**Jurnal Elemen, 12(1), 49-65, January 2026**
https:/doi.org/10.29408/jel.v12i1.32327

e- ISSN: 2442 - 4226

# Conceptual and procedural errors of pre-service elementary teachers in solving educational research statistics problems

**Siti Maryam Rohimah [1 *], Sarah Anida Putri [1], Yullys Helsa [2]**

[1] Department of Elementary School Teacher Education, Universitas Pasundan, West Java, Indonesia
[2] Department of Elementary School Teacher Education, Universitas Negeri Padang, West Sumatra, Indonesia

[*] Correspondence: sitimaryamrohimah@unpas.ac.id


## Abstract

Pre-service elementary teachers often make errors in educational research and problem-solving, highlighting the need to examine the nature and causes of these errors. This study employed a qualitative descriptive case study design to identify the types of errors made by Elementary School Teacher Education (ESTE) students in solving educational research statistics problems and to examine the factors contributing to these errors. The participants were 58 ESTE students enrolled in a statistics course at a private university in Indonesia. Data were collected through problem-solving tests, interviews, and classroom observations. The test instrument was used to assess students' understanding of statistical concepts, while interviews and observations explored cognitive and contextual factors influencing error occurrence. The findings indicate that conceptual errors occurred in 17% of cases and procedural errors in 22%. Conceptual errors were primarily due to difficulties interpreting statistical software outputs, whereas procedural errors were associated with inaccuracies in applying formulas and performing calculations. These findings contribute to a clearer theoretical understanding of conceptual and procedural errors in statistics learning and highlight the need for instructional approaches that balance conceptual and procedural aspects in educational research statistics.

**Keywords:** error analysis; learning difficulties; problem-solving; statistics education research

**How to cite:** Rohimah, S. M., Putri, S. A., & Helsa, Y. (2026). Conceptual and procedural errors of pre-service elementary teachers in solving educational research statistics problems. *Jurnal Elemen, 12*(1), 49-65. https://doi.org/10.29408/jel.v12i1.32327

Received: 16 September 2025 | Revised: 22 December 2025
Accepted: 20 January 2026 | Published: 19 February 2026

## Introduction

Statistics is an important tool in research, especially in education. Statistical mastery supports data-based decision-making in education (Vetten et al., 2023). Educational research requires a systematic approach to collecting, analyzing, and interpreting data using statistics (Rohimah, 2024). Through statistical reasoning, individuals can analyze data objectively, draw valid conclusions, and communicate research findings effectively (Abu-ghalyoun, 2021; Radke et al., 2023; Thanheiser & Mamolo, 2024). For students engaged in research, particularly those completing thesis research, mastery of statistics is essential to ensure accurate data processing and meaningful interpretation of results, which ultimately determine the quality and credibility of educational research.

However, inadequate mastery of statistics often leads to errors in its application, hindering students' ability to use statistical methods effectively. According to Legesse et al. (2020), Errors in statistical learning can be categorized into two main types: conceptual and procedural errors. Conceptual errors occur when students do not understand basic statistical concepts, while procedural errors occur when students make mistakes in mathematical steps during calculations or when using formulas. This aligns with Winarso and Toheri (2021) conceptual errors, namely the inability to determine and apply the correct theorem to solve a problem. At the same time, procedural errors occur when the rules used in the solution process are misapplied. In addition, external factors such as exam anxiety and lack of basic mathematical mastery also influence these errors (Mendes et al., 2024; Dodeen & Alqawasmi, 2025). Errors that emerge during statistical problem-solving reflect a gap between theoretical understanding and practical application, which may subsequently affect the quality of future classroom instruction (Shimizu & Kang, 2025).

Previous research consistently shows that errors in solving statistical problems stem from cognitive factors, such as difficulty understanding formulas, inaccurate calculations, and trouble interpreting statistical results (Leng & Meng, 2023; Silva & Sarnecka, 2025). These errors often appear as misunderstandings or mistakes with procedures, particularly when applying formulas or interpreting software output (Nisa & Dahlan, 2025; Sardareh et al., 2025). For pre-service teachers, affective factors like statistical anxiety often worsen these difficulties by negatively affecting reasoning and problem-solving accuracy (Romero et al., 2023).

Furthermore, students frequently make errors in selecting appropriate statistical tests, interpreting tables and graphs, applying formulas, and drawing valid conclusions (Pallauta et al., 2021; Parks & Yeh, 2021; Lakshmanan, 2022). However, most of these studies focus on general statistics learning or on students in secondary and higher education, with limited attention to pre-service elementary teachers and the specific demands of educational research statistics. In addition, prior research tends to discuss statistical errors in broad terms, without explicitly examining how conceptual and procedural errors manifest in research-oriented statistical tasks or how they relate to the use of statistical software.

Addressing these gaps, the present study offers a focused analysis of conceptual and procedural errors made by Elementary School Teacher Education (ESTE) students in solving educational research statistics problems. The novelty of this study lies in its explicit

differentiation between conceptual and procedural errors in a research-statistics context, as well as its integration of cognitive and affective factors that influence error occurrence. By examining errors in both statistical software interpretation and formula-based calculations, this study seeks to extend existing theoretical frameworks on statistical learning errors and to provide empirical insights relevant to pre-service teacher education.

The study aims to identify the types of conceptual and procedural errors ESTE students make in educational research statistics problems. It also examines the factors that contribute to these errors. The findings are expected to inform the design of statistics instruction in teacher education programs. They should support instructional approaches that balance conceptual understanding and procedural fluency, thereby improving future classroom instruction.

## Methods

### Research design

This study employed a qualitative descriptive case study design to identify the types of cognitive errors made by Elementary School Teacher Education (ESTE) students when solving statistics problems, as well as the factors contributing to these errors. A qualitative approach was chosen because the study aimed to gain an in-depth understanding of students' reasoning processes, error characteristics, and learning experiences. The case study design enabled a detailed exploration of error patterns within a specific instructional context, particularly statistics learning for educational research (Shamsuddin et al., 2021; Lian et al., 2022; Prameshti et al., 2024). The study was conducted at one university offering an ESTE program and involved students enrolled in a statistics course. Participation was voluntary, and confidentiality and anonymity of responses were ensured. The reported percentages of conceptual and procedural errors serve as descriptive indicators of the relative occurrence of each error type and are not intended as inferential statistical measures.

