JARME Journal of Authentic Research on Mathematics Education
Volume 8, No. 1, January 2026

https://doi.org/ 10.37058/jarme.v3i1.15924 e-ISSN 2655-7762

EXPLORING STUDENTS' CONCEPTUAL AND PROCEDURAL ERRORS IN AKM-BASED CENTRAL TENDENCY TASKS

Hamida Haryati[1], Sukirwan[2*], Dedi Muhtadi[3]

[1,2,3] Pendidikan Matematika, Universitas Siliwangi, Tasikmalaya, Jawa Barat, Indonesia
* Corresponding Author : sukirwan@unsil.ac.id

ARTICLE INFO

Article history
Received: 2025-06-11
Revised: 2026-11-18
Accepted: 2025-01-10

Keywords
This study focuses on *students' errors in statistics*, particularly in solving *AKM-based central tendency tasks*, and examines their implications for *students' numeracy literacy*.

ABSTRACT

This study investigates students' errors in solving Asesmen Kompetensi Minimum (AKM)-based problems on measures of central tendency, with a focus on conceptual understanding, procedural reasoning, and technical accuracy. Employing a qualitative descriptive design, data were collected from 20 eighth-grade students through an AKM-based diagnostic test and semi-structured interviews. The findings reveal that students' errors are predominantly conceptual and procedural, particularly in interpreting the meaning of mean and median, planning solution strategies, and systematically analyzing contextual data representations. Technical errors occurred less frequently and were mainly associated with computational inaccuracy and limited reflective checking. These results indicate that students' engagement with AKM-style statistics tasks is characterized by surface-level processing rather than meaningful statistical reasoning. The study contributes to research on statistical literacy by highlighting the need to integrate conceptual understanding, procedural planning, and reflective practices in numeracy instruction. Implications are discussed for strengthening assessment-oriented learning and supporting students' numeracy development within competency-based curricula.

Penelitian ini bertujuan untuk menganalisis kesalahan siswa dalam menyelesaikan soal pemusatan data berbasis Asesmen Kompetensi Minimum (AKM) ditinjau dari pemahaman konseptual, penalaran prosedural, dan ketelitian teknis. Penelitian ini menggunakan pendekatan kualitatif deskriptif dengan melibatkan 20 siswa kelas VIII melalui tes diagnostik berbasis AKM dan wawancara semi-terstruktur. Hasil penelitian menunjukkan bahwa kesalahan siswa didominasi oleh kesalahan konseptual dan prosedural, terutama dalam memahami makna rata-rata dan median, merencanakan langkah penyelesaian, serta menganalisis data kontekstual secara sistematis. Kesalahan teknis muncul dalam frekuensi yang lebih rendah dan umumnya berkaitan dengan ketidaktelitian dalam perhitungan dan kurangnya refleksi terhadap hasil kerja. Temuan ini mengindikasikan bahwa keterlibatan siswa dalam soal statistika berbasis AKM masih bersifat dangkal dan belum mencerminkan penalaran statistik yang bermakna. Penelitian ini memberikan kontribusi pada kajian literasi numerasi dengan menegaskan pentingnya integrasi pemahaman konsep, strategi prosedural, dan praktik reflektif dalam pembelajaran statistika untuk mendukung pencapaian tujuan AKM.

How to Cite:
Haryati, H., Sukirwan, & Muhtadi, D. (2026). Exploring Students' Conceptual And Procedural Errors In Akm-Based Central Tendency Tasks. *Journal of Authentic Research on Mathematics Education, 8*(1), 11-25. https://doi.org/10.37058/jarme.v3i1.15924

INTRODUCTION

The growing emphasis on applying mathematics in real-world contexts has increased attention to statistical literacy in education. Statistical literacy encompasses the ability to understand, interpret, and use quantitative information for data-based decision-making, with measures of central tendency such as mean and median playing a crucial role in understanding data distributions (Gal, 2000; Schield, 2011). However, research has shown that students often master computational procedures without adequate conceptual understanding, leading to difficulties in interpreting contextual data and applying appropriate solution strategies (Karaca & Ay, 2024). Therefore, curriculum design and instructional strategies should prioritize meaningful engagement with real data and statistical reasoning to strengthen students' statistical literacy (Burrill, 2020).

Students frequently face challenges in learning statistics, particularly when engaging with context-based tasks, due to a tendency to rely on procedural rules rather than conceptual reasoning. This often results in misinterpretation of data and inappropriate strategy selection. Previous studies have shown that students struggle to interpret data representations, differentiate between statistical measures, and select relevant evidence to support claims, reflecting weak metacognitive regulation in statistical reasoning (Dijke-Droogers et al., 2021, 2024; Frischemeier & Schnell, 2021; Kuhn & Modrek, 2021). To address these issues, various instructional approaches have been recommended, including authentic, data-driven assessments, interactive learning tools, and realistic problem situations that integrate multiple representations. Such approaches have been found to enhance students' engagement, conceptual understanding, and statistical reasoning (Conlon & Wilson, 2025; Pai, 2024; Orozco-Rodríguez et al., 2023). Nevertheless, the effectiveness of statistics instruction largely depends on balancing conceptual and procedural instruction and maintaining context sensitivity, which requires adaptive instructional design to accommodate students' diverse abilities (Lee, 2024; Supply et al., 2023).

