Kognitif
Jurnal Riset HOTS Pendidikan Matematika
Volume- 5 No- 3 Halaman 1159 – 1170
ISSN 2776-9704 P-ISSN 2776-9984
https://doi.org/10.51574/kognitif.v5i3.1184

Practicality and Effectiveness of the COMESC Learning Model in Fostering Statistical Thinking Skills

Tan Hian Nio, Hardi Suyitno, Kartono, Scolastika Mariani

How to cite : Nio, T. H., Suyitno, H., Kartono, K., & Mariani, S. (2023). Practicality and Effectiveness of the COMESC Learning Model in Fostering Statistical Thinking Skills. *Kognitif: Jurnal Riset HOTS Pendidikan Matematika*, *5*(3), 1159–1170. https://doi.org/10.51574/kognitif.v5i3.1184

To link to this artcle : https://doi.org/10.51574/kognitif.v5i3.1184

Published Online on 13 August 2025

Kognitif: Jurnal Riset HOTS Pendidikan Matematika
https://doi.org/10.51574/kognitif.v5i3.1184 Volume 5, No 3, Juli - September 2025, pp. 1159– 1170

Practicality and Effectiveness of the COMESC Learning Model in Fostering Statistical Thinking Skills

Tan Hian Nio[1*], Hardi Suyitno[2], Kartono[3], Scolastika Mariani[4]

[1]Department of Mathematics Education, Faculty of Teacher Training and Education, Universitas Kristen Indonesia
[2,3,4]Department of Mathematics Education, Faculty of Mathematics and Natural Sciences, Universitas Negeri Semarang

Article Info

Article history:

Received Feb 05, 2025
Accepted Jul 26, 2025
Published Online Aug 13, 2025

Keywords:

COMESc Model
Practicality
Effectiveness
Statistical Thinking

ABSTRACT

The COMESc learning model is a specialized instructional approach developed through research and development to address university students' weaknesses in statistical thinking. These weaknesses include challenges in problem identification, hypothesis formulation, selection of appropriate statistical methods, interpretation of statistical relationships, application of significance testing, execution of calculations, comparison of sample and critical values, decision-making, and result interpretation. This study aimed to evaluate the practicality and effectiveness of the COMESc model in improving students' statistical thinking skills. A limited field trial was conducted with 27 participants over 10 class sessions. Qualitative data were collected through open-ended questionnaires and in-depth interviews, while quantitative data were obtained from closed-ended questionnaires and tests. Results indicated that the COMESc model is highly practical (95.65%) and effective, with 100% of participants achieving a minimum average test score of 73. The model also improved learning outcomes, enhanced performance, and stimulated active learning dynamics. Its strengths lie in ease of implementation, structured problem-solving syntax, and capacity to foster deeper statistical reasoning, while its main limitation is the need for instructor patience during initial implementation.

Corresponding Author:

Tan Hian Nio,
Department of Mathematics Education,
Faculty of Teacher Training and Education,
Universitas Kristen Indonesia
UKI Cawang Campus, Jl. Mayjen Sutoyo, Jakarta 13630, Indonesia
Email: hiannio.tan@gmail.com

Introduction

Statistics is a core subject in school curricula, spanning from elementary to higher education, and plays a crucial role in developing students' analytical and reasoning skills. Despite its importance, many students perceive statistics as complex and convoluted (Hollebrands & Lee, 2020; Parrill et al., 2018; van Dijke-Droogers et al., 2022), with most concepts being abstract (Kosko, 2020). This abstract nature often results in persistent difficulties in comprehension, leading to statistical thinking errors and conceptual misunderstandings, particularly in inferential statistics (Rellensmann et al., 2017; van Dijke-Droogers et al., 2022). If left unaddressed, these issues can hinder learning, limit motivation, and reduce students' engagement with statistical content.

The researcher's six years of experience teaching introductory inferential statistics reveal several recurring misconceptions. Students often struggle to measure data variability, formulate research questions, construct hypotheses, select appropriate statistical methods, link degrees of freedom with significance levels, determine critical values, compare sample statistics with critical values, apply decision-making criteria, ensure computational accuracy, critique analytical processes, evaluate results, and draw conclusions. Similar findings are reported by van Dijke-Droogers et al. (2022), who highlighted student errors in understanding correlation, confidence intervals, hypothesis testing, data interpretation, and regression analysis.