### Participants and sampling

The participants in this study were 58 seventh-semester Elementary School Teacher Education (ESTE) students, comprising 6 males and 52 females, enrolled in an educational research statistics course at a private university in Indonesia. The main data source was students' responses to a course-based statistics assessment used as a problem-solving test for error analysis. The assessment consisted of essay-type questions designed to evaluate students' conceptual and procedural understanding of core topics in educational research statistics, including fundamental concepts and roles of statistics in education, data concepts and quality criteria, measures of central tendency and data presentation, measures of dispersion, standard scores and Z-scores, inferential statistics, hypothesis formulation, and tests of normality and homogeneity. Using a course-based assessment ensured that the identified errors reflected authentic student performance in an academic learning context.

## Instruments and data collection

The research instrument consisted of six essay questions (Table 1) designed to measure students' understanding of basic statistical concepts, including Measurement Scale (Q1), Frequency Distribution (Q2), Measure of Data Centralization and Measure of Data Location (Q3), Standard Score (Q4), Normality Test (Q5), and Homogeneity Test (Q6). The instrument was validated by experts in educational statistics and linguistics based on content relevance, language clarity, and alignment with the research objectives (Ismail et al., 2025; Sugiyono, 2017).

Data collection in this study employed three complementary methods: tests, observations, and interviews. The test instrument was used to assess students' understanding of statistical concepts and their ability to solve statistics problems, while observations were conducted concurrently to examine students' problem-solving processes. In-depth interviews were conducted with five selected students to explore the underlying reasons for their errors, using guiding questions such as "How did you interpret this statistical output?" "What difficulties did you encounter when applying the formula?", and "How confident were you when answering this question?" The integration of tests, observations, and interviews supported data triangulation and enabled a comprehensive analysis of students' cognitive processes and error patterns in solving statistical problems.

## Validity and reliability

Data validity was ensured through triangulation by comparing data from multiple sources, namely tests, interviews, and classroom observations, to obtain a more comprehensive and accurate understanding of the findings (Sugiyono, 2017). To further enhance credibility, peer debriefing and member checking were conducted by discussing preliminary interpretations with colleagues and confirming them with research participants. Data reliability was maintained through an audit trail, which systematically documented the data collection and analysis processes to ensure transparency and consistency, allowing the research procedures to be traced and reviewed (Miles et al., 2014).

## Data analysis

Data were analyzed inductively using thematic analysis. Students' written responses were first read holistically and then coded based on the type of error identified. Coding focused on three main categories: conceptual errors, procedural errors, and factors causing errors. For example, a response in Q1 that incorrectly classified interval and ratio scales was coded as a conceptual error, while an incorrect calculation of the mean in Q3 was coded as a procedural error. Interview data explaining difficulties due to time pressure or exam anxiety were coded under the factor 'causing errors'. Similar codes were subsequently grouped into broader themes to identify recurring error patterns and contributing factors. The results of the coding and thematic interpretation were presented in tables and diagrams to enhance clarity (Miles et al., 2014).

**Table 1.** Statistics questions

**Q1**

Lengkapilah tabel berikut dengan skala pengukuran yang sesuai (nominal, ordinal, interval dan rasio)

| No | Peubah | Skala Pengukuran |
|---|---|---|
| 1 | Tingkat gizi balita (kurang, cukup, baik, sangat baik) | |
| 2 | Nilai indeks prestasi kumulatif (IPK) | |
| 3 | Taraf penyakit: ringan, sedang, berat, sangat berat | |
| 4 | Tingkat inflasi negara pada selang waktu tertentu (%) | |
| 5 | Nilai kecerdasan (IQ) | |
| 6 | Penghasilan rumah tangga per bulan (Rp) | |
| 7 | Letak bujur dan lintang suatu tempat (derajat) | |
| 8 | Penyebab penyakit: Virus, bakteri, bahan kimia dll | |
| 9 | Jenis pekerjaan kepala rumah tangga (petani, peternak, PNS, wiraswasta, dll) | |
| 10 | Dalam suatu penelitian tingkat kesukaan konsumen terhadap suatu produk, sering digunakan skala 3 untuk menunjukkan suka, 2 menunjukkan sedang, dan 1 untuk menunjukkan tidak suka. | |

Complete the table along with the scale-appropriate measurements (nominal, ordinal, interval, and ratio)

| No | Variable | Measurement Scale |
|---|---|---|
| 1 | Nutritional level toddler (less, enough, good, perfect) | |
| 2 | Index value performance cumulative (IPK) | |
| 3 | Disease level: mild, moderate, severe, very severe | |
| 4 | The country's inflation rate over time is certain (%) | |
| 5 | Intelligence quotient (IQ) | |
| 6 | Household income per month (Rp) | |
| 7 | The longitude and latitude location of a place (degree) | |
| 8 | Reason disease: viruses, bacteria, chemicals, etc. | |
| 9 | Type of work head of household (farmer, rancher, civil servant, self-employed, etc.) | |
| 10 | At a study level favorite consumers of products, often used a scale of 3 to show like, 2 to show moderate, and 1 to indicate dislike. | |

**Q2**

Berikut ini adalah nilai Ujian Akhir Semester suatu mata kuliah dari 25 mahasiswa di PGSD di Universitas A:

| | | | | |
|---|---|---|---|---|
| 75 | 67 | 71 | 65 | 90 |
| 80 | 83 | 71 | 64 | 66 |
| 85 | 82 | 72 | 79 | 75 |
| 90 | 87 | 80 | 82 | 72 |
| 61 | 63 | 80 | 89 | 89 |

Dari data tersebut buatlah:

a. Tabel distribusi frekuensi
b. Histogram
c. Diagram lingkaran

The following are the Final Semester Exam scores for a course from 25 students in ESTE at University A:

| | | | | |
|---|---|---|---|---|
| 75 | 67 | 71 | 65 | 90 |
| 80 | 83 | 71 | 64 | 66 |
| 85 | 82 | 72 | 79 | 75 |
| 90 | 87 | 80 | 82 | 72 |
| 61 | 63 | 80 | 89 | 89 |