In the Indonesian context, the implementation of the Minimum Competency Assessment (AKM) aims to strengthen numeracy literacy through context-based reasoning and problem-solving. However, students still encounter difficulties with tasks requiring data analysis and statistical reasoning, highlighting the need for a deeper understanding of how students conceptualize statistical ideas in assessment-based learning environments. Research suggests that integrating realistic problems and multiple data representations can enhance students' statistical reasoning (Orozco-Rodríguez et al., 2023), while early exposure to data literacy fosters the development of critical thinking skills in numeracy tasks (Sickler et al., 2024). Furthermore, innovative learning designs that promote higher-order thinking skills (Liu et al., 2024) and contextual project-based learning, such as *STEAM-integrated ethnomathematics*, have been shown to improve students' problem-solving abilities (Afifah et al., 2025). From an evaluation perspective, the use of domainappropriate assessment instruments and robust analytical frameworks enables more accurate measurement of students' reasoning (Тарасова et al., 2025; Groth & Choi, 2023), while also considering the influence of cognitive and sociodemographic factors on numeracy performance (Knabbe et al., 2024).

Against this backdrop, investigating students' errors in solving AKM-based problems on measures of central tendency is essential to advancing statistical and numeracy literacy. Such analysis not only identifies the types of errors but also uncovers the underlying cognitive causes. By analyzing students' written responses and reasoning processes, this study explores the relationships among conceptual understanding, procedural reasoning, and technical accuracy in statistical problem solving, addressing the literature's call to move beyond mere error categorization. Recent studies continue to show that conceptual, procedural, and technical errors are prevalent in mathematics learning, encompassing conceptual comprehension, solution planning, and computational precision (Sihotang et al., 2025; Sinaga et al., 2025; Elagha & Pellegrino, n.d.; Lestari et al., 2025). These findings underscore the importance of error-based learning interventions, adaptive scaffolding, and statistical visualization to enhance students' reasoning (Sihotang et al., 2025). Additionally, the use of analytical frameworks such as the Toulmin argumentation model and SOLO taxonomy provides deeper insights into the quality of students' reasoning and argumentation (Groth & Choi, 2023), while technology integration and attention to assessment equity contribute to more holistic development of statistical literacy (Gök & Gök, 2024; McCracken et al., 2024).

METHODS

Research Design

This study employed a qualitative descriptive design to examine students' errors in solving AKM-based problems on measures of central tendency. This approach enables an in-depth understanding of students' conceptual, procedural, and technical reasoning in contextualized numeracy tasks. Qualitative analysis is suitable for identifying error patterns and exploring students' reasoning processes through their written responses (Creswell & Poth, 2018; Garfield & Ben-Zvi, 2008). Previous research in mathematics education supports the use of such methods to analyze misconceptions and statistical reasoning, particularly in measures of central tendency (Groth & Bergner, 2013; Jacobbe & Carvalho, 2011), aligning with AKM's emphasis on literacy, reasoning, and contextual problem-solving (Kemendikbud, 2020; Stacey, 2020).

Participants

The participants were 20 eighth-grade students from a public junior high school in Tasikmalaya, Indonesia. A purposive sampling technique was used to select students who had completed instruction on measures of central tendency and were familiar with context-based assessment tasks. The sample size was considered sufficient for qualitative analysis aimed at identifying recurring error patterns rather than statistical generalization. Students represented varied levels of mathematical achievement to ensure heterogeneity of responses.

Instruments

Two primary instruments were utilized in this study: an AKM-based diagnostic test and semi-structured interviews. The diagnostic test comprised three contextual problems developed in alignment with the AKM numeracy framework, focusing on the concepts of mean, median, and data interpretation using tables and graphs. The items were constructed to elicit students' reasoning processes and potential misconceptions rather than merely assess answer accuracy. To establish content validity, the test was reviewed by two experts in mathematics education and one experienced secondary mathematics teacher, with revisions made based on their feedback concerning clarity, contextual appropriateness, and construct alignment. Furthermore, semi-structured interviews were conducted with selected students who demonstrated representative error patterns in their written responses. These interviews sought to further explore students' reasoning, decision-making processes, and interpretations of the task requirements, thereby supporting a deeper understanding of their statistical thinking.

Data Collection Procedure

Data collection was conducted in three stages. First, students completed the AKM-based diagnostic test in a regular classroom setting. Second, students' written responses were collected and initially reviewed to identify recurring error patterns. Third, follow-up interviews were conducted to gain deeper insights into students' reasoning and to triangulate the findings from the written data.

Data Analysis

Data analysis employed an iterative thematic approach (Braun & Clarke, 2006) to examine students' conceptual, procedural, and technical errors. Students' written responses were open-coded and categorized based on similarities in reasoning patterns (Miles, Huberman, & Saldaña, 2014). Interview data were used to refine these categories and provide insights into students' thinking (Cohen, Manion, & Morrison, 2018). The analysis was inductive, allowing themes to emerge from the data while guided by established frameworks on mathematical error analysis and statistical literacy (Newman, 1977; Radatz, 1980; Garfield & Ben-Zvi, 2008).

