MJP Journal of Education and Teaching Learning
Vol.4, No.2 (2026): 730 – 734
https://jurnalilmiah.co.id/index.php/MJPJETL



Development of Mathematics Problems to Measure Critical Thinking Skills in Fifth-Grade Elementary School Students

[1)]**Septhia Indah Putri,** [2)]**Nila Kesumawati,** [3)]**Destiniar**

[1,2,3)] Primary School Teacher Education, PGRI Palembang University

***Correspondence Author:** septiaindahputri004@gmail.com

Article Info

Keywords:
Mathematical Question Development, Critical Thinking Skills, Fifth-Grade Elementary School Students, Tessmer Model.

ABSTRACT

This study aims to produce mathematics questions that can be used to measure the critical thinking skills of fifth-grade elementary school students on the material of data presentation and interpretation and meet the criteria of validity and reliability. The method used is Research and Development (R&D) with the Tessmer development model consisting of preliminary stages, self-evaluation, expert review, one-on-one, small group, and field testing. This study was conducted involving 3 students in the one-on-one stage, 6 students in the small group stage, and 30 students in the field testing stage. Data collection was carried out through expert validation sheets, interviews, and tests, then analyzed using qualitative and quantitative approaches. The results of the study showed that the assessment of the three validators obtained an average percentage of 81.3% which was included in the valid category. In the one-on-one and small group stages, the questions developed were considered easy to understand by students and had met the aspects of readability and clarity so that they were worthy of being tested in the field testing stage. The results of the validity test showed that 10 questions (83.33%) were declared valid and 2 questions (16.67%) were declared invalid. Meanwhile, the reliability test results obtained a coefficient of 0.967, which is categorized as reliable. Therefore, it can be concluded that the developed mathematics questions have met the validity and reliability criteria, making them suitable for measuring the critical thinking skills of fifth-grade elementary school students.

Informasi Artikel

Kata Kunci:
Pengembangan Soal Matematika, Keterampilan Berpikir Kritis, Siswa Kelas Lima Sekolah Dasar, Model Tessmer

ABSTRAK

Penelitian ini bertujuan untuk menghasilkan soal matematika yang dapat digunakan untuk mengukur kemampuan berpikir kritis siswa kelas lima SD pada materi penyajian dan interpretasi data serta memenuhi kriteria validitas dan reliabilitas. Metode yang digunakan adalah Penelitian dan Pengembangan (R&D) dengan model pengembangan Tessmer yang terdiri dari tahap pendahuluan, evaluasi diri, tinjauan ahli, satu lawan satu, kelompok kecil, dan uji lapangan. Penelitian ini dilakukan dengan melibatkan 3 siswa pada tahap satu lawan satu, 6 siswa pada tahap kelompok kecil, dan 30 siswa pada tahap uji lapangan. Pengumpulan data dilakukan melalui lembar validasi ahli, wawancara, dan tes, kemudian dianalisis menggunakan pendekatan kualitatif dan kuantitatif. Hasil penelitian menunjukkan bahwa penilaian dari tiga validator memperoleh persentase rata-rata 81,3% yang termasuk dalam kategori valid. Pada tahap satu lawan satu dan kelompok kecil, soal yang dikembangkan dianggap mudah dipahami oleh siswa dan telah memenuhi aspek keterbacaan dan kejelasan sehingga layak diuji pada tahap uji lapangan. Hasil uji validitas menunjukkan bahwa 10 pertanyaan (83,33%) dinyatakan valid dan 2 pertanyaan (16,67%) dinyatakan tidak valid. Sementara itu, hasil uji reliabilitas memperoleh koefisien 0,967, yang dikategorikan reliabel. Oleh karena itu, dapat disimpulkan bahwa soal-soal matematika yang dikembangkan telah memenuhi kriteria validitas dan reliabilitas, sehingga cocok untuk mengukur kemampuan berpikir kritis siswa kelas lima sekolah dasar.