From this data, create:

a. Distribution table frequency
b. Histogram
c. Pie chart

**Q3**

Berikut ini data tinggi badan 25 mahasiswa (dalam centimeter) yang diambil secara acak.

| No. | Tinggi | Gender | No. | Tinggi | Gender |
|---|---|---|---|---|---|
| 1 | 170 | Pria | 14 | 165 | Wanita |
| 2 | 172 | Pria | 15 | 165 | Wanita |
| 3 | 160 | Wanita | 16 | 170 | Pria |
| 4 | 162 | Pria | 17 | 175 | Pria |
| 5 | 159 | Wanita | 18 | 166 | Wanita |
| 6 | 168 | Pria | 19 | 168 | Pria |
| 7 | 162 | Wanita | 20 | 164 | Wanita |
| 8 | 164 | Wanita | 21 | 150 | Wanita |
| 9 | 168 | Pria | 22 | 155 | Wanita |
| 10 | 159 | Wanita | 23 | 168 | Pria |
| 11 | 160 | Wanita | 24 | 164 | Pria |
| 12 | 161 | Wanita | 25 | 165 | Wanita |
| 13 | 162 | Pria | | | |

Tentukan:

a. Mean
b. Median

The following is data on the height of 25 students (in centimeters) taken randomly.

| No. | Tall | Gender | No. | Tall | Gender |
|---|---|---|---|---|---|
| 1 | 170 | Man | 14 | 165 | Woman |
| 2 | 172 | Man | 15 | 165 | Woman |
| 3 | 160 | Woman | 16 | 170 | Man |
| 4 | 162 | Man | 17 | 175 | Man |
| 5 | 159 | Woman | 18 | 166 | Woman |
| 6 | 168 | Man | 19 | 168 | Man |
| 7 | 162 | Woman | 20 | 164 | Woman |
| 8 | 164 | Woman | 21 | 150 | Woman |
| 9 | 168 | Man | 22 | 155 | Woman |
| 10 | 159 | Woman | 23 | 168 | Man |
| 11 | 160 | Woman | 24 | 164 | Man |
| 12 | 161 | Woman | 25 | 165 | Woman |
| 13 | 162 | Man | | | |

Determine:

a. Mean
b. Median

| Code | Questions (in Bahasa) | Questions (in English) |
|---|---|---|
| | c. Kuartil ke-tiga<br>d. Percentil 10<br>e. Std. Deviasi<br>f. Varians | c. Third quartile<br>d. 10th percentile<br>e. Std. Deviation<br>f. Variance |
| Q4 | Hasil ujian dua kelompok yang berbeda diperoleh, kelompok 1 memperoleh rata- rata 25 dan simpangan baku 4. Sedangkan kelompok 2 memperoleh rata-rata 30 dan simpangan baku 3. Rahman adalah siswa dari kelompok 1 memperoleh skor 33, dan Bayu dari kelompok 2 memperoleh skor 33. Siapakah yang memperoleh skor baku lebih tinggi? | Test results of two different groups obtained, group 1 obtained an average of 25 and a deviation of standard 4. Meanwhile, group 2 obtained an average of 30 and a deviation of standard 3. Rahman is student from group 1 obtained a score of 33, and Bayu from group 2 got a score of 33. Who gets the score standard moretall? |
| Q5 | Tabel berikut ini nilai ujian statistika antara kelas A dan kelas B di Universitas B. Penelitian dengan menggunakan sampel sebanyak 20 responden yang diambil dari kelas A dan kelas B. Kelas A sebanyak 8 orang dan kelas B sebanyak 12 orang. Data-data yang didapat sebagai berikut.<br>Tabel. Tabulasi Data<br>No \| Nilai Ujian \| Kelas<br>1 \| 32 \| Kelas A<br>2 \| 35 \| Kelas A<br>3 \| 41 \| Kelas A<br>4 \| 39 \| Kelas A<br>5 \| 45 \| Kelas A<br>6 \| 43 \| Kelas A<br>7 \| 42 \| Kelas A<br>8 \| 47 \| Kelas A<br>9 \| 42 \| Kelas A<br>10 \| 37 \| Kelas A<br>11 \| 35 \| Kelas B<br>12 \| 36 \| Kelas B<br>13 \| 30 \| Kelas B<br>14 \| 28 \| Kelas B<br>15 \| 26 \| Kelas B<br>16 \| 27 \| Kelas B<br>17 \| 32 \| Kelas B<br>18 \| 35 \| Kelas B<br>19 \| 38 \| Kelas B<br>20 \| 41 \| Kelas B<br>Tentukanlah:<br>a. Hasil uji normalitas kelas A<br>b. Hasil uji normalitas kelas B | The following table shows the statistics of test scores between class A and class B at University B. Research using a sample of a total of 20 respondents was taken from class A and class B. Class A has 8 people and Class B has 12 people. The data obtained are as follows.<br>Table. Data Tabulation<br>No \| Test scores \| Class<br>1 \| 32 \| Class A<br>2 \| 35 \| Class A<br>3 \| 41 \| Class A<br>4 \| 39 \| Class A<br>5 \| 45 \| Class A<br>6 \| 43 \| Class A<br>7 \| 42 \| Class A<br>8 \| 47 \| Class A<br>9 \| 42 \| Class A<br>10 \| 37 \| Class A<br>11 \| 35 \| Class B<br>12 \| 36 \| Class B<br>13 \| 30 \| Class B<br>14 \| 28 \| Class B<br>15 \| 26 \| Class B<br>16 \| 27 \| Class B<br>17 \| 32 \| Class B<br>18 \| 35 \| Class B<br>19 \| 38 \| Class B<br>20 \| 41 \| Class B<br>Determine:<br>a. Normality test results class A<br>b. Normality test results class B |
| Q6 | Tabel di bawah ini merupakan data Nilai hasil belajar dengan menggunakan Metode Ceramah (Kelas A) dan metode diskusi (Kelas B): | The table below shows the learning outcome values using the lecture method (Class A) and discussion method (Class B): |

| Code | Questions (in Bahasa) | Questions (in English) |
|---|---|---|

| Siswa Ke. | Metode Ceramah (Kelas A) |
|---|---|
| 1. | 95 |
| 2. | 98 |
| 3. | 76 |
| 4. | 90 |
| 5. | 87 |
| 6. | 89 |
| 7. | 77 |
| 8. | 92 |
| 9. | 78 |
| 10. | 82 |

| Siswa Ke. | Metode diskusi (Kelas B) |
|---|---|
| 1. | 100 |
| 2. | 94 |
| 3. | 78 |
| 4. | 98 |
| 5. | 90 |
| 6. | 85 |
| 7. | 86 |
| 8. | 87 |
| 9. | 80 |
| 10. | 83 |

