Jurnal Perspektif
Vol. 10 No. 1 Mei 2026
Page 152-165

Analysis of Thematic Exam Questions for Fifth Grade Elementary School Students, Theme VI, Even Semester in Sibolga City

Tiflah Addina Khairiah Nasution[1], Sumarno[2], Ajat Sudrajat[3]

[1,2,3] *Postgraduate Program in Elementary Education, Medan State University.*
Jl. Willem Iskandar, Pasar V Medan Estate, Deli Serdang, Medan, Sumatera Utara, Indonesia

[*] *addina_k@yahoo.com*

Received: 12 Januari 2026; Accepted: 28 Februari 2026; First Available Online 02 Maret 2026;
Published: 20 Mei 2026

DOI:10.15575/jp.v10i1.439

Abstrak

Tujuan dilakukannya penelitian ini adalah untuk mengetahui kualitas butir soal ujian yang diujikan pada siswa semester genap kelas V Sekolah Dasar. Penelitian ini merupakan penelitian kualitatif. Adapun pengumpulan data berupa soal ujian dan kunci jawaban siswa yang akan di teliti dengan memanfaatkan teknologi media Anates ver.0.9 untuk mengetahui taraf kualitas tes yang diujikan pada siswa. Hasil penelitian ini membuktikan bahwa terdapat tingkat kesesuaian yang cukup antara kisi-kisi soal yang telah disediakan dengan soal yang telah dievaluasi. Jika dilihat dari ciri-ciri butir soal berdasarkan Taksonomi Bloom pada soal ulangan semester genap, baik soal tematik maupun topik, banyak soal yang menggunakan ciri rendah, bahkan tidak ditemukan satu soal pun yang masuk ke dalam kategori ciri tinggi. Terdapat kontras atau perbedaan keefektifan antara soal ujian yang menggunakan tematik dan ujian yang menggunakan mata pelajaran. Jumlah soal valid yang berkaitan dengan topik tes tematik di SD Negeri 081234 dan SD Negeri 084081 lebih sedikit dibandingkan dengan soal yang berkaitan dengan mata pelajaran yang terdapat di SD Negeri 084087 dan SD Negeri 081224. Skor reliabilitas lebih besar untuk penilaian tingkat, serta untuk soal-soal yang lebih terfokus pada tema daripada mata pelajaran tertentu. Tingkat kesulitan soal ditentukan oleh banyaknya soal selain materi pelajaran yang termasuk dalam kategori sulit.

***Kata Kunci**: Analisis butir soal, kualitas tes, kualitatif.*

Abstract

The purpose of this study was to determine the quality of the exam questions administered to students in the even semester of grade 5 in elementary school. This study is a qualitative study. The data collection, in the form of exam questions and student answer keys, will be examined using Anates ver.0.9 media technology to determine the quality of the tests administered to students. The results of this study demonstrate sufficient correspondence between the provided question grids and the evaluated questions. When viewed through the characteristics of the questions based on Bloom's Taxonomy in the even semester exam, both thematic and topic questions, many questions use low characteristics; in fact, not a single question was found that fell into the high characteristic category. There is a contrast in effectiveness between exam questions that use themes and those that use subjects. The number of valid questions on thematic test topics in SD Negeri 081234 and SD Negeri 084081 is fewer than the number of questions on subjects in SD Negeri 084087 and SD Negeri 081224. Reliability scores are higher for

level assessments and for questions that focus more on themes than on specific subjects. The level of difficulty of the questions is determined by the number of questions, other than subject-matter questions, that are included in the difficult category.

Keywords: Item analysis, test quality, qualitative.