Previous studies have proposed various solutions, including integrating information technology into instruction and adopting constructivist-based learning designs (Günster & Weigand, 2020; Hollebrands & Lee, 2020; Rocha, 2020; Weldeana et al., 2023). While these approaches emphasize real data, technological tools, and classroom discussions, few have addressed the combined challenges of cognitive misconceptions and instructional limitations in teaching inferential statistics at the higher education level. Moreover, empirical studies on models that systematically integrate constructivist, meaningful, and scientific learning principles remain scarce. This gap underscores the need for an instructional design that not only improves statistical thinking but is also validated for practicality and effectiveness in authentic classroom contexts.

To address this need, the researcher developed the COMESc learning model—an acronym for *Constructivist, Meaningful, and Scientific* learning. This model is specifically designed to overcome students' academic weaknesses in statistical thinking and to guide instructors through a structured syntax for problem-solving, from initial planning to drawing conclusions (Habsyi et al., 2022; Rahayuningsih et al., 2022). It builds on the constructivist paradigm by encouraging learners to actively construct knowledge through observation, questioning, association, analysis, and conclusion-making, and on meaningful learning principles by promoting deep connections between concepts and real-world applications (Harris et al., 2023; Verzosa, 2020; Yao, 2020). This study addresses two research questions: (1) How practical is the COMESc learning model in fostering students' statistical thinking skills?; and (2) Is the COMESc learning model effective in meeting the minimum acceptable learning outcome for the "good" category, defined as a minimum score of 70?. Accordingly, the objectives are: (i) to describe the practicality and effectiveness of the COMESc model in teaching *Introduction to Inferential Statistics*, and (ii) to develop COMESc learning materials that meet the minimum "good" quality standard.

The practical and effective application of the COMESc model has the potential to advance both theory and practice in statistics education. Academically, it contributes to the literature by offering an empirically validated instructional model that addresses known misconceptions in inferential statistics. Practically, it provides instructors with a structured, evidence-based framework for improving students' statistical reasoning and learning outcomes. Ultimately, the COMESc model offers a promising alternative for higher education settings seeking to cultivate deeper statistical thinking and sustained high-quality learning performance.

Kognitif: Jurnal Riset HOTS Pendidikan Matematika
https://doi.org/10.51574/kognitif.v5i3.1184 Volume 5, No 3, Juli - September 2025, pp. 1159– 1170

Method

Settings

This study employed a research and development (R&D) approach. The article represents one part of the broader development of the COMESc learning model. The variables developed included learning tools such as the Teaching Guide, Semester Learning Plan (RPS), Student Worksheet (LKMD), and a statistical thinking ability test. As part of the model development process, the activities conducted in this study focused on testing the model's practicality and effectiveness to meet the criteria for being practical and effective. To describe the practicality of the learning model, ten class sessions were conducted, which included formative assessments and a final test. Formative assessments consisted of open-ended questionnaires, formative tests, and the final test. The purpose of the formative assessment was to evaluate the implementation and operability of the model, while the formative and final tests aimed to assess the progress and effectiveness of the model in the learning process. Assessments in R&D are used to evaluate a product through a series of field trials, with evaluations conducted at each stage (both on outcomes and processes) and model revisions made based on the trial findings. During the trials, both formative and final evaluations were conducted. The initial prototype of the model was assessed by experts and practitioners. Following several corrections and notes for improvement, the model was declared theoretically valid and ready for field testing. The first trial was conducted with a small group of 10 participants during the odd semester, followed by a limited trial with 27 mathematics education students during the even semester.

Data, Evaluation, and Analysis

To collect the required data, open-ended questionnaires, a statistical thinking test, and in-depth interviews were employed. The data consisted of qualitative and quantitative components used to test and evaluate both the model and its learning tools. The questionnaire was administered in two stages.