Ujilah apakah dua kelompok data di atas memiliki varians yang homogen dengan menggunakan $\alpha$= 5%!

| Student No. | Lecture Method (Class A) |
|---|---|
| 1. | 95 |
| 2. | 98 |
| 3. | 76 |
| 4. | 90 |
| 5. | 87 |
| 6. | 89 |
| 7. | 77 |
| 8. | 92 |
| 9. | 78 |
| 10. | 82 |

| Student No. | Discussion method (Class B) |
|---|---|
| 1. | 100 |
| 2. | 94 |
| 3. | 78 |
| 4. | 98 |
| 5. | 90 |
| 6. | 85 |
| 7. | 86 |
| 8. | 87 |
| 9. | 80 |
| 10. | 83 |

Test whether the two groups of data above have homogeneous variance by using a = 5%!

## Results

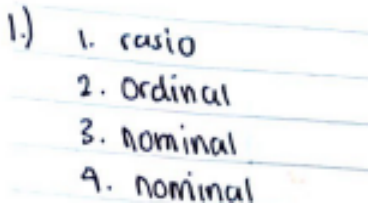
Based on analyses of test results, observations, and interview data from 58 Elementary School Teacher Education (ESTE) students, two major categories of errors were identified: conceptual and procedural errors. Conceptual errors refer to misunderstandings of fundamental statistical concepts, whereas procedural errors involve incorrect application of formulas, methods, or statistical software procedures. Conceptual errors were found in the answers to questions Q1, Q4, and Q6, while procedural errors were found in the answers to questions Q2, Q3, and Q5.

### Conceptual errors

Conceptual errors were found at an average of 17%. Conceptual errors refer to students' misunderstandings of fundamental statistical ideas, definitions, or interpretations. These errors occur when students fail to correctly grasp the meaning of a concept. In this study, conceptual errors were in Q1 (Measurement Scale), followed by Q4 (Standard Score) and Q6 (Homogeneity Test).

**Conceptual errors in Q1: Measurement scale**

Conceptual errors in Q1 were found in 27%. These errors mainly involved incorrect classification of variables into nominal, ordinal, interval, and ratio scales. Figure 2 presents a representative student response illustrating this type of error.




Translation:
1. ratio
2. ordinal
3. nominal
4. nominal

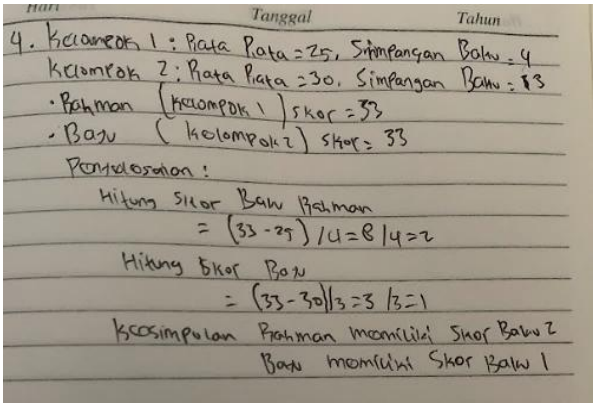
**Figure 2.** Student misclassification of measurement scales

Figure 2 shows one student's answer that illustrates a conceptual error in identifying measurement scales. In this response, the student misclassified variables such as nutritional status, GPA, disease severity, and inflation rate, indicating confusion in distinguishing between nominal, ordinal, interval, and ratio scales. These errors suggest that the student relied on

surface features or intuitive judgments rather than the defining properties of each scale, such as the presence of order, equal intervals, and meaningful zero points. This misclassification highlights a lack of understanding of measurement scale theory.

**Conceptual errors in Q4: Standard score**

Conceptual errors in Question 4 were identified in 10%. This question was designed to assess students' understanding of standard scores (z-scores) and their interpretation when comparing scores from different groups with different means and standard deviations. Figure 3 shows a sample of a student's written response that illustrates this conceptual error.




Translation:

4. Group 1: Average = 25, Standard Deviation = 4
Group 2: Average = 30, Standard Deviation = 3

- Rahman (group 1) score = 33
- Bayu (Group 2) score = 33

Solution:

Calculate Rahman's Standard score
= (33-25)/4 = 8/4 = 2
Calculate Bayu's score
= (33-30)/3 = 3/3 = 1

Conclusion Rahman has a Standard score of 2
Bayu has a Standard Score of 1

**Figure 3.** Conceptual error in standard score interpretation

The responses in Figure 3 reflect a conceptual error in interpreting standard scores. The student treated the standard score merely as a calculated value, without demonstrating an understanding of its meaning as a relative position within a distribution. The conclusion was stated without conceptual justification regarding what the standard scores imply about each student's performance relative to their respective groups. This indicates that the student did not fully grasp the underlying concept of standardization and its role in making meaningful comparisons across different distributions.

**Conceptual errors in Q6: Homogeneity test**

Conceptual errors in Question 6 were identified in 14% of students. This question aimed to assess students' conceptual understanding of the homogeneity of variance test, particularly their ability to interpret the output from Levene's Test generated by SPSS. Figure 4 presents a representative student response illustrating this conceptual error.

| Test of Homogeneity of Variance | | | | | |
|---|---|---|---|---|---|
| | | Levene Statistic | df1 | df2 | Sig. |
| Nilaihasilbelajar | Based on Mean | .126 | 1 | 18 | .727 |
| | Based on Median | .126 | 1 | 18 | .727 |
| | Based on Median and with adjusted df | .126 | 1 | 17.907 | .727 |
| | Based on trimmed mean | .140 | 1 | 18 | .713 |

**Kesimpulan** : Nilai F hitung (1,12) < F tabel (3,18) maka, $H_0$ diterima. Artinya dengan 5%, kedua kelompok hasil belajar yang menggunakan metode diskusi dengan hasil belajar yang menggunakan metode ceramah memiliki varians yang homogen

Translation:

**Conclusion:** The calculated F value (1.12) < F table (3.18) therefore, H0 is accepted. This means that with 5%, both groups of learning outcomes using the discussion method and learning outcomes using the lecture method have homogeneous variance.