Trustworthiness

To ensure the trustworthiness of findings, several strategies were applied (Lincoln & Guba, 1985; Shenton, 2004). Credibility was established through triangulation of test and interview data (Cohen, Manion, & Morrison, 2018). Dependability was supported by documenting analytical procedures and coding decisions (Miles, Huberman, & Saldaña, 2014). Confirmability was enhanced via peer debriefing with another mathematics education researcher (Creswell & Poth, 2018). Ethical standards were maintained through informed consent and participant confidentiality (BERA, 2018).

RESULTS

This study identified distinct patterns of students' errors in solving AKM-based problems on measures of central tendency. Through analysis of students' written responses and follow-up interviews, three dominant categories of errors emerged: conceptual errors, procedural errors, and technical errors. These errors appeared consistently across the three AKM tasks involving mean, data interpretation, and median.

1. Conceptual Errors

Misconceptions in statistical concepts, particularly in calculating the mean and median, remain common in mathematics education. Students frequently commit errors such as computing an overall mean instead of a group-based mean, indicating limited contextual understanding of statistical measures ("Error in Figure 1," 2022; "Errors in Table 1," 2022). These mistakes reveal persistent gaps in statistical learning, where students tend to memorize procedures rather than grasp conceptual principles ("Errors in Table," 2023; "Error in Figures," 2022). As a result, their data interpretation and decision-making can become inaccurate both in academic settings and real-world applications (Ramdani & Audemard-Verger, 2023; "Error in Table and Figure," 2022). Addressing this issue requires instructional practices that emphasize conceptual understanding through authentic contexts, problem-based learning, and continuous assessment and feedback, enabling students to apply statistical measures accurately and meaningfully ("Error in Results," 2023; "Error in Discussion Section," 2022).

Difficulties in determining the median—especially when handling datasets with an even number of values—highlight students' reliance on procedural knowledge rather than conceptual understanding. This challenge aligns with findings from other mathematical domains, such as mathematical induction, where inadequate conceptual foundations lead to errors in procedure selection and interpretation (Sinaga et al., 2025). The median should be viewed not merely as a computed value but as a positional measure that conveys contextual meaning across various statistical situations (Goibert et al., 2022). Procedural challenges in finding the median for even-numbered data, which require averaging two middle values, can be reduced through explicit instruction and purposeful practice (Dey & Chaudhuri, 2024). Effective strategies such as scaffolding and collaborative group discussions are recommended to reinforce conceptual understanding (Sinaga et al., 2025). Furthermore, introducing robust estimators and connecting the median to real-world applications, including clinical research, can deepen students' appreciation of its practical relevance and limitations (Hussain et al., 2024; Xiang et al., 2023). Finally, integrating data modeling and informal statistical reasoning may further enhance students' conceptual grasp and their ability to apply median concepts effectively in diverse contexts (Kazak et al., 2021).

2. Procedural Errors

Procedural errors in problem-solving tasks represent a common challenge encountered by many students, particularly those related to planning, sequencing, and systematic implementation of solution steps. Students often experience difficulties in identifying relevant data and determining appropriate problem-solving stages, which in turn leads to inaccurate or unsupported conclusions. These challenges arise across various types of tasks—including arithmetic operations, data interpretation, and statistical computations—indicating that procedural weaknesses are cross-contextual rather than confined to a single mathematical topic.

In arithmetic operations, students frequently perform calculations without first determining the quantities that should be averaged, reflecting deficiencies in planning and step sequencing (Sinaga et al., 2025; Chinofunga et al., 2024). In data interpretation tasks, particularly those involving air quality index graphs, students tend to rely on visual impressions rather than systematically counting category frequencies, resulting in inaccurate interpretations (Albarracín et al., 2021; Chang et al., 2024). Similar procedural errors are also evident in statistical tasks, such as determining the median, where students often select a middle value without arranging the data in order—demonstrating insufficient understanding of proper data analysis procedures (Peng et al., 2021).

These procedural errors are closely associated with weak metacognitive skills, particularly in planning strategies, monitoring problem-solving processes, and selecting relevant evidence (Kuhn & Modrek, 2021). Therefore, instructional strategies that explicitly foster procedural and metacognitive development are essential. Approaches such as adaptive scaffolding, collaborative group discussions, and the use of procedural flow diagrams can support students in organizing problem-solving steps more systematically and generating accurate solutions (Sinaga et al., 2025; Chinofunga et al., 2024).

3. Technical Errors

Technical errors in arithmetic calculations—such as mistakes in addition, multiplication, and division—are a common issue in mathematics learning, even among students who have demonstrated a sound understanding of the underlying concepts. These errors frequently manifest as operational inaccuracies and symbolic manipulation mistakes, occurring both in basic arithmetic operations and in more complex algebraic computations (Lestari et al., 2025; Sinaga et al., 2025). Contributing factors include carelessness, time pressure, task complexity involving multiple procedural steps, and the lack of reflective verification habits, all of which significantly affect the accuracy of final answers (Sinaga et al., 2025; Wennberg-Capellades et al., 2022).

Although technical in nature, such errors present valuable opportunities for instructional improvement when appropriately addressed. Research has emphasized the use of scaffolding and collaborative group discussions to enhance students' accuracy and procedural understanding, as well as targeted instruction focusing on frequently occurring error types (Sinaga et al., 2025; Wennberg-Capellades et al., 2022). Furthermore, encouraging students to analyze their own mistakes and providing constructive feedback have proven effective in fostering both computational accuracy and confidence, aligning with the learning-from-errors approach, which views error detection and correction as integral components of the learning process (Zhang & Fiorella, 2022).