A. Introduction

Teachers are essential to education quality. Teachers may improve education in many ways. Teachers need assessment planning skills. Teachers must develop, administer, and evaluate learning outcomes in accordance with Law No. 14 of 2005, Article 20, paragraph 1. According to (Maulidya, D., Shaleh, S., & Rohmi, 2022) teachers must be honest and evaluate not just the product. Teachers must establish and develop techniques to evaluate student learning outcomes to be professional. Teachers must know how well kids learn. Unfortunately, teachers don't know their kids' reasoning abilities. Exam questions don't take into account pupils' higher-level and lower-level thinking skills. Lack of development of students' thinking skills results in a lack of critical thinking skills in problem-solving (Karo-Karo, M. S., Maulana, I., Nasution, S. K., 2023). Participants in learning evaluations always want to know the findings. They want to know if the exercise worked. Students and teachers want to know the results of their learning activities. Teachers must evaluate the procedure to determine its success. Evaluations include learning outcomes and evaluations (Muchlizani, N., Mania, S., & Rasyid, 2022).

Teachers test students after each lesson to assess their abilities. Tests at the end of each class assess students' comprehension. Create a student test measurement instrument for this testing. (Sukmawati, D., 2024) defines a test as a tool or technique used to determine or measure something in a specific situation using preset rules. Based on the principle above, while constructing a test, it is vital to follow established guidelines to maximize student benefit and accurately measure their abilities in a learning environment. (Darmawan, M., Sudarsono, Riyanti, D., Yuliana, Y. G. S., 2022) believes that a good measurement tool must be valid, reliable, objective, practicable, and economical. This theory requires teachers to construct assessments that meet these characteristics. (Nawir, M. S., Nur, S., Abubakar, M., Asmawati, A., 2023) defines a standardized exam as one with tested quality, validity, reliability, difficulty, discriminatory power, and distractor analysis. Once validity and reliability are confirmed, item characteristic analysis should be done. Item characteristic analysis includes discriminatory power, difficulty level, and distractor analysis for multiple-choice tests (Lutfina, A., Susapti, P., & Rahman, 2022).

Over time, society must think critically, creatively, logically, and rationally to filter information and address societal problems. Critical thinking may be used in daily life because everyone has difficulties, big and little. Critical thinking is essential for problem-solving. Students can learn critical thinking and problem-solving in primary school. Students must

actively develop knowledge through conceptual learning for advanced thinking. (Oktavia, F. T. A., Maharani, D., & Qudsiyah, 2023) aims to improve attitudes, knowledge, and skills to develop productive, creative, inventive, and effective workers (Hartati, N., & Suprayogi, 2021). Thus, the curriculum stresses constructivist scientific learning to reach this purpose. This should help pupils understand things, remember them, and understand learning.

Scientific learning involves assessment tools to examine higher-order thinking skills (HOTS) like analysis, synthesis, evaluation, and creativity (Marlinda & Budiman, 2025). Thus, questions must be crafted so students answer them using Bloom's taxonomy's operational verbs to test their thinking skills (Sijabat, M. P., Hutabarat, K., Sitorus, L., 2023). According to (Alfani, 2022) there are numerous principles for higher-order thinking questions. This involves measuring the test material using Bloom's cognitive domains—analyzing, evaluating, and producing (Yusuf, M., & Anwar, 2021). Each test question is based on reading for information to develop higher-order thinking. The School Exam (AS) includes questions that encourage higher-order thinking. Elementary school pupils are easily encouraged to think creatively and imaginatively. Critical thinking should start early. Students can develop higher-order thinking skills by mastering ower-order thinking and critical thinking (Asri, S. A., Syahrul, & Ishak, 2025).