**Figure 4.** Conceptual error in interpreting homogeneity test results

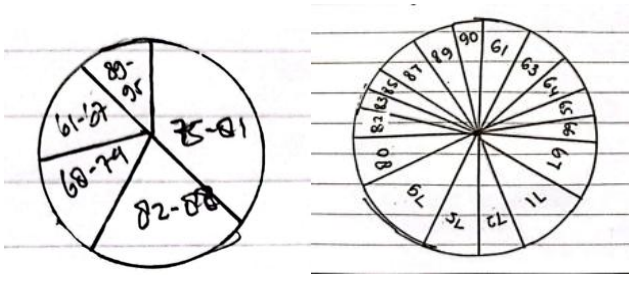
The response shown in Figure 4 indicates a conceptual error in statistical decision-making. The student based the conclusion on a comparison between the calculated F value and the F table value, revealing a misunderstanding of the decision rule used in SPSS-based hypothesis testing. In Levene's Test, conclusions should be drawn from the significance value (Sig. or p-value), not from manual F-value comparisons. Although the student reached the correct conclusion regarding variance homogeneity, the reasoning reflects an incorrect conceptual framework, indicating a misunderstanding of the underlying principles of hypothesis testing rather than a purely procedural or technical error.

## Procedural errors

Procedural errors refer to mistakes in applying formulas, performing calculations, and following correct solution steps. In this study, procedural errors were identified in Q2, Q3, and Q5, with an overall average of 22%.

### Procedural Errors in Q2: Frequency Distribution

Procedural errors in Question 2 were identified in 16%. In this question, students were required to construct a frequency distribution table, a histogram, and a pie chart based on a given dataset. This task assessed students' ability to correctly apply procedural steps in organizing data and transforming it into appropriate statistical representations. Figures 5a and 5b present two student responses that illustrate common procedural errors in constructing pie charts.



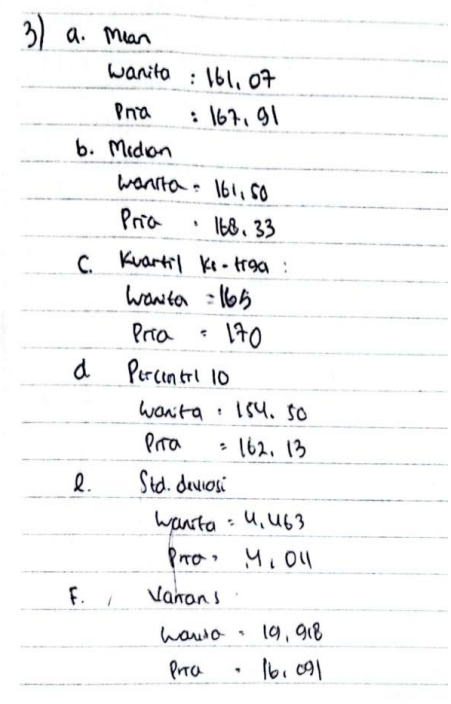
(a) (b)

**Figure 5.** Procedural errors in constructing Pie Charts (a, b)

The responses in Figures 5a and 5b show procedural errors in constructing pie charts. The students did not follow the correct steps for converting data into proportional sectors, such as grouping data and calculating frequencies or angles. One pie chart was constructed using unequal sectors without proportional calculations, while the other was divided into many equal parts based on raw data.

**Procedural errors in Q3: Measures of central tendency and data location**

Procedural errors in Question 3 were the most frequent, occurring in 29% of students. This question required students to compute measures of central tendency, data location, and variability based on the given dataset. Figure 6 presents a representative student response that illustrates errors in applying formulas and following the required calculation procedures.




Translation:
3) a. Mean
Female: 161.07
Male: 167.91
b. Median
Female: 161.50
Male: 168.33
c. Third Quartile
Female: 165
Male: 170
d. 10th Percentile
Female: 154.50
Male: 162.13
e. Standard Deviation
Female: 4.463
Male: 4.011
f. Variance
Female: 19.918
Male: 16.091

**Figure 6.** Procedural Errors in Applying Descriptive Statistics Formulas

The student's answer in Figure 6 demonstrated procedural errors in performing the required statistical calculations. Although the question required analyzing the data as a whole, the student unnecessarily separated the data by gender, leading to incorrect application of formulas and calculation steps. This demonstrates difficulty following correct procedures and instructions, such as selecting the correct formula and applying it consistently. These errors reflect weaknesses in procedural implementation.

**Procedural Errors in Q5: Normality Test**

Procedural errors in Question 5 were identified in 22%. This question required students to perform and interpret a normality test using SPSS output. Figure X presents a representative student response that illustrates procedural errors in selecting and using the appropriate SPSS results.

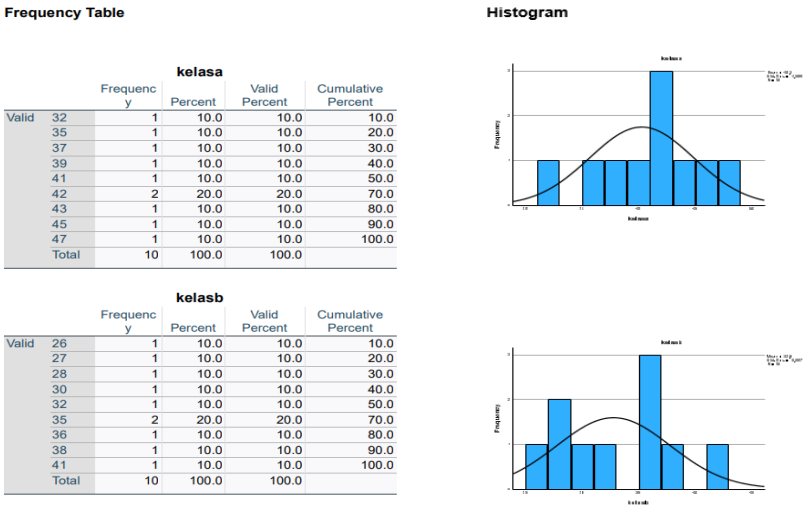



**Figure 7.** Procedural errors in selecting SPSS output for normality testing

The responses in Figure 7 demonstrate procedural errors in performing normality tests. Instead of using the appropriate normality test output (such as the Kolmogorov–Smirnov or Shapiro–Wilk test), students were presented with frequency tables and histograms. This indicates that students did not follow the procedural steps required to correctly perform and report normality tests in SPSS. Although students appeared to understand the purpose of normality testing, they encountered difficulties in carrying out the technical procedures, reflecting procedural errors.