Article History

Received : 15/08/2026
Revised : 10/09/2026
Accepted : 21/10/2026

✉ **Corresponding Author:** (1) Septhia Indah Putri, (2) Department of Pendidikan Guru Sekolah Dasar, (3) Institution of Universitas PGRI Palembang, (4) Palembang, Indonesia, (5) Email: septiaindahputri004@gmail.com

Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.
Copyright © 2026 - [1)] Septhia Indah Putri, [2)] Nila Kesumawati, [3)] Destiniar.

1. Introduction

Education is a planned and continuous process aimed at comprehensively developing learners' potential across intellectual, emotional, social, and moral dimensions. Rahman et al. (2022) explain that education is a conscious and systematic effort to create a learning environment and process that enables learners to actively develop their inherent potential. Within the context of the *Kurikulum Merdeka* (Independent Curriculum), a key competency being fostered is Higher-Order Thinking Skills (HOTS), specifically critical thinking.

Mathematics, as a subject playing a strategic role in the development of logic and reasoning, should serve as a primary vehicle for cultivating students' critical thinking skills. As explained by Febriano et al. (2021), mathematics instruction should not merely focus on the mastery of concepts and formulas but should also train students to analyze, evaluate, and draw logical conclusions from problems. Ideally, teachers implement assignments and assessments that encourage higher-order thinking through contextual and open-ended problems.

Critical thinking ability in mathematics learning is reflected in students' capacity to understand problems, analyze provided data or information, select appropriate solution strategies, and evaluate the results obtained. Iskandar (2020) states that mathematical critical thinking requires students not only to arrive at a final answer but also to understand and explain the solution process. In line with this, Sapitri et al. (2025) argue that the assessment of critical thinking skills in mathematics must encompass reasoning, decision-making, and reflection on the problem-solving process to ensure a comprehensive measurement of students' critical thinking abilities.

Based on observations conducted on October 31, 2025, a needs analysis involving fifth-grade students and their teacher at SDN 74 Palembang revealed that most students' learning outcomes remained low and that they struggled to answer questions requiring critical thinking. Consequently, the learning and assessment processes in primary schools must focus not only on final outcomes but also on the students' thinking processes.

The urgency of this research lies in the need for an assessment instrument that comprehensively measures both mathematical knowledge and critical thinking skills among primary school students. This aligns with the *Profil Pelajar Pancasila* (Pancasila Student Profile) guidelines within the *Kurikulum Merdeka* (Independent Curriculum), which call for learners who are reflective and capable of critical reasoning. Research by Hastuti et al. (2024) underscores the importance of assessments that encourage students to formulate generalizations, draw conclusions, and evaluate mathematical solutions. By developing assessment instruments based on real-world contexts, teachers can evaluate students' thinking processes more objectively and meaningfully.

Therefore, to address this need, this research focuses on developing a mathematics assessment instrument—specifically an open-ended test—designed to measure students' critical thinking skills, particularly through essay-style questions that allow students to demonstrate their thought processes.

More recent research conducted by Trisnani et al. (2024) developed an assessment instrument for critical thinking skills in the form of open-ended questions for elementary school students. The results indicated that open-ended questions were more effective than objective questions at revealing students' thought processes, as they required students to logically explain their reasoning and steps toward a solution. This aligns with findings by Hollingshead (2024), who emphasized that open-ended instruments possess a higher level of sensitivity in measuring critical thinking skills.

The instrument developed in this study was designed to measure four indicators of critical thinking skills: clarification, interpretation, inference, and evaluation. Clarification relates to the ability to understand and identify information within a problem. Interpretation refers to the ability to interpret presented data or information. Inference is the ability to draw conclusions based on available information, while evaluation concerns the ability to assess the correctness or accuracy of an answer. These indicators are based on the critical thinking framework proposed by Facione and adapted by Agustine et al. (2020) (citing Surya & Zuhur, 2017). These four indicators were selected because they align with the cognitive development characteristics of fifth-grade students and can be measured through open-ended mathematics questions.