Research demonstrates that those who constantly apply thinking skills finish tasks more often (Arifin, Z., & Retnawati, 2024). These abilities include lower-order and higher-order thinking. Six levels of Bloom's cognitive process dimensions, revised by Anderson & Krathwohl, include remembering, comprehending, applying, analyzing, evaluating, and creating. Provide practice questions that target lower-order and higher-order thinking to improve students' problem-solving skills. Exams are part of the graduate and advancement requirements (Sari, N., & Hidayat, 2023). Students need HOTS-level tests. It will improve their skills. Assessments include standardized and teacher-made tests. Professionals create and test standardized examinations numerous times before giving them to pupils, assuring accountability (Marlinda, S., & Budiman, 2025). Teaching tests are created by teachers who lack professional skills in test selection or who have expertise but lack the time to develop, evaluate, and assess them, underestimating their quality. Good tests have validity, reliability, objectivity, practicality, and economy (An-Vo, D. A., Mushtaq, S., Nguyen-Ky, T., Bundschuh, J., Tran-Cong, T., Maraseni, T. N., & Reardon-Smith, 2015).

Based on the author's interviews with class teachers at SD Negeri 081234 Sibolga, semester exams are structured in multiple-choice and essay formats. Question items are generally compiled using textbooks, student worksheets from various publishers, or revisions of previous years' questions, with a focus on measuring cognitive aspects, particularly understanding of the material. However, item analysis before and after the test is rarely conducted. This is influenced

by teachers' workloads, which encompass not only teaching activities but also administrative duties and other responsibilities (Kurniawan, D., 2024).

In fact, educational evaluation studies indicate that item analysis is an essential procedure for ensuring the quality of assessment instruments. Without analysis of difficulty, discriminatory power, validity, and reliability, test results may not accurately represent student abilities (Surapranata, 2016). Previous research has generally focused on secondary school levels or specific subjects, whereas research on the quality of thematic exam questions in elementary schools remains limited. Furthermore, empirical studies focusing on the local context of elementary schools in Sibolga are relatively rare.

The urgency of this research lies in the practical need to improve learning evaluation by providing empirical data on the quality of semester exam questions. The analysis of the Grade 5 thematic exam questions, Theme 6, for the Even Semester of the 2021/2022 Academic Year is expected to provide an objective overview of the quality of assessment instruments and serve as a basis for improving question development in elementary schools.

Thus, this research not only examines the quality of the test items but also provides practical contributions to teachers to improve the accuracy of measuring student learning outcomes. Based on the description above, a more in-depth analysis of thematic exam items for Grade 5 students, Theme 6, in the Even Semester of the 2021/2022 Academic Year in Sibolga City is deemed necessary.

B. Research Method

This research used a descriptive method, a naturalistic approach, and collected data in the form of words or pictures (Sugiyono, 2018). The objective of the research was to describe the quality of test items as they were without any treatment or manipulation of variables. This approach was chosen to enable objective analysis of the characteristics of the test questions teachers use, based on actual conditions in schools. Furthermore, a combination of qualitative (question construction and language) and quantitative (difficulty level, discriminating power, validity, and reliability) analysis was used to obtain a comprehensive picture of the quality of the learning evaluation instrument (Putra, R. A., & Sudrajat, 2022).

This research was conducted by analyzing the questions, examining their construction and language in a multiple-choice format. Quantitative analysis was also conducted, including difficulty level, discriminating power, reliability, and validation. The odd semester exam questions for grade 5 elementary schools in Sibolga City for the 2021/2022 academic year were used. Four schools were selected as research subjects: two schools with theme-based exam questions and two schools with subject-based exam questions. This research was conducted in the even semester of the 2022/2023 academic year, and was implemented in October 2022.

In this study, data was obtained from document collection. The documents used in this study were official school documents in the form of exam questions for the even semester of the 2021/2022 academic year. These documents were requested from homeroom teachers who had completed checking students' answer sheets. A research instrument is a tool used to measure observed natural or social phenomena (Sugiyono, 2016:148). A research instrument can also be defined as a tool used to obtain, manage, and interpret information from respondents using the same measurement pattern. Research instruments are designed for a single purpose and cannot be used in other studies (Fero, D, 2011:52).