## Factors contributing to students' errors

The analysis of students' written responses, observations, and interview data indicates that students' errors were influenced by both internal and external factors. In-depth interviews were conducted with five selected students, and representative excerpts are presented to explain the dominant sources of error identified in the test results.

Internal factors emerged as the most dominant contributors to students' errors, particularly their limited understanding of fundamental statistical concepts. This limitation was evident in students' inability to differentiate between closely related measurement scales, such as interval and ratio scales. The following interview excerpt exemplifies this conceptual difficulty.

*R : Why did you categorize GPA as an ordinal variable?*

*M25 : Because GPA is numerical and shows levels of achievement, I thought it could be ordered, so I classified it as ordinal.*

The response demonstrates a clear conceptual misunderstanding of measurement scales. The student relied primarily on the numerical and ordered nature of GPA, while overlooking the defining characteristics of measurement scales, particularly the concept of an absolute zero point that distinguishes ratio data. This indicates that the error stems from insufficient conceptual knowledge rather than from procedural misapplication, highlighting a fundamental weakness in students' understanding of statistical measurement theory.

Procedural factors were also identified as contributors to students' errors, particularly in tasks that required the use of statistical software. Several students experienced difficulties in selecting and interpreting appropriate SPSS outputs to answer statistical questions. This procedural challenge is illustrated in the following interview excerpt with Student M8:

*R : What output did you use to answer Question 5?*

*M8 : I used a frequency table and a histogram because I assumed these outputs represented the results of a normality test*

The interview excerpt indicates that the student conflated descriptive statistical outputs with inferential test results. While the student demonstrated an awareness of the need to assess data normality, the reliance on frequency tables and histograms reveals a procedural error in navigating and interpreting SPSS output.

In addition to internal factors, external factors such as exam anxiety and time pressure also contributed to students' errors. Several interviewed students reported that limited time during the test led them to rely on intuitive reasoning or shortcut methods instead of applying appropriate statistical procedures. These external pressures exacerbated existing conceptual and procedural weaknesses, leading to a higher incidence of errors. Overall, the findings suggest that students' errors arose from an interaction among insufficient conceptual understanding, limited procedural proficiency, and external constraints during assessment.

## Discussion

The findings of this study indicate that ESTE students experience persistent difficulties in both conceptual understanding and procedural application when solving statistical problems. These difficulties cannot be viewed merely as isolated mistakes, but rather as reflections of how students construct and apply statistical knowledge. From the perspective of conceptual and procedural knowledge theory (Hiebert, 2013), The errors identified suggest that many students rely on procedural knowledge that is weakly connected to underlying conceptual understanding. Conceptual errors, such as misclassifying measurement scales or misinterpreting hypothesis testing outputs, reveal that students often rely on surface-level characteristics of data rather than underlying statistical principles. This supports previous international studies that highlight weak conceptual structures as a major source of error in statistics learning (Legesse et al., 2020), although the present findings further show that such weaknesses are particularly evident when students are required to interpret results rather than perform routine calculations.

These results further illustrate how students struggle to operationalize statistical knowledge in problem-solving contexts. Procedural knowledge, as defined by Hiebert (2013), includes both the use of formal symbolic representations and the algorithms or rules required to complete mathematical tasks. The procedural errors observed in this study—such as incorrect application of formulas, unnecessary data stratification, and inappropriate selection of SPSS outputs—indicate that students often execute procedures mechanically without fully understanding the conditions, assumptions, or purposes underlying those procedures. This pattern supports findings from previous studies showing that procedural errors are frequently linked to weak conceptual foundations rather than mere carelessness or lack of practice (Layn

et al., 2023; Ikram & Rosidah, 2024). In this sense, the present findings reinforce recent international evidence suggesting that procedural proficiency in statistics cannot be developed independently of conceptual understanding, as insufficient conceptual foundations tend to result in fragmented and ineffective procedural execution (Lenz et al., 2024).

These findings can also be interpreted through Nooijen et al. (2024), Ouwehand et al. (2025), who argue that excessive cognitive load exceeds the limited capacity of working memory and interferes with effective learning of complex knowledge. In tasks involving statistical software, students are required to simultaneously interpret output tables, recall statistical concepts, and make statistical decisions, which places a high cognitive demand on working memory. When students' conceptual understanding is fragile, the additional cognitive load imposed by software navigation and output interpretation can overwhelm cognitive resources, resulting in incorrect conclusions. This contrasts with findings from studies conducted in learning environments that provide extensive scaffolding in software use, suggesting that instructional design plays a critical role in determining whether technology supports or hinders students' statistical reasoning (Arifin & Aprisal, 2020; Pujiarti et al., 2024).

In addition to cognitive factors, affective and contextual factors such as exam anxiety and time pressure were found to exacerbate students' errors. Consistent with affective–cognitive interaction frameworks, anxiety consumes cognitive resources that would otherwise be allocated to conceptual reasoning and procedural execution, prompting students to rely on intuition or shortcut strategies during assessments (Ferreira et al., 2025; Uğraş, 2025). While some prior studies emphasize repeated practice as a primary solution for reducing procedural errors, the present findings suggest that practice alone is insufficient unless accompanied by explicit conceptual clarification, reflective activities, and supportive learning environments, particularly in technology-assisted statistical tasks.