4. Summary of Error Patterns

Overall, the analysis reveals that procedural errors were the most dominant, followed by conceptual errors, while technical errors occurred sporadically. Table X summarizes the distribution of error types across the three AKM tasks.

Table 1. Summary of Students' Error Types in AKM-Based Central Tendency Tasks

Task	Dominant Error Type	Description of Error Pattern
1	Conceptual & Procedural	Misinterpretation of mean per group and incorrect averaging steps
2	Procedural	Failure to systematically classify and count AQI categories
3	Conceptual & Procedural	Incorrect identification of median due to unsorted data

These findings indicate that students' difficulties in AKM-based statistics tasks stem not only from computational weaknesses but also from limited conceptual understanding and inadequate procedural planning when engaging with contextual data problems.

DISCUSSION

The findings of this study indicate that students' difficulties in solving AKM-based central tendency tasks are not merely computational but are rooted in limited conceptual understanding, procedural reasoning, and data interpretation skills. This aligns with international research reporting that students frequently exhibit both conceptual and procedural errors across various contexts of statistical and mathematical learning (Koerfer & Gregorcic, 2024; Sinaga et al., 2025; Kranz et al., 2022). The integration of realistic problem situations and visual learning tools has been shown to strengthen statistical reasoning by fostering coordination among multiple data representations, although individual differences in ability remain evident (Orozco-Rodríguez et al., 2023; Oslington et al., 2023). Consequently, effective statistics instruction should emphasize conceptual discussions, guided scaffolding, and diagnostic approaches that help students connect statistical concepts with real-world contexts in a more meaningful and enduring way (Säfström et al., 2023).

Conceptual Understanding and Statistical Meaning-Making

Studies on students' conceptual misunderstandings of statistical measures such as the mean and median reveal a fundamental gap between formulaic application and representational understanding of data. Students often perceive statistical measures as mechanical procedures, as reflected in their misinterpretations of terms such as "group mean" and their limited conception of the median as merely the "middle number," without considering data ordering or distributional context. These findings align with international research highlighting students' fragmented conceptual knowledge and the weak connections they form between statistical concepts and the meanings represented in data (Sinaga et al., 2025; Schreiter & Vogel, 2024).

These difficulties are closely associated with low levels of statistical literacy, particularly in understanding data distribution as a whole. Students frequently fail to integrate information about variability and distributional patterns, both of which are essential for meaningful interpretation of mean and median values. Furthermore, data visualization practices may inadvertently reinforce misconceptions—for example, through bar or line graphs that prompt interpretive errors such as the *Bar-Tip Limit Error*, *Dichotomization Fallacy*, or *mean estimation bias* in high-variability data (Wilmer & Kerns, 2022; Moritz et al., 2023). This suggests that while visualizations can appear intuitive, they do not necessarily ensure accurate conceptual understanding.

The educational implications of these findings underscore the need for a more integrated approach to statistics education, emphasizing conceptual scaffolding, reflective discussion, and exploration of data distribution meaning. Instructional strategies that foreground variability through informative visualizations, such as *jitter plots* or *prediction intervals*, can help students develop more accurate interpretations of statistical measures (Holder & Xiong, 2022). By systematically linking theoretical understanding with contextual application, statistics instruction has the potential to enhance students' statistical literacy and reduce the persistence of long-standing misconceptions.

Procedural Reasoning and Problem-Solving Planning

Students' struggles in solving AKM-style contextual problems reveal a significant gap in their ability to plan and implement effective problem-solving strategies. Procedural errors frequently emerge due to disorganization in sequencing solution steps and an overreliance on heuristic shortcuts and visual impressions, rather than systematic data handling and analytical reasoning (Sinaga et al., 2025; Chang et al., 2024). These findings are consistent with international studies showing that students often possess fragmented procedural skills without the accompanying strategic competence necessary to apply those procedures effectively in complex problem contexts (Chinofunga et al., 2024).

This weakness is further reflected in students' limited procedural fluency and data visualization literacy, particularly in tasks that demand the integration of contextual and statistical reasoning (Schulz, 2023; Groth & Choi, 2023). To address these challenges, prior research recommends the use of scaffolding, collaborative group discussions, procedural flow diagrams, and adaptive instruction tailored to students' planning typologies to enhance their strategic and metacognitive competencies (Sinaga et al., 2025; Chinofunga et al., 2024; Zhang et al., 2024). Such approaches have the potential to create a more reflective and flexible learning environment, enabling students to meaningfully integrate procedures, strategies, and reasoning in solving contextualized mathematical problems (Nicolay et al., 2023).

Technical Accuracy and Reflective Thinking

Although technical errors were less frequent, their presence underscores the importance of accuracy and reflective checking in numeracy tasks (Goos, Dole, & Geiger, 2012). Errors in basic arithmetic operations occurred even when students demonstrated partial conceptual and procedural understanding, indicating gaps between knowledge and execution (Clements & Sarama, 2014; Verschaffel, Greer, & De Corte, 2000). This finding aligns with research emphasizing the role of metacognitive regulation—particularly self-monitoring and verification—in supporting numeracy and mathematical problem solving (Schoenfeld, 1992; Desoete, 2007).