C. Result and Discussion

Researchers aligned the question grids by assigning item numbers to each indicator. The following is the analysis of the correspondence between the question items and the provided grid indicators:

Table 1. Conformity of question items to the SDN 081234 Sibolga

No	Subjects	Basic competencies	Indicator	Question No.	Number of questions	Total (%)
1	PKN	1.2	1 2	1 2	2	7,70
		2.2	1 2	3 4	2	7,70
		3.2	1	5	1	3,84
		4.2	1	6	1	3,84
2	Indonesian	3.2	1	1 2	2	7,70
		3.3	1	3	1	3,84
			2	4	1	3,84
			3	5	1	3,84
3	IPA	3.6	1	1 2	2	7,70
			2	3	1	3,84
			3	4 5	2	7,70
4	IPS	3.2	1	1 2	2	7,70
			2	3	1	3,84
			3	4	1	3,84
			4	5	1	3,84

5	SBDP	3.1	1	12	2	7,70
		3.2	2	34	2	7,70
		3.3	3	5	1	3,84
		Total	26 test items			100

The data in the table of item suitability with the exam framework is presented based on an analysis of official documents obtained from SD Negeri 081234 Sibolga. The documents analyzed included the semester exam framework, multiple-choice test scripts, and the learning indicators listed for each core competency. The primary source of analysis was the match between the indicators in the exam framework and the competency requirements measured in each item.

The analysis process began with an inventory of all question items for each subject within a single exam theme. Each question item was then mapped to the core competencies and indicators listed in the exam framework. The mapping was carried out by matching the cognitive demands of the questions, the material tested, and the operational verbs used in the learning indicators. The results of this matching served as the basis for determining the number of question items for each indicator.

After the mapping process was complete, the number of question items representing each indicator was calculated. Frequency data was then converted to percentages to illustrate the proportion of indicator representation in the overall exam. Presenting the data in tabular form aims to systematically show the distribution of question items across each core competency and indicator and to facilitate interpretation of the instrument's suitability for planning learning evaluations.

Conclusions regarding suitability were determined based on the fulfillment of indicator representation in the question items and the absence of indicators tested outside the prepared grid. Thus, the results of the analysis in the table show that the questions used refer to the learning indicators planned in the question grid.

In the exam questions for SD Negeri 081234 in Sibolga City, there are 5 subjects in 1 theme, each covering several KD and Indicators. For Civics, there are 4 KD and 6 indicators. For Indonesian Language, there are 2 KD and 4 indicators; for Science, 1 KD and 3 indicators; for Social Studies, 1 KD and 4 indicators; and 3 KD and 3 indicators. Based on the existing grid and the questions tested, it can be concluded that the grid and questions for SD Negeri 081234 are in accordance with the requirements.

Table 2. Suitability of the question items to the SDN 084081 Sibolga

No	Subjects	Basic competencies	Indicator	Question format	Question No.	Number of questions	Total (%)
1	PKN	3.2	1	Multiple choice	1	1	17,5
			2	Multiple choice	2	1	17,5
			3	Multiple choice	3	1	17,5
			4	Multiple choice	4	1	17,5
			5	Multiple choice	5	1	17,5
			6	short answer	1	1	17,5
			7	short answer	2	1	17,5
			8	Essay	1	1	17,5
2	Bahasa Indonesia	3.3	1	Multiple choice	6	1	17,5
			2	Multiple choice	7	1	17,5
			3	Multiple choice	8	1	17,5
			4	Multiple choice	9	1	17,5
			5	Multiple choice	10	1	17,5
			6	short answer	3	1	17,5
			7	short answer	4	1	17,5
			8	Essay	2	1	17,5
3	IPA	3.6	1	Multiple choice	11	1	17,5
			2	Multiple choice	12	1	17,5