Furthermore, the topic of data presentation and interpretation in fifth-grade mathematics plays a strategic role in training students to understand and process quantitative information related to daily life. This material emphasizes not only the ability to read and present data in various forms but also encourages students to analyze, compare, and draw logical conclusions. Through activities involving data interpretation and data-based decision-making, instruction on this topic holds significant potential for developing students' critical thinking and reasoning abilities. Therefore, the topic of data presentation and interpretation serves as an effective learning context for fostering critical thinking skills among elementary school students. Based on this background, this development research aims to produce a set of mathematical problems for the fifth-grade mathematics curriculum. The study is expected to yield a valid and reliable assessment instrument while making a tangible contribution to supporting the implementation of the *Kurikulum Merdeka* (Independent Curriculum) and enhancing the quality of mathematics instruction in elementary schools.

2. Method

This study employs a Research and Development (R&D) approach aimed at producing an educational product specifically, a set of mathematical test items and systematically evaluating its quality. The development model utilized is Tessmer's formative evaluation model. This model was selected because it emphasizes a phased formative evaluation process to refine the product through continuous feedback. The study population consists of 124 fifth-grade students at SD Negeri 72 Palembang during the 2025/2026 academic year.

Data collection techniques were aligned with the characteristics of R&D, which require data on both the final outcome and the development process to assess product quality; these techniques included interviews, validation questionnaires, and tests. Prototype validation was conducted through expert judgment to assess the instrument's suitability prior to student trials. Data analysis involved both quantitative and qualitative descriptive approaches.

3. Results and Discussion

3.1 Result

The research was conducted over two sessions, with each session lasting two hours. The study aimed to determine the validity and reliability of the mathematical problem set developed to measure the critical thinking skills of fifth-grade elementary school students. The subjects consisted of two fifth-grade classes: Class V.C (9 students) served as the pilot group, while Class V.D (30 students) served as the experimental group. The development model employed was the Tessmer model, which comprises several stages: preliminary, self-evaluation, prototyping (expert review, one-to-one, and small group), and field test.

Preliminary Stage

The results of the preliminary stage served as the basis for designing and developing the mathematical critical thinking assessment instrument, ensuring alignment with student characteristics and learning objectives. Regarding the research schedule, arrangements were made through prior communication and mutual agreement with the school.

Self-Evaluation

At this stage, the researcher conducts an assessment by carrying out the following steps:

a. Curriculum Analysis

This research is grounded in the *Kurikulum Merdeka* (Independent Curriculum), which emphasizes strengthening students' critical thinking skills. However, the assessment methods currently used in schools tend to be procedural and are not yet optimal for measuring critical thinking skills. Therefore, it is necessary to develop open-ended mathematics problems that align with the requirements of the *Kurikulum Merdeka* (Ministry of Education, Culture, Research, and Technology, 2023).

b. Needs Analysis

Based on interviews with a Grade 5 elementary school teacher, it was revealed that the learning process—particularly in mathematics—relies heavily on teacher and student guidebooks, focusing primarily on routine problems. Furthermore, teachers lack instruments to assess critical thinking; this results in students failing to think critically and instead resorting to guessing when solving problems.

c. Student Analysis

The second analysis focuses on the students. This analysis aims to understand student characteristics—such as ability levels, prior knowledge, cognitive development rates, and readiness—in the context of developing mathematics problems. Interview results indicate that students possess varying abilities in absorbing learning material; the use of engaging problems can alleviate boredom and enhance student attention during lessons.

d. Content Analysis

Content analysis is conducted to identify and organize the material to be taught to students, based on the initial analysis. The topic selected for developing mathematics problems is "data," covering skills such as reading data, interpreting data, and drawing conclusions. The data topic supports authentic assessment of critical thinking because it involves processing complex information—akin to real-life scenarios like analyzing surveys or sales graphs—which elementary students rarely master, while also aligning with established learning outcomes and objectives.

Expert Riview

This stage involves the validation process for the developed assessment instrument. The instrument, designed to measure critical thinking skills, was first validated by two mathematics lecturers and one elementary school teacher.