The occurrence of technical errors suggests that students may not consistently engage in reflective practices during problem solving (Panaoura, 2017). Without structured opportunities to monitor and evaluate their solutions, learners remain prone to avoidable computational mistakes that diminish overall performance in assessment contexts (Efklides, 2011; Stacey, 2020).

Implications for Statistical Literacy and AKM-Oriented Learning

Taken together, the dominance of procedural and conceptual errors indicates that students' engagement with AKM-based statistics tasks reflects surface-level processing rather than deep statistical reasoning (Leavy & Hourigan, 2021; Prodromou, 2020). From a statistical literacy perspective, this suggests that students have not yet developed the ability to interpret, reason with, and critically evaluate data within contextual situations (Ben-Zvi & Makar, 2021; Watson & English, 2022).

The findings further imply that current instructional practices may overemphasize procedural execution and formulaic computation, offering limited opportunities for students to engage in conceptual exploration and meaning-making of statistical measures (Garfield, Ben-Zvi, & Zieffler, 2020; Batanero & Díaz, 2021). In the context of Asesmen Kompetensi Minimum (AKM), which assesses students' ability to apply mathematics in real-life contexts, such instructional orientations appear insufficient to achieve the intended numeracy and reasoning outcomes (Stacey, 2020; Widjaja, 2021; Sumartojo et al., 2022).

Toward Conceptually Driven and Contextualized Instruction

Taken together, the dominance of procedural and conceptual errors indicates that students' engagement with AKM-based statistics tasks reflects surface-level processing rather than deep statistical reasoning (Leavy & Hourigan, 2021; Prodromou, 2020). From a statistical literacy perspective, this suggests that students have not yet developed the ability to interpret, reason with, and critically evaluate data within contextual situations (Ben-Zvi & Makar, 2021; Watson & English, 2022).

The findings further imply that current instructional practices may overemphasize procedural execution and formulaic computation, offering limited opportunities for students to engage in conceptual exploration and meaning-making of statistical measures (Garfield, Ben-Zvi, & Zieffler, 2020; Batanero & Díaz, 2021). In the context of Asesmen Kompetensi Minimum (AKM), which assesses students' ability to apply mathematics in real-life contexts, such instructional orientations appear insufficient to achieve the

intended numeracy and reasoning outcomes (Stacey, 2020; Widjaja, 2021; Sumartojo et al., 2022).

CONCLUSION

This study reveals that students' errors in solving AKM-based central tendency problems are predominantly conceptual and procedural, rather than merely computational. Students frequently misinterpret the meaning of mean and median and apply inappropriate strategies when engaging with contextual data tasks, indicating limited statistical reasoning and surface-level numeracy processing (Leavy & Hourigan, 2021; Ben-Zvi & Makar, 2021; Watson & English, 2022). These findings contribute to the growing body of literature on statistical literacy, emphasizing the critical role of conceptual understanding and procedural planning in assessment-based learning contexts (Batanero & Díaz, 2021; Prodromou, 2020).

Practically, the results highlight the need for instructional approaches that integrate conceptual explanation, systematic problem solving, and reflective verification to foster students' numeracy development aligned with AKM objectives (Stacey, 2020; Widjaja, 2021; Sumartojo et al., 2022). Future research should explore instructional interventions that strengthen statistical reasoning and metacognitive awareness across diverse learning contexts (Panaoura, 2020; Rach, Ufer, & Heinze, 2020).

REFERENCES

Afifah, A. D. N., Indiati, I., & Rahmawati, N. D. (2025). Keefektifan model project based learning terintegrasi steam berbasis etnomatematika terhadap kemampuan berpikir kreatif matematis ditinjau dari efikasi diri siswa. *Science*, *5*(3), 1437–1446. https://doi.org/10.51878/science.v5i3.6754

Antoniassi, M., & Dias, F. A. da S. (2023). Letramento Estatístico: Uma Revisão de Literatura com Artigos da Base de Dados ERIC. *Jornal Internacional de Estudos Em Educação Matemática*, *15*(3), 304–313. https://doi.org/10.17921/2176-5634.2022v15n3p304-313

Aziz, A. M., & Rosli, R. (2021). *A systematic literature review on developing students' statistical literacy skills*. *1806*(1), 012102. https://doi.org/10.1088/1742-6596/1806/1/012102

Batanero, C., & Díaz, C. (2021). Developing statistical literacy: What can we learn from research? *Statistics Education Research Journal, 20*(2), 1–18. https://doi.org/10.52041/serj.v20i2.304

Ben-Zvi, D., & Makar, K. (2021). Preparing students for reasoning about uncertainty in the data era. *Educational Studies in Mathematics, 108*(3), 431–449. https://doi.org/10.1007/s10649-021-10074-7

Bergner, R. M. (2023). Conceptual misunderstandings in mainstream scale construction: Suggestions for a better approach to concepts. *Theory & Psychology*, 095935432311776. https://doi.org/10.1177/09593543231177696

British Educational Research Association (BERA). (2018). *Ethical guidelines for educational research* (4th ed.). London: BERA.

