S., Suryadi, D., & Turmudi, T. (2022). Students' Prior Knowledge as an Ontogenic Obstacle on the Topic of Ratio and Proportion. *Jurnal Pendidikan Matematika*, *17*(1), 55–68. https://doi.org/10.22342/jpm.17.1.18866.55-68

Wang, J.-F., Wang, T.-H., & Huang, C.-H. (2022). Investigating Students' Answering Behaviors in a Computer-Based Mathematics Algebra Test: A Cognitive-Load Perspective. *Behavioral Sciences*, *12*(8), 293. https://doi.org/10.3390/bs12080293

Wennberg-Capellades, L., Fuster-Linares, P., Rodríguez-Higueras, E., Fernández-Puebla, A. G., & Llaurado-Serra, M. (2022). Where do nursing students make mistakes when calculating drug doses? A retrospective study. *BMC Nursing*, *21*(1), 309. https://doi.org/10.1186/s12912-022-01085-9

Woodcock, J., & Woosley, R. (2008). The FDA Critical Path Initiative and Its Influence on New Drug Development. *Annual Review of Medicine*, *59*(1), 1–12. https://doi.org/10.1146/annurev.med.59.090506.155819