Overall, this study contributes to the literature on statistics education by demonstrating that students' errors arise from an interaction between limited conceptual understanding, procedural difficulties, cognitive load, and affective constraints. For prospective elementary school teachers, these findings highlight the importance of instructional approaches that integrate conceptual explanations with procedural practice, provide structured guidance in the use of statistical software, and address affective factors such as anxiety and time management. Strengthening these aspects is essential not only for improving students' statistical performance, but also for preparing them to teach statistical concepts accurately and confidently in future classroom contexts.

## Conclusion

This study identified conceptual and procedural errors made by Elementary School Teacher Education students when solving statistics problems, indicating persistent difficulties in understanding fundamental statistical concepts and in accurately applying appropriate procedures. Conceptual errors reflect weaknesses in students' comprehension of core statistical ideas, while procedural errors indicate challenges in executing formulas, calculations, and statistical software procedures. These findings show that students' statistical problem-solving

performance is influenced not only by technical skills but also by the quality of their conceptual understanding.

This study contributes to statistics education research by clarifying the interaction between conceptual and procedural errors in the context of pre-service elementary teacher education. The findings provide important implications for curriculum design in teacher education programs, highlighting the need to integrate conceptual explanations with procedural practice, offer structured guidance in the use of statistical software, and consider affective factors such as test anxiety and time pressure. Although conducted within a single institutional context, this study offers insights that can inform efforts to improve statistics instruction and better prepare future teachers to understand and teach statistical concepts effectively.

## Acknowledgment

The authors would like to thank the ESTE students of Pasundan University who participated in this research and made a very significant contribution to the smooth running and success of this research.

## Declarations

Conflicts of Interest : The authors declare no conflict of interest.

Generative AI Statement : AI Used for Limited, Non-Substantive Support:
Generative AI tools, such as Grammarly, were employed solely for language editing and minor phrasing enhancements. All conceptualization, analysis, and scholarly content were independently developed and verified by the authors.

Funding Statement : This work received no specific grant from any public, commercial, or not-for-profit funding agency.

Author Contributions : **Siti Maryam Rohimah:** Conceptualization, writing - original draft, editing, and visualization; **Sarah Anida Putri:** Writing - review & editing, formal analysis, and methodology; **Yullys Helsa:** Validation and visualization.

## References

Abu-ghalyoun, O. (2021). Pre-service teachers' difficulties in reasoning about sampling variability. *Educational Studies in Mathematics*, *108*(3), 553–577. https://doi.org/10.1007/s10649-021-10067-8

Arifin, S., & Aprisal. (2020). Analisis tingkat pemahaman konsep statistika mahasiswa calon guru menggunakan two tier test berbasis online [Analysis of pre-service teachers' conceptual understanding of statistics using an online two-tier test]. *Delta: Jurnal Ilmiah Pendidikan Matematika*, *8*(2), 201–208. https://doi.org/10.31941/delta.v8i2.1059

Dodeen, H., & Alqawasmi, A. A. (2025). Exploring statistical anxiety, attitudes, and self-efficacy among social sciences students: The impact of gender, academic progression, and

achievement. *Educational Process: International Journal*, *18*, e2025437. https://doi.org/10.22521/edupij.2025.18.437

Ferreira, R. A., Rodríguez, C., Guzmán, B., Sepúlveda, F., & Peake, C. (2025). The interplay of working memory, vocabulary, and math anxiety in early mathematical learning. In *Journal of Intelligence*, 13(10), 1-20. https://doi.org/10.3390/jintelligence13100125

Hiebert, J. (2013). Conceptual and procedural knowledge: The case of mathematics. In *Conceptual and Procedural Knowledge: The Case of Mathematics*. https://doi.org/10.4324/9780203063538

Ikram, F. Z., & Rosidah. (2024). Kesalahan mahasiswa fakultas keguruan dan ilmu pendidikan dalam menggunakan SPSS [Errors of students in the faculty of teacher training and education in using SPSS]. *Jurnal Media TIK: Jurnal Media Pendidikan Teknik Informatika dan Komputer*, *7*(2), 144–149. https://doi.org/10.59562/mediatik.v7i2.2819

Ismail, E. N., Ramadhani, D. A., Ramadhan, Anie, E. E., Akil, & Azis, A. (2025). Validitas alat ukur dalam evaluasi pembelajaran: Studi faktor yang mempengaruhi validitas alat ukur [Validity of measurement tools in learning evaluation: A study of factors affecting instrument validity]. *Jurnal Kajian Ilmiah Interdisipliner*, *9*(5), 391–400. https://sejurnal.com/pub/index.php/jkii/article/view/7428

Lakshmanan, M. (2022). Common errors in using statistical tools and data presentation. In *Introduction to Basics of Pharmacology and Toxicology: Volume 3: Experimental Pharmacology: Research Methodology and Biostatistics*, 897–910. Springer Nature Singapore. https://doi.org/10.1007/978-981-19-5343-9_63

Layn, M. R., Arsyad, R. Bin, Mulyono, Sira'a, Y., & Kadtabalubun, C. (2023). Analisis kesalahan menyelesaikan soal statistika dan pengolahan data ditinjau dari kemampuan mahasiswa Universitas Muhammadiyah Sorong [Analysis of errors in solving statistical problems and data processing: a review of the abilities of students at Muhammadiyah University of Sorong]. *KAMBIK: Journal of Mathematics Education*, *1*(2), 43–53. https://doi.org/10.33506/jme.v1i2.3068

Legesse, M. Y., Kakoma, L., & Ejigu, T. (2020). Analyzing the effects of mathematical discourse-based instruction on eleventh-grade students' procedural and conceptual understanding of probability and statistics. *Studies in Educational Evaluation, 67*, 100918. https://doi.org/10.1016/j.stueduc.2020.100918

Leng, N., & Meng, C. (2023). Making sense of students' errors in solving problems related to measures of dispersion. *International Journal of Evaluation and Research in Education (IJERE)*, *12*(2), 924–940. https://doi.org/10.11591/ijere.v12i2.24580

Lenz, K., Reinhold, F., & Wittmann, G. (2024). Transitions between conceptual and procedural knowledge profiles. Patterns in understanding fractions and indicators for individual differences. *Learning and Individual Differences*, *116*(102548), 1–12. https://doi.org/10.1016/j.lindif.2024.102548