Burrill, G. (2020). *Statistical literacy and quantitative reasoning: rethinking the curriculum*. https://doi.org/10.52041/srap.20104

Chang, H.-Y., Chang, Y.-J., & Tsai, M.-J. (2024). Strategies and difficulties during students' construction of data visualizations. *International Journal of STEM Education*. https://doi.org/10.1186/s40594-024-00463-w

Clements, D. H., & Sarama, J. (2014). *Learning and teaching early math: The learning trajectories approach.* Routledge.

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). Routledge.

Conlon, B., & Wilson, S. (2025). Enhancing statistics education through student-generated data and authentic assessment. *Assessment & Evaluation in Higher Education*, 1–20. https://doi.org/10.1080/02602938.2025.2553894

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches.* SAGE Publications.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design.* SAGE.

Desoete, A. (2007). Evaluating and improving the mathematics teaching–learning process through metacognition. *Electronic Journal of Research in Educational Psychology, 5*(3), 705–730.

Dey, A., & Chaudhuri, P. (2024). Quantile processes and their applications in finite populations. *Annals of Statistics*, *52*(5). https://doi.org/10.1214/24-aos2432

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning. *Educational Psychologist, 46*(1), 6–25.

Elagha, N., & Pellegrino, J. W. (n.d.). Understanding error patterns in students' solutions to linear function problems to design learning interventions. *Learning and Instruction*. https://doi.org/10.1016/j.learninstruc.2024.101895

Error in Figure 1. (2022). *JAMA Network Open*, *5*(6), e2221479. https://doi.org/10.1001/jamanetworkopen.2022.21479

Errors in Table. (2023). *JAMA Dermatology*. https://doi.org/10.1001/jamadermatol.2023.4273

Frischemeier, D., & Schnell, S. (2021). Statistical investigations in primary school – the role of contextual expectations for data analysis. *Mathematics Education Research Journal*, 1–26. https://doi.org/10.1007/S13394-021-00396-5

Gal, lddo. (2000). *Statistical Literacy: Conceptual and Instructional issues* (pp. 135–150). Springer, Dordrecht. https://doi.org/10.1007/0-306-47221-X_8

Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning.* Springer.

Garfield, J., Ben-Zvi, D., & Zieffler, A. (2020). *Developing students' statistical reasoning: Connecting research and teaching practice.* Springer.

Goibert, M., Cl'emenccon, S., Irurozki, E., & Mozharovskyi, P. (2022). Statistical Depth Functions for Ranking Distributions: Definitions, Statistical Learning and Applications. *International Conference on Artificial Intelligence and Statistics*, 10376–10406.

Gök, S., & Gök, S. (2024). How do students reason about statistical sampling with computer simulations? An integrative review from a grounded cognition perspective. *Cognitive*

Research: Principles and Implications, *9*(1). https://doi.org/10.1186/s41235-024-00561-x

Goos, M., Dole, S., & Geiger, V. (2012). *Numeracy across the curriculum: Research-based strategies for enhancing teaching and learning.* ACER Press.

Groth, R. E., & Bergner, J. A. (2013). Building on students' informal inferential reasoning. *Statistics Education Research Journal, 12*(2), 82–108.

Groth, R. E., & Choi, Y. (2023). A method for assessing students' interpretations of contextualized data. *Educational Studies in Mathematics*, 1–18. https://doi.org/10.1007/s10649-023-10234-z

Holder, E., & Xiong, C. (2022). Dispersion vs Disparity: Hiding Variability Can Encourage Stereotyping When Visualizing Social Outcomes. *IEEE Transactions on Visualization and Computer Graphics*, *29*, 624–634. https://doi.org/10.1109/TVCG.2022.3209377

Hussain, M. A., Javed, M., Zohaib, M., Shongwe, S. C., Awais, M., Zaagan, A. A., & Irfan, M. (2024). Estimation of population median using bivariate auxiliary information in simple random sampling. *Heliyon*. https://doi.org/10.1016/j.heliyon.2024.e28891

Jacobbe, T., & Carvalho, C. (2011). Teachers' understanding of averages. *Educational Studies in Mathematics, 76*(2), 155–167.

Karaca, Ş. A., & Ay, Z. S. (2024). Investigation of Eighth Grade Students' Performance on Tasks Involving Statistical Thinking About Measures of Central Tendency. *Participatory Educational Research*, *12*(1), 18–42. https://doi.org/10.17275/per.25.2.12.1

Kazak, S., Fujita, T., & Pifarré Turmo, M. (2021). Students' informal statistical inferences through data modeling with a large multivariate dataset. *Mathematical Thinking and Learning*, 1–21. https://doi.org/10.1080/10986065.2021.1922857

Kementerian Pendidikan dan Kebudayaan (Kemendikbud). (2020). *Panduan Asesmen Kompetensi Minimum (AKM).* Jakarta: Pusat Asesmen dan Pembelajaran.