Stage 1: Practicality Assessment

The first-stage formative (partial) assessment by participants was conducted four times—during the 2nd, 4th, 6th, and 8th class meetings. All ten sessions, including tests, were completed. The aim was to gain an overview of the model's practicality based on participant evaluations. At this stage, the same questionnaire was administered, consisting of 12 statements regarding the COMESc model's learning syntax, with four response options: Strongly Agree (SA), Agree (A), Disagree (D), and Strongly Disagree (SD), scored 4, 3, 2, and 1 respectively. With 27 participants, the maximum possible score was (12 items × 4 × 27) = 1,296, and the minimum possible score was (12 items × 1 × 27) = 324. The formative assessment aimed to document and explain participants' responses during each session.

Practicality was determined by:

$$\text{Practicality (\%)} = \frac{\text{Score obtained per session}}{\text{Maximum Score}} \times 100\%$$

The score range between maximum and minimum values was divided using quartile intervals to categorize practicality as follows: minimum score to Q1 = not practical (low score), Q1–Q2 = fairly practical (medium score), Q2–Q3 = practical (high score), and above Q3 =

highly practical (very high score), as shown in Table 1. In this study, a model was considered practical if its practicality score was in the high or very high category.

Table 1. Formative Assessment Criteria

Score Interval	Category	Description
1050 – 1296	Very High	Model is highly practical
810 – 1052	High	Model is practical
567 – 809	Medium	Model is fairly practical
324 – 566	Low	Model is not practical

Stage 2: Practicality Assessment

The second-stage practicality assessment used a more comprehensive questionnaire covering all components of the model, administered during the 10th (final) session. The purpose was to evaluate overall practicality and to compare the results with those from the first-stage formative assessment. The questionnaire covered components such as syntax (five phases), social system, reaction principles, support systems, and both instructional and nurturant effects. The practicality indicators were aligned with these components and included 10 aspects, such as understanding the COMESc model, constructing knowledge in the context of statistical thinking, linking concepts in inferential statistics, applying STAD methods, applying scientific learning to foster statistical thinking, completing learning activities, fostering collaboration, providing recognition, informing about required learning facilities, and demonstrating both instructional and nurturant effects. From these indicators, 40 Likert-scale items were developed, each with four response options (SA=4, A=3, D=2, SD=1). With 27 participants, the maximum score was (40 × 4 × 27) = 4,320 and the minimum score was (40 × 1 × 27) = 1,080. Quartile intervals were used to categorize practicality: minimum–Q1 = not practical (low), Q1–Q2 = fairly practical (medium), Q2–Q3 = practical (high), and above Q3 = highly practical (very high). A model was deemed practical if it scored at least in the high category.

Table 2. Complete Practicality Assessment Criteria

Score Interval	Category	Description
3511 – 4320	Very High	Model is highly practical
2701 – 3510	High	Model is practical
1891 – 2700	Medium	Model is fairly practical
1080 – 1890	Low	Model is not practical

Model Effectiveness

Effectiveness was evaluated through formative and final test scores. The series of formative tests provided data on learning progress and academic improvement, indicating the strength of the COMESc model's influence. The normalised gain (*n-gain*) score was calculated using Hake's (1999) formula:

$$\langle \mathrm{n-gain} \rangle = \frac{\mathrm{post-test\ score - pre-test\ score}}{maximum\ ideal\ score - pre_{test}\ Score}$$

with benchmarks: 0.70 < *n-gain* ≤ 1.00 = high (very effective), 0.30 < *n-gain* ≤ 0.70 = medium (effective), and *n-gain* ≤ 0.30 = low (less effective). The minimum acceptable level was medium effectiveness. Effect size was also calculated to assess the practical or theoretical significance of the model's impact, following Cohen's (1988) and Becker's (2001) formula:

Kognitif: Jurnal Riset HOTS Pendidikan Matematika
https://doi.org/10.51574/kognitif.v5i3.1184 Volume 5, No 3, Juli - September 2025, pp. 1159– 1170

$$d = \frac{\bar{x}_{post\ test} - \bar{x}_{pre\ test}}{SD_{pooled}}$$

with interpretation (Lipsey & Wilson, 2001): 0.00–0.19 = trivial, 0.20–0.49 = small, 0.50–0.79 = medium, and ≥ 0.80 = large effect.