			3	Multiple choice	13	1	17,5
			4	Multiple choice	14	1	17,5
			5	Multiple choice	15	1	17,5
			6	short answer	5	1	17,5
			7	short answer	6	1	17,5
			8	Essay	3	1	17,5
4	IPS	3.2	1	Multiple choice	16	1	17,5
			2	Multiple choice	17	1	17,5
			3	Multiple choice	18	1	17,5
			4	Multiple choice	19	1	17,5
			5	Multiple choice	20	1	17,5
			6	short answer	7	1	17,5
			7	short answer	8	1	17,5
			8	Essay	4	1	17,5
5	SBDP	3.1	1	Multiple choice	21	1	17,5
		3.2	2	Multiple choice	22	1	17,5
		3.3	3	Multiple choice	23	1	17,5
			4	Multiple choice	24	1	17,5
			5	Multiple choice	25	1	17,5
			6	short answer	9	1	17,5

		7	short answer	10	1	17,5
		8	Essay	5	1	17,5
	Total		40 Question Items			100

The item-grid correspondence table shows that the semester exam questions at SD Negeri 084081 Sibolga were structured around the core competencies and planned learning indicators. Each subject within a theme is represented by several multiple-choice, short-answer, and essay questions, all of which are mapped to relevant indicators. This representation pattern indicates that learning evaluation planning has been systematically implemented in the development of the test instrument.

The distribution of question items within each subject indicates that each indicator receives relatively balanced representation throughout the test. This indicates that the material tested does not focus solely on one aspect of competency but encompasses various learning indicators targeted for student mastery. Thus, the test covers sufficient material to reflect mastery of overall competency. The presence of multiple question formats within each subject also plays a crucial role in measurement quality. Multiple-choice questions measure broad conceptual understanding, while short-answer and essay questions provide opportunities to assess students' ability to express their answers in greater depth. This combination of question formats demonstrates an effort to develop tests that assess learning outcomes quantitatively and consider the diversity of students' cognitive abilities.

The alignment between the test framework and the test items confirms that the assessment instrument has a strong foundation of content validity. Each test item corresponds to the designated indicators, ensuring that no content is assessed beyond the learning objectives. This condition supports the test's proper function as a tool for measuring student learning outcomes in accordance with learning objectives. Overall, the table's main implications indicate that the test questions were designed following a systematic learning evaluation plan, covering a range of representative indicators, and using various question formats to support a more comprehensive measurement of students' abilities. These findings provide a basis for assessing the suitability of the assessment instrument for evaluating learning outcomes, although further analysis of the empirical quality of the test items is still needed to assess its effectiveness in differentiating among students' abilities. The following is the suitability of the questions to the grid from SD Negeri 084087 Sibolga.

Table 3. Suitability of the question items to the SDN 084087 Sibolga

No	Subjects	Basic competencies	Indicator	Question format	Question No.	Number of questions	Total (%)

1	IPS (theme 6)	3.1	1 2 3 4 5	Multiple choice	1 2 3 4 5	1 1 1 1 1	6,7 6,7 6,7 6,7 6,7
2	IPS (theme 7)	3.4	1 2 3 4 5	Multiple choice	6 7 8 9 10	1 1 1 1 1	6,7 6,7 6,7 6,7 6,7
3	IPS (theme 8)	3.3	1 2 3 4 5	Multiple choice	11 12 13 14 15	1 1 1 1 1	6,7 6,7 6,7 6,7 6,7
			Total		15 Question Items		100

The information in the item-to-item correspondence table is based on an analysis of official documents from SD Negeri 084087 Sibolga. The documents analyzed included the semester exam outline, multiple-choice test papers for social studies for themes 6, 7, and 8, and the learning indicators listed for each core competency. The primary source of analysis was the relationship between the indicators in the outline and the competency requirements measured by each item.

The analysis process began with document collection and verification to ensure that the outline and test papers came from the same exam period. Next, all test items were inventoried and grouped based on themes and core competencies. Each item was then mapped to relevant indicators by matching the material tested, cognitive demands, and operational verbs to the learning indicators. The mapping results served as the basis for determining the item number for each indicator.