Knabbe, A., Leiss, D., & Ehmke, T. (2024). Reality-Based Tasks with Complex-Situations: Identifying Sociodemographic and Cognitive Factors for Solution. *International Journal of Science and Mathematics Education*. https://doi.org/10.1007/s10763-024-10463-5

Koerfer, E., & Gregorcic, B. (2024). Exploring student reasoning in statistical mechanics: Identifying challenges in problem-solving groups. *Physical Review*. https://doi.org/10.1103/physrevphyseducres.20.010105

Kranz, J. N., Baur, A., & Möller, A. (2022). Learners' challenges in understanding and performing experiments: a systematic review of the literature. *Studies in Science Education*, *59*, 321–367. https://doi.org/10.1080/03057267.2022.2138151

Kuhn, D., & Modrek, A. S. (2021). Choose Your Evidence. *Science Education*, 1–11. https://doi.org/10.1007/S11191-021-00209-Y

Leavy, A., & Hourigan, M. (2021). Examining statistical reasoning in middle school classrooms: Challenges and opportunities. *International Journal of Science and Mathematics Education, 19*(6), 1225–1242. https://doi.org/10.1007/s10763-020-10086-9

Lee, J. (2024). Concept-Focused and Procedure-Focused Instruction on the Algebra Performance of Grade 9 Students With and Without Mathematics Difficulty. *Journal* *of Learning Disabilities*, 222194241249960. https://doi.org/10.1177/00222194241249960

Lestari, M. I., Lusiana, L., & Wahyuningsih, S. (2025). Analisis kesulitan belajar matematika berbasis masalah pada materi operasi hitung perkalian dan pembagian. *Science*, *5*(3), 1285–1293. https://doi.org/10.51878/science.v5i3.6670

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry.* SAGE.

Liu, J., Liu, Z., Wang, C., Xu, Y., Chen, J., & Cheng, Y. (2024). K-12 students' higher-order thinking skills: Conceptualization, components, and evaluation indicators. *Thinking Skills and Creativity*. https://doi.org/10.1016/j.tsc.2024.101551

Marchy, F., & Juandi, D. (2023). Student's Statistical Literacy Skills (1980-2023): A Systematic Literature Review with Bibliometric Analysis. *Journal of Education and Learning Mathematics Research*, *4*(1), 31–45. https://doi.org/10.37303/jelmar.v4i1.105

McCracken, M., Bostic, J. D., & Folger, T. (2024). Middle School Students' Conceptualizations and Reasoning about the Fairness of Math Tests. *Techtrends*. https://doi.org/10.1007/s11528-024-00988-5

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE.

Moritz, D., Padilla, L., Nguyen, F., & Franconeri, S. (2023). Average Estimates in Line Graphs Are Biased Toward Areas of Higher Variability. *IEEE Transactions on Visualization and Computer Graphics*, 1–10. https://doi.org/10.1109/tvcg.2023.3326589

Musarurwa, D. C., Chigeza, P., & Taylor, S. (2024). How can procedural flowcharts support the development of mathematics problem-solving skills? *Mathematics Education Research Journal*. https://doi.org/10.1007/s13394-024-00483-3

Nicolay, B. F., Krieger, F., Kuhn, J.-T., Graesser, A. C., Ifenthaler, D., Baker, R., & Greiff, S. (2023). Unsuccessful and successful complex problem solvers – A log file analysis of complex problem solving strategies across multiple tasks. *Intelligence*. https://doi.org/10.1016/j.intell.2023.101793

Orozco-Rodríguez, C., Palafox González, A., & Valenzuela García, C. (2023). A realistic situation, proposed from data generated within the GeOrder Simulator to elicit the statistical reasoning. *Heliyon*, *9*. https://doi.org/10.1016/j.heliyon.2023.e19330

Orozco-Rodríguez, C., Palafox González, A., & Valenzuela García, C. (2023). A realistic situation, proposed from data generated within the GeOrder Simulator to elicit the statistical reasoning. *Heliyon*, *9*. https://doi.org/10.1016/j.heliyon.2023.e19330

Oslington, G., Mulligan, J., & Van Bergen, P. (2023). Shifts in students' predictive reasoning from data tables in years 3 and 4. *Mathematics Education Research Journal*. https://doi.org/10.1007/s13394-023-00460-2

Pai, G. (2024). Using Formative Assessment and Feedback from Student Response Systems (SRS) to Revise Statistics Instruction and Promote Student Growth for All. *Journal of Statistics and Data Science Education*. https://doi.org/10.1080/26939169.2024.2321241

Panaoura, A. (2020). Metacognitive regulation in problem solving: Implications for mathematics education. *Frontiers in Psychology*, *11*, 575089. https://doi.org/10.3389/fpsyg.2020.575089

Peng, R. D., Chen, A., Bridgeford, E. W., Leek, J. T., & Hicks, S. C. (2021). Diagnosing Data Analytic Problems in the Classroom. *Journal of Statistics and Data Science Education*. https://doi.org/10.1080/26939169.2021.1971586

Prodromou, T. (2020). Students' statistical reasoning with data: Emerging perspectives in the era of big data. *Mathematics Education Research Journal, 32*(3), 371–393. https://doi.org/10.1007/s13394-019-00278-1

Rach, S., Ufer, S., & Heinze, A. (2020). Learning from errors: Effects of teachers' training on students' attitudes toward errors. *Journal of Mathematical Behavior, 57*, 100732. https://doi.org/10.1016/j.jmathb.2019.100732

Ramdani, Y., & Audemard-Verger, A. (2023). Drs. Ramdani and Audemard-Verger reply. *The Journal of Rheumatology*, jrheum.221295. https://doi.org/10.3899/jrheum.221295

Säfström, A. I., Lithner, J., Palm, T., Palmberg, B., Sidenvall, J., Andersson, C., Boström, E., & Granberg, C. (2023). Developing a diagnostic framework for primary and secondary students' reasoning difficulties during mathematical problem solving. *Educational Studies in Mathematics*. https://doi.org/10.1007/s10649-023-10278-1

Schield, M. (2011). Statistical literacy: A new mission for data producers. *Statistical Journal of the IAOS*, *27*, 173–183. https://doi.org/10.3233/SJI-2011-0732

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 334–370). Macmillan.