Interviews

Qualitative data were obtained through triangulation of techniques (different sources and methods) using open-ended questionnaires followed by interviews. Five participants, selected randomly, were interviewed with probing questions to elicit in-depth responses. All interviews were recorded using laptops and mobile phones, transcribed verbatim, and classified for analysis. The aim was to verify and complement responses from the open-ended questionnaires. Responses from interviews and questionnaires were cross-checked, reduced by grouping similar information into subcategories, and described. Data saturation was used as the stopping criterion, when responses were consistent and no new information emerged. A total of 15 interview questions addressed various aspects of the model. In this article, the researcher presents one example item: *"Are you interested and motivated to learn using the COMESc learning approach as presented in the LKMD? Please explain your reasons."*

Results

Practicality Assessment

First-Stage Formative Assessment

The model development results were obtained through a limited trial involving 27 mathematics education students. The instructional trial was conducted over 10 class meetings using the COMESc learning model, with four formative assessments administered via questionnaires. The results of the formative practicality questionnaire are presented in Table 3. The percentage of practicality ranged from 96.5% to 97.7%, with all four formative assessments classified as very high in practicality. The average practicality percentage was 97.1%.

These results indicate that, in all four formative assessments, the COMESc learning model was perceived by participants as highly practical. The lowest practicality score occurred in the first assessment, suggesting some initial weaknesses in understanding certain learning elements during the early sessions. This was evident from spontaneous questions raised by participants regarding prerequisite knowledge for the topic of homogeneity of variance and how to construct new knowledge by linking it to prior knowledge. The trend showed an increase from the first to the fourth assessment, reaching a maximum practicality of 97.7%. While the practicality level remained very high, these results imply that minor weaknesses in the model's concept, implementation, or operability still existed. Nevertheless, overall, the analysis of the first-stage formative assessments confirmed that the COMESc model is practical and feasible for use in the *Introduction to Inferential Statistics* course, based on participants' learning experiences.

Table 3. Formative Assessment Scores for Class Meetings 1–4

No.	Statement	Score per Meeting
1	The lecturer's explanation of the application of the COMESc model's instructional syntax motivates me to learn and understand Homogeneity of Variance.	103
2	The lecturer's explanation of the learning outcomes for today's topic helps me stay focused on learning.	98
3	I can explain the components of the COMESc model syntax application as a guide to understanding today's material.	100
4	I am able to recall prerequisite knowledge to understand the topic of Homogeneity of Variance (F-Test) after receiving the lecturer's explanation.	102
5	I am able to construct new knowledge by integrating my prior knowledge with today's topic as required in the COMESc model.	104
6	I am better able to link one concept to another after participating in the COMESc learning activities provided by the lecturer.	107
7	I am better able to formulate problems and hypotheses in statistical tasks after practicing with the COMESc learning model.	108
8	I can select the appropriate statistical method and perform calculations in statistical tasks after practicing with the COMESc model.	108
9	Every discussion group member in the LKMD actively participates and spontaneously shares opinions on the material provided by the lecturer.	104
10	The COMESc learning model helps me better understand statistical material by following the guidance provided in the LKMD.	109
11	The COMESc learning model encourages me to prepare thoroughly for the presentation I am working on.	104
12	Learning with the COMESc model captures my attention and clearly guides me in solving statistical problems.	103

Second-Stage Final Assessment

In addition to the first-stage formative assessment, a comprehensive (final) practicality assessment was conducted during the 10th and final meeting, covering all components and processes of the COMESc model. The purpose was to evaluate the overall practicality of the model and to compare the results with those from the first-stage formative assessments. Based on the assessment criteria, this falls into the very high category. These results are consistent with the first-stage formative assessment findings, confirming that the COMESc learning model is highly practical and feasible for use in the Introduction to Inferential Statistics course.