Table 1. Overall Validation Results

No	Validator	Aspek				
		Materi	Kontruksi	Bahasa	Rata-Rata	Kategori
1.	Validator 1	85,7%	81,25%	95%	87,3%	Valid
2.	Validator 2	75%	75%	75%	75%	Valid
3.	Validator 3	78,57%	81,25%	85%	81,60%	Valid
		79,75%	79,16%	85%	81,3%	Valid

Based on the validation conducted by expert validators, the mathematical critical thinking assessment instrument was rated as valid. The results indicate that the developed questions met suitability criteria regarding content, construction, and language.

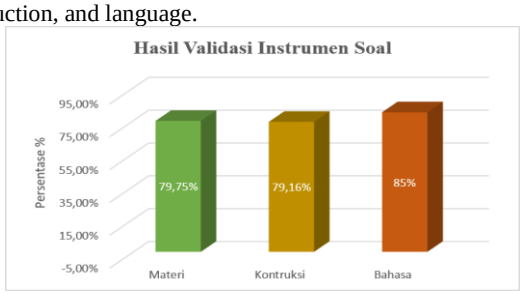


Figure 1. Validation Results from Three Validators

Once the expert validation process was complete and the instrument was deemed valid and suitable for use, the next steps involved "one-to-one" and "small-group" trials.

Research Stages

a. One-to-One

The one-to-one stage was conducted after the instrument had been validated by experts. During this stage, the researcher administered a trial to three fifth-grade elementary school students representing different ability levels: high, medium, and low. This activity aimed to assess the readability of the questions, the clarity of the language, and any difficulties students encountered in understanding the instrument. In this stage, students were asked to complete 12 open-ended questions.



Figure 2. One-to-One Trial

Based on the results of the one-to-one stage, students provided feedback regarding the clarity of certain parts of the questions. This feedback was taken into account to refine the instrument, thereby improving the quality of the questions before the small-group trial.Berdasarkan hasil tahap *one to one*, terdapat beberapa masukan dari siswa terkait kejelasan beberapa bagian soal. Masukan tersebut dijadikan bahan pertimbangan dalam penyempurnaan instrumen soal sehingga kualitas soal menjadi lebih baik sebelum diuji cobakan pada tahap small group.

b. Small Group

The small-group stage is conducted after the instrument has been revised based on the results of the one-to-one trial. In this stage, the trial involves six fifth-grade students, selected based on their top six mathematics scores achieved during regular classroom instruction. Students are asked to complete 12 open-ended questions assessing mathematical critical thinking skills—items previously deemed suitable—and their responses are analyzed to evaluate the quality of each question. The objective of this stage is to assess the instrument's practicality and gauge student response regarding whether the questions are understandable and manageable, including factors such as completion time efficiency and the alignment of difficulty levels with student abilities. This stage aims to identify potential issues before the instrument undergoes large-scale testing.



Figure 3. Small-Group Trial

The small-group trial results indicate that all questions were well understood by the six participating students. No issues arose during the activity suggesting that the questions were difficult to comprehend or caused student misconceptions. Furthermore, the students were able to follow the instructions and successfully complete the questions. These results demonstrate that the instrument, having been refined in the previous stage, meets the criteria for readability and clarity, making it suitable for use in the field test stage.

c. Field Test

The field test stage is conducted after the assessment instrument has undergone revision or refinement during the "one-to-one" and "small group" phases. This stage aims to evaluate the quality of the developed mathematical critical thinking instrument before it is utilized as a research measurement tool.

The field test was administered to 30 fifth-grade elementary school students. During this stage, students were asked to complete 12 open-ended questions developed based on critical thinking ability indicators. While the students worked on the questions, the researcher observed their ability to comprehend the items and noted any difficulties they encountered. Once all student responses were collected, the researcher analyzed the results to determine the suitability of each developed test item. Tahap *field test* merupakan tahap uji coba lapangan yang dilakukan setelah instrumen soal melalui tahap perbaikan atau penyempurnaan pada *one to one* dan *small group.* Tahap ini bertujuan untuk mengetahui kualitas instrumen soal kemampuan berpikir kritis matematis yang dikembangkan sebelum digunakan sebagai alat ukur penelitian.