Schreiter, S., & Vogel, M. (2024). Students' local vs. global views of data distributions: a cross-grade-level analysis using eye-tracking. *Educational Studies in Mathematics*. https://doi.org/10.1007/s10649-024-10352-2

Schulz, A. (2023). Assessing student teachers' procedural fluency and strategic competence in operating and mathematizing with natural and rational numbers. *Journal of Mathematics Teacher Education*. https://doi.org/10.1007/s10857-023-09590-7

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information, 22*(2), 63–75.

Sickler, J., Lentzner, M. E., Goldsmith, L. T., Brase, L. E., & Kochevar, R. (2024). Reasoning about data in elementary school: student strategies and strengths when reasoning with multiple variables. *International Journal of Science Education*. https://doi.org/10.1080/09500693.2023.2298214

Sihotang, H., Sitindaon, D. M., Saing, N. M. T., Silalahi, L. G. L., Sinaga, D., & Simanullang, M. C. (2025). Analisis miskonsepsi mahasiswa dalam menyelesaikan soal supremum dan infimum berdasarkan teori newman. *Science*, *5*(2), 591–606. https://doi.org/10.51878/science.v5i2.5090

Sinaga, D., Silalahi, L. G. L., Saing, N. M. T., & Manurung, S. L. (2025). Analisis kesalahan mahasiswa pendidikan matematika dalam memahami dan menyelesaikan soal induksi matematika pada materi pengantar grup berdasarkan teori kastolan. *Science*, *5*(1), 350–362. https://doi.org/10.51878/science.v5i1.4658

Stacey, K. (2020). Mathematics teaching and assessment to support mathematical literacy for all. *Journal on Mathematics Education, 11*(3), 405–420. https://doi.org/10.22342/jme.11.3.12944.405-420

Sumartojo, S., Suryani, N., & Wulandari, R. (2022). Analysis of students' numeracy skills in solving AKM-like problems. *International Journal of Evaluation and Research in Education, 11*(2), 874–883. https://doi.org/10.11591/ijere.v11i2.22894

Supply, A.-S., Vanluydt, E., Van Dooren, W., & Onghena, P. (2023). Out of proportion or out of context? Comparing 8- to 9-year-olds' proportional reasoning abilities across fair-sharing, mixtures, and probability contexts. *Educational Studies in Mathematics*, *113*(3), 371–388. https://doi.org/10.1007/s10649-023-10212-5

van Dijke-Droogers, M., Drijvers, P., & Bakker, A. (2021). Introducing Statistical Inference: Design of a Theoretically and Empirically Based Learning Trajectory. *International Journal of Science and Mathematics Education*, 1–24. https://doi.org/10.1007/S10763-021-10208-8

van Dijke-Droogers, M., Drijvers, P., & Bakker, A. (2024). Effects of a Learning Trajectory for statistical inference on 9th-grade students' statistical literacy. *Mathematics Education Research Journal*. https://doi.org/10.1007/s13394-024-00487-z

Verschaffel, L., Greer, B., & De Corte, E. (2000). Making sense of word problems. *Educational Studies in Mathematics, 42*(3), 211–231.

Watson, J., & English, L. (2022). Statistical literacy for students in the age of data science. *ZDM – Mathematics Education, 54*(1), 145–158. https://doi.org/10.1007/s11858-021-01284-x

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

Widjaja, W. (2021). Rethinking mathematical literacy in Indonesia: Implications for curriculum and assessment. *Indonesian Journal on Learning and Advanced Education, 3*(1), 15–27. https://doi.org/10.23917/ijolae.v3i1.11073

Wilmer, J., & Kerns, S. (2022). How bar graphs deceive: readout-based measurement reveals three fallacies. *Journal of Vision*, *22*(14), 3968. https://doi.org/10.1167/jov.22.14.3968

Xiang, Q., Bosch, R. J., & Lok, J. J. (2023). The survival-incorporated median vs the median inthesurvivors or in the always-survivors: What arewemeasuring? and Why? *Statistics in Medicine*. https://doi.org/10.1002/sim.9922

Zhang, Q., & Fiorella, L. (2022). An integrated model of learning from errors. *Educational Psychologist*, *58*(1), 18–34. https://doi.org/10.1080/00461520.2022.2149525

Zhang, X.-Q., Gao, Q.-H., Tian, W., & Xin, T. (2024). Dynamic and typological explanations of planning in complex problem-solving. *Learning and Individual Differences*. https://doi.org/10.1016/j.lindif.2024.102417

Тарасова, К., Gracheva, D., Талов, Д., Орел, Е., & Dementiev, A. (2025). Measuring changes in critical thinking skills among university economics students: insights from domain-specific assessment. *Studies in Higher Education*, 1–16. https://doi.org/10.1080/03075079.2025.2484423