Table 4. Final Practicality Assessment Scores for the COMESc Learning Model

No	Objective	Statement	Score
1	Understanding the concept of the COMESc learning model	**A. Opening Phase**	
		1. The lecturer greets all students to begin the class.	109
		2. The lecturer explains the activities in the COMESc learning model.	105
		3. The lecturer motivates students to learn inferential statistics.	98
2	Ability to connect concepts in inferential statistics and their applications	**B. Implementation Phase of the COMESc Model**	
		4. The lecturer explains the learning objectives for the topic at the beginning of the session.	104
		5. The lecturer and students engage in Q&A about prerequisite knowledge needed for the new topic of simple correlation.	101

Kognitif: Jurnal Riset HOTS Pendidikan Matematika
https://doi.org/10.51574/kognitif.v5i3.1184 Volume 5, No 3, Juli - September 2025, pp. 1159– 1170

		6. I can construct and explain simple correlation in my own words.	99
		7. I can formulate questions and hypotheses for problems related to simple correlation.	104
		8. I can outline the steps for solving correlation and simple linear regression problems according to the COMESc model.	100
3	Ability to connect concepts in inferential statistics and their applications	**C. Linking Concepts Phase**	
		9. I can identify the link between prerequisite material and the new concepts introduced.	102
		10. I can state the relationship between one concept and another in correlation topics.	101
		11. I can provide examples of correlation relationships in daily life.	101
		12. I can link critical value tables with sample statistics for the topic taught.	102
4	Applying the STAD method in learning	**D. Group Discussion Phase**	
		13. I actively participate in small group (3–4 people) discussions on the topic provided.	102
		14. I demonstrate teamwork in solving problems during group discussions.	104
		15. I enjoy using the LKMD as it helps guide me in solving inferential statistics problems.	102
5	Applying scientific learning to foster statistical thinking	**E. Scientific Learning Phase**	
		16. I actively review and formulate questions about the given problems.	104
		17. I can choose the correct statistical method to test hypotheses for the given problems.	105
		18. I can perform calculations to determine the sample statistics value.	104
		19. I critically assess the problem-solving process completed by a group or individual.	104
		20. I can evaluate the hypothesis testing process for the given problems.	103
		21. I can determine the statistical test criteria for accepting/rejecting a hypothesis and draw conclusions.	105
		22. I present my group's work results earnestly.	104
		23. My motivation to learn statistics increases after the COMESc model is applied.	105
6	Concluding the learning activities	**F. Closing Phase**	
		24. At the end of the lesson, the lecturer guides students to make conclusions based on the statistical thinking process.	103
		25. The lecturer provides reinforcement and reflection to students individually or in groups.	101

		26. I support the implementation of the COMESc model to be disseminated across the university.	104
7	Demonstrating lecturer–student collaboration in learning	**G. Social Principle Implementation**	
		27. The lecturer and students agree to attend class punctually on a voluntary basis.	103
		28. Students complete group assignments conscientiously.	104
		29. The lecturer and students discuss problem-solving together.	105
8	Providing recognition in the learning process	**H. Reaction Principle Implementation**	
		30. The lecturer gives recognition to students who are critical during the lesson.	102
		31. The lecturer gives recognition to students who complete assignments conscientiously.	102
		32. The lecturer gives grades and feedback on student worksheets.	102
9	Providing information on learning facilities needed to support learning outcomes	**I. Supporting System Availability Implementation**	
		33. The lecturer provides the LKMD for group/individual work.	104
		34. The lecturer assists students in accessing references via the web.	116
		35. The lecturer provides and distributes the PPT for study purposes.	104
10	Demonstrating instructional and nurturant effects in learning	**J. Achievement of Instructional and Nurturant Effects Implementation**	
		36. I aim to maintain at least 90% attendance across all face-to-face sessions.	105
		37. I am enthusiastic in completing the problems given by the lecturer.	102
		38. I aim to achieve at least 85 for each assignment.	99
		39. I take responsibility for completing the work assigned by the lecturer.	103
		40. The COMESc learning model can foster students' knowledge and statistical thinking skills.	105

Effectiveness Assessment Results

The effectiveness of the COMESc model was measured using data from formative and final tests, consisting of pre-test scores, Test 1, Test 2, Test 3, and the final post-test. Learning progress and improvement in statistical thinking were analyzed using the normalized gain (*n-gain*), while the magnitude of the model's effect was determined using Cohen's *d*.