Field Test Item Validity Results

The validity results for the test items—based on a field test involving 30 students from Grade V at SDN 74 Palembang—indicate that out of the 12 items tested, 10 met the validity criteria, while 2 did not. Validity was determined by comparing the calculated correlation coefficient (r_{count}) with the critical table value (r_{table}) of 0.361. An item was deemed valid if $r_{count} > r_{table}$ and invalid if $r_{count} < r_{table}$. Based on these results, items 1, 2, 3, 4, 5, 6, 7, 9, 10, and 11 were declared valid, whereas items 8 and 12 were declared invalid. Valid items demonstrate the ability to measure the intended construct and show a strong correlation with the total score. Conversely, invalid items exhibited correlation coefficients below the established threshold, indicating a need for revision or rendering them unsuitable for measuring students' critical thinking skills. Consequently, the 10 valid items were selected for use as the research instrument in the subsequent stage.

The graph illustrating the validity test results is presented below:

Large-Group Validity Test Results (Field Test)

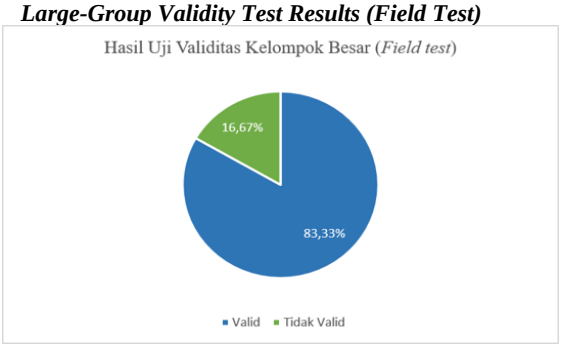


Figure 4. Large-Group Validity Test Results (Field Test)

Field Test Reliability Results

The reliability results for the large group, consisting of 30 students from class V.D, are as follows:

Table 2. Large Group Reliability Test Results

Reliability Statistics

Cronbach's Alpha	N of Items
.967	10

Based on the reliability test results presented in Table 2, Cronbach's Alpha value of 0.967 was obtained, falling within the "Reliable" category. This classification is based on a satisfactory reliability value that meets the established minimum coefficient threshold for a 10-item instrument. These results indicate that the instrument yields consistent measurements and meets the criteria for use. The reliability test was conducted on an instrument that had previously undergone validity testing using SPSS. The validity test identified 10 valid items and 2 invalid items. Consequently, the 10 valid items were deemed suitable for use as an instrument to measure the critical thinking skills of fifth-grade elementary students, as they met validity criteria and were supported by reliability results indicating the instrument was reliable.

3.2 Discussion

This study is titled "Development of Mathematical Problems to Measure Critical Thinking Skills in Fifth-Grade Elementary School Students." It aims to develop mathematical

problems designed to measure the critical thinking skills of fifth-grade students regarding data-related topics (data collection and presentation). The instrument was developed using the Tessmer model, comprising the following stages: preliminary, expert review, one-to-one, small group, and field test.

The results indicate that the developed instrument meets validity and reliability criteria, making it suitable for measuring students' critical thinking skills. Validation by three experts yielded an average score of 81.3%, placing it in the "valid" category. These results demonstrate that the instrument satisfies requirements regarding content, construction, and language, thereby qualifying it for the student trial phase. The findings show that the development process successfully produced an instrument aligned with the objective of measuring students' mathematical critical thinking skills.

During the one-to-one and small group stages, all test items were clearly understood by students following revisions based on expert feedback and readability assessments. Analysis confirmed that all items fell within the valid category, making them suitable for the field test stage. These findings demonstrate that the Tessmer model facilitates the gradual refinement of the instrument through a continuous revision process. This staged approach allowed for the correction of instrument weaknesses prior to large-scale testing.

The field test involving 30 students revealed that 83.33% of the test items were valid, while 16.67% were invalid. Of the 12 items tested, 10 met the validity criteria, whereas two required revision. These results indicate that the majority of the items effectively measured students' critical thinking skills in accordance with the established indicators. The fact that two items failed to meet validity criteria suggests that not all students grasped the intent of the questions or provided answers aligned with the expected indicators. Nevertheless, the instrument demonstrated overall high quality as a tool for measuring the mathematical critical thinking skills of elementary school students.