Lian, L. H., Yew, W. T., & Meng, C. C. (2022). Assessing lower secondary school students' common errors in statistics. *Pertanika Journal of Social Sciences and Humanities*, *30*(3), 1427–1450. https://doi.org/10.47836/pjssh.30.3.26

Mendes, R. A., Loxton, N. J., Stuart, J., Donnell, A. W. O., & Stainer, M. J. (2024). Statistics anxiety or statistics fear? A reinforcement sensitivity theory perspective on psychology students' statistics anxiety, attitudes, and self-efficacy. *European Journal of Psychology of Education*, *39*(3), 2461–2480. https://doi.org/10.1007/s10212-024-00802-z

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A methods sourcebook* (3 ed.). https://www.metodos.work/wp-content/uploads/2024/01/Qualitative-Data-Analysis.pdf

Nisa, R. K., & Dahlan, J. A. (2025). Analyzing students' errors on the topic of statistics using Watson's criteria. Pi-Radian: Journal of Mathematics Education. *Pi-Radian: Journal of*

*Mathematics Education*, *3*(1), 61–76. https://doi.org/10.63214/piradian.v3i1.pp61-76

Nooijen, C. C. A. Van, Koning, B. B. De, Bramer, W. M., Isahakyan, A., Asoodar, M., Kok, E., Merrienboer, J. J. G. Van, & Paas, F. (2024). A cognitive load theory approach to understanding expert–novice differences: scaffolding and working memory limitations. In *Educational Psychology Review*, 36(1). Springer US. https://doi.org/10.1007/s10648-024-09848-3

Ouwehand, K., Lespiau, F., Tricot, A., & Paas, F. (2025). Cognitive load theory: Emerging trends and innovations. *Education Sciences*, *15*, 458. https://doi.org/10.3390/educsci15040458

Pallauta, J. D., Arteaga, P., & Garzón-Guerrero, J. A. (2021). Secondary school students' construction and interpretation of statistical tables. *Mathematics*, *9*(24). https://doi.org/10.3390/math9243197

Parks, J., & Yeh, D. D. (2021). How to Lie with Statistics and Figures. *Surgical Infections*, *22*(6), 611–619. https://doi.org/10.1089/sur.2021.065

Prameshti, N. L., Darmawan, P., & Dejarlo, J. O. (2024). Analysis of students' errors in solving statistics problems based on Newman's Error theory : a study on high school students. *Polyhedron International Journal in Mathematics Education*, *2*(11), 56–63. https://doi.org/10.59965/pijme.v2i2.150

Pujiarti, T., Mahdin, M., & Ilham, I. (2024). Analisis kemampuan pemahaman konsep pada mata kuliah dasar-dasar statistik mahasiswa PGSD STKIP Yapis Dompu [Analysis of conceptual understanding in basic statistics courses for PGSD students at STKIP Yapis Dompu]. *JagoMIPA: Jurnal Pendidikan Matematika dan IPA*, *4*(2), 345–351. https://doi.org/10.53299/jagomipa.v4i2.600

Radke, S. C., Krishnamoorthy, R., Ma, J. Y., & Kelton, M. L. (2023). "Your truth isn't the truth": Data activities and informal inferential reasoning. In *The Journal of Mathematical Behavior*, 69, 1–17. Elsevier Science. https://doi.org/10.1016/j.jmathb.2023.101053

Rohimah, S. M. (2024). *Statistika penelitian pendidikan (analisis manual dan IBM SPSS) [Educational research statistics: Manual and IBM SPSS analysis]*. Rajawali Press.

Romero, E. P. J., Laguerta, M., & Andrade, R. (2023). Perceived statistics self-efficacy, research anxiety, and research confidence of mathematics pre-service teachers in one state university in the Philippines. *Jurnal Pendidikan Progresif*, *13*(2), 708–722. https://doi.org/10.23960/jpp.v13.i2.202343

Sardareh, S. A., Brown, G., & Denny, P. (2025). Statistical software usability for novice research students in the social sciences: An eye-tracking study. *Journal of Statistics and Data Science Education*, *00*(0), 1–25. https://doi.org/10.1080/26939169.2025.2497550

Shamsuddin, M., Mahlan, S. B., Alias, F. A., Hamat, M., & Mohamed, S. A. (2021). Analysis of student error in statistical subject: A case study for online learning. *International Journal of Academic Research in Progressive Education and Development*, *10*(3), 73–83. https://doi.org/10.6007/IJARPED/v10-i3/10714

Shimizu, Y., & Kang, H. (2025). Research on classroom practice and students ' errors in mathematics education: A scoping review of recent developments for 2018-2023. *ZDM Mathematics Education*, *57*, 695–710. https://doi.org/10.1007/s11858-025-01704-0

Silva, P. N., & Sarnecka, B. W. (2025). What do your students struggle with? A survey of statistics instructors. *Journal of Statistics and Data Science Education*, *00*(0), 1–12. https://doi.org/10.1080/26939169.2025.2455560

Sugiyono. (2017). *Metode penelitian kuantitatif, kualitatif, dan R&D [Quantitative, qualitative, and R&D research methods]*. Alfabeta, CV. https://digi-lib.stekom.ac.id/assets/dokumen/ebook/feb_35efe6a47227d6031a75569c2f3f39d44fe2db43_1652079047.pdf

Thanheiser, E., & Mamolo, A. (2024). Introduction to the virtual special issue: Mathematics

that underpins social issues. *Journal of Mathematical Behavior*, *75*, 101176. https://doi.org/10.1016/j.jmathb.2024.101176

Uğraş, H. (2025). Research on mathematics anxiety in primary school: bibliometric analysis and evaluation of trends. *Frontiers in Psychology*, *16*(1545556), 1–21. https://doi.org/10.3389/fpsyg.2025.1545556

Vetten, A. De, Keijzer, R., Schoonenboom, J., & Oers, B. Van. (2023). Pre-service primary school teachers' knowledge during teaching informal statistical inference. *Statistics Education Research Journal*, *22*(1), 1–16. https://doi.org/10.52041/serj.v22i2.424

Winarso, W., & Toheri, T. (2021). An analysis of students' errors in learning mathematical problem solving: The perspective of David Kolb's theory. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *12*(1), 139–150. https://doi.org/10.16949/turkbilmat.753899