Table 5. Formative and Final Test Results

Test	Pre-test	Tf 1	Tf 2	Tf 3	Post-test
Mean score	67.4	79.4	85.6	91.3	82.6
Standard deviation	6.846				5.483
n-gain	0.47 (medium)				
Cohen's *d*	3.45 (strong)				

Kognitif: Jurnal Riset HOTS Pendidikan Matematika
https://doi.org/10.51574/kognitif.v5i3.1184 Volume 5, No 3, Juli - September 2025, pp. 1159– 1170

The results show an improvement in academic performance from the pre-test through Test 1, Test 2, and Test 3, with a slight decrease in the post-test. The average score across Tests 1, 2, 3, and the post-test was 84.735, indicating an overall increase from pre-test to post-test. This value falls into the strong effect category. Thus, the COMESc learning model had a strong influence on improving students' statistical thinking skills, confirming its effectiveness for use in teaching Introduction to Inferential Statistics.

Participant Interviews

All participants stated that they were interested and motivated to learn using the COMESc learning approach as presented in the LKMD, citing various reasons. They explained that the approach made it easier to solve problems collaboratively in groups and to answer questions more systematically. It also helped them place answers in the correct sections, better understand the questions, and complete tasks in an orderly manner. The structured nature of the approach, supported by the LKMD, enabled them to organize their work more effectively, review lessons, and produce clearer, more focused answers point by point. Participants noted that their written work became neater and more coherent, avoiding confusion about which tasks to prioritize. They also appreciated the novelty of the LKMD method, which they had not previously experienced, and valued the opportunity to engage in peer discussions. Overall, they found that the COMESc model facilitated a more organized, structured, and efficient problem-solving process.

Discussion

The implementation of the COMESc model in the introductory inferential statistics course proved beneficial for participants by guiding them through the model's syntax as a foundation for solving statistical problems. As Callingham & Siemon (2021), Nolte & Pamperien (2017), and Soneira et al. (2018), notes, students often require assistance in problem-solving to avoid relying on irrelevant trial-and-error methods they devise themselves. This model offers structured technical guidance that directs participants to follow a series of syntax steps aimed at building a conceptual understanding of statistical thinking. Rather than repeatedly applying familiar formulas (an approach that, as Tanudjaya & Doorman (2020) observed, often fails to guide learners toward sound conclusions) participants were encouraged to follow an ordered problem-solving pathway. This shift was particularly relevant for those with weak statistical backgrounds, language barriers, or conceptual difficulties.

The COMESc approach recommends that students reconsider their prior learning and problem-solving methods in favor of a sequential, step-by-step model that leads from basic tasks to final conclusions, as outlined in the student group worksheets (LKMD). The practicality trials demonstrated consistently high ratings from participants, indicating that the model can be effectively used in teaching introductory inferential statistics. Participant interviews further reinforced this finding: students reported that the COMESc model helped them approach problems more systematically, organize answers point-by-point, and avoid confusion regarding task order. The unique use of LKMD was seen as engaging, promoting collaborative discussion, and enabling efficient and well-structured problem-solving.

These perceptions were corroborated by measurable academic gains, as evidenced by formative and final test results. All participants achieved an average score of at least 73 across all assessments, indicating effective learning outcomes attributable to the COMESc model. The structured problem-solving steps in the LKMD appear to have played a crucial role in enhancing statistical reasoning. This aligns with Öçal et al. (2020) assertion that structured problem-solving in statistics benefits both instructors and students, and supports a re-examination of current assessment and teaching approaches in statistics education.

Conclusion

This study demonstrates that the COMESc model offers substantial benefits for teaching introductory inferential statistics, both in practice and in theory. Beyond its practical utility, the model contributes to the theoretical discourse on structured learning by embodying key principles of constructivism, cognitive load theory, and scaffolding. Its sequenced tasks facilitate active knowledge construction, reduce extraneous cognitive load through clear procedural guidance, and provide graduated support that enables learners to progress from basic concepts to more complex reasoning in a controlled manner. Despite these promising results, the study is limited to a single subject area within one institutional context, with participants sharing similar academic backgrounds. This may constrain the generalizability of the findings. Furthermore, the reliance on self-reported data in the practicality evaluation could introduce subjective bias. Future research should extend the application of the COMESc model across multiple institutions, disciplines, and learning environments to validate its broader applicability. The integration of digital tools, such as learning management systems and interactive statistical software, holds potential to further enhance its efficiency. Longitudinal studies are also recommended to assess whether the structured problem-solving skills developed through COMESc can be transferred to more advanced statistical topics or other fields that demand systematic reasoning. In sum, the COMESc model emerges as an effective and efficient pedagogical strategy that strengthens statistical reasoning through structured syntax, scaffolding, and collaborative problem-solving, while also laying a robust foundation for future innovation in instructional design.