In summary, the research findings indicate that the developed mathematical test items met the criteria for a high-quality instrument. The instrument was not only valid—based on expert assessment and field test results—but also reliable for measuring the critical thinking skills of fifth-grade elementary students. These findings reinforce previous research (Lestari & Kurniawan, 2024) emphasizing the importance of developing mathematics assessment instruments capable of measuring higher-order thinking skills, as required by the *Kurikulum Merdeka* (Independent Curriculum). Therefore, the developed instrument serves as a viable alternative for evaluating mathematics learning and measuring critical thinking skills among elementary school students.

4. Conclusion

Based on the results of the research and development of mathematics problems designed to measure the critical thinking skills of fifth-grade elementary school students, the following conclusions can be drawn:

The developed mathematics problems demonstrate a good level of validity. Validation results from three validators yielded an average percentage of 81.3%, placing them in the "valid" category. Following the "one-to-one" and "small-group" stages—which served to assess readability and clarity and to refine the instrument—a validity test was conducted during the "field test" stage. Analysis revealed that 10 items (83.33%) met the validity criteria, while 2 items (16.67%) were invalid. Thus, the majority of the developed problems are capable of measuring students' critical thinking skills in accordance with the established indicators.

The developed mathematics problems demonstrate a good level of reliability. Reliability testing during the field test stage yielded a coefficient of 0.967, placing the instrument in the "reliable" category. These results indicate that the developed instrument possesses a high level of consistency, making it suitable for measuring the critical thinking skills of fifth-grade elementary school students.

References

[1] Agustine, D., Surya, E., & Zuhur, M. (2020). Pengembangan Instrumen Kemampuan Berpikir Britis Siswa Sekolah Dasar. *Jurnal Pendidikan Dasar*, 11(2), 45–56.

[2] Febriano, R., & Enawaty, E. (2021). Analisis Pembelajaran Matematika Berbasis Pemecahan Masalah Di Sekolah Dasar. *Jurnal Pendidikan Matematika Indonesia*, 6(1), 1–8.

[3] Hastuti, Y., Hapsari, P., & Gularso, D. (2024). Pengembangan Kepraktisan Instrumen Penilaian Berbasis Keterampilan Berpikir Kritis Siswa Kelas IV SD Negeri Wirosaban. *Jurnal Cendekia: Jurnal Pendidikan Matematika*, *08*, 1351–1357.

[4] Hollingshead, A. B. (2024). *Assessing Critical Thinking Skills In Elementary Education. Journal of Educational Assessment*, 18(2), 101–115.

[5] Iskandar. (2020). *Metodologi Penelitian Pendidikan Dan Sosial*. Jakarta: Referensi.

[6] Kemendikbudristek. (2022). *Capaian Pembelajaran Kurikulum Merdeka: Sekolah Dasar*. Jakarta: Kementerian Pendidikan, Kebudayaan, Riset, Dan Teknologi.

[7] Lestari, S., & Kurniawan, D. (2024). Pengembangan Asesmen Matematika Berbasis Kurikulum Merdeka Fase C. *Jurnal Pendidikan Dasar Indonesia*, 8(2), 95–105.

[8] Sapitri, D. (2025). *Systematic Literature Review*: Kemampuan Berpikir Kritis Dalam Pembelajaran Matematika. Gammath: Jurnal Ilmiah Program Studi Pendidikan Matematika, 10(1), 46-62.

[9] Sari, R., & Widodo, S. (2021). Pengembangan Instrumen Berpikir Kritis Matematis Di Sekolah Dasar. *Jurnal Pendidikan Matematika*, 12(1), 55–65.

[10] Trisnani, N., Retnawati, H., & Wuryandani, W. (2024). Pengembangan Instrumen Berpikir Kritis Berbentuk Soal Uraian. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 28(1), 60–75.