Conflict of Interest

The authors declare that there is no conflict of interest.

Author Contributions

T.H.N. conceptualized the research idea presented and collected the data. The other fourth authors (H.S., K, and S.M.) actively contributed to the development of the theory, methodology, data organization and analysis, discussion of results, and approval of the final version of the work. All authors confirm that they have read and approved the final version of this manuscript. The percentage contributions to the conceptualization, drafting, and revision of this paper are as follows: T.H.N.: 70%, H.S.: 10%, K.: 10%, and S.M.: 10%.

Data Availability Statement

The authors state that the data supporting the findings of this study are available from the corresponding author, [T.H.N], upon reasonable request.

References

Callingham, R., & Siemon, D. (2021). Connecting multiplicative thinking and mathematical reasoning in the middle years. *Journal of Mathematical Behavior*, *61*(December 2020), 100837. https://doi.org/10.1016/j.jmathb.2020.100837

Günster, S. M., & Weigand, H. G. (2020). Designing digital technology tasks for the development of functional thinking. *ZDM - Mathematics Education*, *52*(7), 1259–1274. https://doi.org/10.1007/s11858-020-01179-1

Habsyi, R., R. M. Saleh, R., & Isman M. Nur. (2022). Pengembangan E-LKPD Berbasis Guided Dicovery Learning untuk Meningkatkan Kemampuan Berpikir Kritis Siswa. *Kognitif: Jurnal Riset HOTS Pendidikan Matematika*, *2*(1), 1–18. https://doi.org/10.51574/kognitif.v2i1.385

Harris, D., Logan, T., & Lowrie, T. (2023). Spatial visualization and measurement of area: A case study in spatialized mathematics instruction. *Journal of Mathematical Behavior*, *70*(May 2022), 101038. https://doi.org/10.1016/j.jmathb.2023.101038

Hollebrands, K. F., & Lee, H. S. (2020). Effective design of massive open online courses for mathematics teachers to support their professional learning. *ZDM - Mathematics Education*, *52*(5), 859–875. https://doi.org/10.1007/s11858-020-01142-0

Kosko, K. W. (2020). The multiplicative meaning conveyed by visual representations. *Journal of Mathematical Behavior*, *60*(January), 100800. https://doi.org/10.1016/j.jmathb.2020.100800

Nolte, M., & Pamperien, K. (2017). Challenging problems in a regular classroom setting and in a special foster programme. *ZDM - Mathematics Education*, *49*(1), 121–136. https://doi.org/10.1007/s11858-016-0825-5

Öçal, M. F., Şen, C., Güler, G., & Kar, T. (2020). The investigation of prospective mathematics teachers' non-algebraic solution strategies for word problems. *International Journal of Mathematical Education in Science and Technology*, *51*(4), 563–584. https://doi.org/10.1080/0020739X.2019.1597936

Parrill, F., McKim, A., & Grogan, K. (2018). Gesturing standard deviation: Gestures undergraduate students use in describing their concepts of standard deviation. *Journal of Mathematical Behavior*, *February 2017*, 0–1. https://doi.org/10.1016/j.jmathb.2018.05.003

Rahayuningsih, S., Nurasarawati, N., & Nurhusain, M. (2022). Komparasi Efektivitas Model Pembelajaran Project Based Learning (PjBL) dan Konvensional: Studi Pada Siswa Menengah Pertama. *Kognitif: Jurnal Riset HOTS Pendidikan Matematika*, *2*(2), 118–129.

Rellensmann, J., Schukajlow, S., & Leopold, C. (2017). Make a drawing. Effects of strategic knowledge, drawing accuracy, and type of drawing on students' mathematical modelling performance. *Educational Studies in Mathematics*, *95*(1), 53–78. https://doi.org/10.1007/s10649-016-9736-1

Rocha, H. (2020). Using tasks to develop pre-service teachers' knowledge for teaching mathematics with digital technology. *ZDM - Mathematics Education*, *52*(7), 1381–1396. https://doi.org/10.1007/s11858-020-01195-1

Soneira, C., González-Calero, J. A., & Arnau, D. (2018). An assessment of the sources of the reversal error through classic and new variables. *Educational Studies in Mathematics*, *99*(1), 43–56. https://doi.org/10.1007/s10649-018-9828-1

Tanudjaya, C. P., & Doorman, M. (2020). Examining higher order thinking in Indonesian lower secondary mathematics classrooms. *Journal on Mathematics Education*, *11*(2), 277–300. https://doi.org/10.22342/jme.11.2.11000.277-300

van Dijke-Droogers, M., Drijvers, P., & Bakker, A. (2022). Introducing Statistical Inference: Design of a Theoretically and Empirically Based Learning Trajectory. *International Journal of Science and Mathematics Education*, *20*(8), 1743–1766. https://doi.org/10.1007/s10763-021-10208-8

Verzosa, D. M. B. (2020). Pre-service elementary teachers' reasoning in introducing multidigit subtraction. *International Journal of Mathematical Education in Science and Technology*, *51*(6), 876–892. https://doi.org/10.1080/0020739X.2019.1647573

Weldeana, H. N., Sbhatu, D. B., & Berhe, G. T. (2023). Freshman STEM students' misconceptions in a basic limit theorem of limx→0(sinx/x). *Heliyon*, *9*(12), 1–16. https://doi.org/10.1016/j.heliyon.2023.e22359

Yao, X. (2020). Unpacking learner's growth in geometric understanding when solving problems in a dynamic geometry environment: Coordinating two frames. *Journal of Mathematical Behavior*, *60*(April), 100803. https://doi.org/10.1016/j.jmathb.2020.100803

Author Biographies

	**Tan Hian Nio.** is a permanent lecturer in the Mathematics Education Study Program, Faculty of Teacher Training and Education, Universitas Kristen Indonesia, Jakarta, Indonesia. Her area of expertise is statistics, with research interests in statistical education, applied statistics, and quantitative research methods. Dr. Tan has been actively involved in teaching, curriculum development, and publishing research in national and international journals.E-mail*:* hiannio.tan@gmail.com
	**Hardi Suyitno.** is a Professor in the Mathematics Education Study Program, Faculty of Teacher Training and Education, Universitas Ahmad Dahlan, Indonesia. He earned his Bachelor's degree in Pure Mathematics from IKIP Yogyakarta in 1978, his Master's degree in Mathematics Education from IKIP Malang in 1988, and his Doctoral degree in Philosophy from Universitas Gadjah Mada in 2008. His primary research interests lie in the philosophy of mathematics, with a particular focus on the epistemological and ontological foundations of mathematical thought and their implications for mathematics education. Suyitno has extensive teaching experience in courses such as Linear Programming, History and Philosophy of Mathematics, and Analysis of Secondary School Mathematics Curriculum. He has published widely in national and international journals, presented in various academic forums, and actively contributes to the advancement of mathematics education in Indonesia. E-mail*:* hardisuyitno@pmat.uad.ac.id
	**Kartono** is a Professor in the Master's Program of Educational Research and Evaluation, Graduate School, Universitas Negeri Semarang, Indonesia. He holds a Master's degree in Statistics and is currently pursuing a Doctoral degree in Educational Research and Evaluation. His research interests include educational assessment, statistical modeling, program evaluation, and measurement in education. Kartono has extensive experience in supervising graduate students, publishing in national and international journals, and serving as a reviewer for various academic publications. Email: kartono@unnes.ac.id
	**Scolastika Mariani, M.Si.** is an Associate Professor in the Undergraduate Program of Statistics and Data Science, Faculty of Mathematics and Natural Sciences, Universitas Negeri Semarang, Indonesia. She holds a Master's degree in Statistics and a Doctoral degree in Statistics. Her research interests include applied statistics, data science, statistical modeling, and educational measurement. Dr. Mariani has published in national and international journals, supervised undergraduate and graduate theses, and actively contributes as a reviewer for various academic publications. Email: scolastika.mariani@unnes.ac.id