After the mapping was completed, the number of test items representing each indicator for the core competency was calculated. The frequency of indicator representation was then converted to a percentage to illustrate each indicator's contribution to the overall test. The purpose of presenting data in tabular form is to systematically demonstrate the distribution of test items across themes, core competencies, and indicators, making it easier for readers to understand the scope of the material being measured.

Suitability is determined by examining whether each test item refers to the indicators listed in the test grid and whether any indicators are tested outside the planned scope. Furthermore, checking the consistency of the test format with the indicator characteristics is also considered

in interpreting the results. Based on these steps, the data presented in the table represents the results of an empirical mapping between the test grid and the test items used in the exam. Thus, the tabular presentation not only displays the number and distribution of test items but also illustrates the analytical process that confirms that the assessment instrument has been designed in accordance with the planned learning indicators. The following table shows the correspondence between the questions and the outline for SD Negeri 081224 Sibolga

Table 4. Conformity of question items to the SDN 081224 Sibolga

No	Subjects	Basic competencies	Indicator	Question format	Question No.	Number of questions	Total (%)
1	IPS (theme 6)	3.1	1	Multiple choice	7	1	6,7
2	IPS (theme 7)	3.4	1	Multiple choice	1	1	6,7
			2		2	1	6,7
			3		8	1	6,7
			4		12	1	6,7
3	IPS (theme 8)	3.4	1	Multiple choice	3	1	6,7
			2		4	1	6,7
			3		9	1	6,7
			4		13	1	6,7
4	IPS (theme 9)	3.4	1	Multiple choice	5	1	6,7
			2		10	1	6,7
			3		6	1	6,7
			4		14	1	6,7
			5		15	1	6,7
			6		11	1	6,7
			Total	15 Question Items			100

From the data above, it can be concluded that there are 15 test items for students and 4 core competencies in a single subject, each with 15 indicators. The questions tested by SD Negeri 081224 are social studies questions. Therefore, the data obtained is stated to be in accordance with the question outline for the questions tested.

Discussion

HOTS test items have distinctive characteristics. The levels encompass students' abilities or skills in analyzing, evaluating, and creating. Indicators of analysis, evaluation, and creation skills are based on the theory outlined in the revised Bloom's Taxonomy. The results of the analysis of 26 test items from SD Negeri 081234 Sibolga are as follows: a) Low characteristics:

20 low-characteristic test items, b) Medium characteristics: 6 medium-characteristic test items, c) High characteristics: 0 characteristic test items.

The results of the analysis of the characteristics of 40 test items from SD Negeri 084081 Sibolga are: a) Low characteristics: 29 low-characteristic test items, b) Medium characteristics: 11 medium-characteristic test items, c) High characteristics: 0 characteristic test items. The results of the characteristic analysis of 15 questions in the Sibolga 084087 Elementary School exam are: a) Low characteristics: 15 low characteristic items, b) Medium characteristics: 0 medium characteristic items, c) High characteristics: 0 characteristic items. The results of the characteristic analysis of 15 questions in the Sibolga 0841224 Elementary School exam are: a) Low characteristics: 12 low characteristic items, b) Medium characteristics: 3 medium characteristic items, c) High characteristics: 0 characteristic items. This is in line with several previous studies examining the quality of learning evaluation instruments in elementary schools. Research conducted by (Pratiwi, I., & Widoyoko, 2023) on the analysis of thematic exam items for grade 5 elementary school showed that not all items met the criteria for good validity and discrimination.

These findings confirm that the preparation of thematic exams still requires improvement, particularly in planning the question framework and aligning competency indicators with the exam format. These results are relevant to research in Sibolga City, which also found an imbalance in the quality of exam questions, impacting the accuracy of measuring student learning outcomes. Furthermore, research by (Rahmawati, D., 2022), which analyzed HOTS-based exam questions in elementary school thematic learning, showed that most of the questions teachers used were still at the lower cognitive level (LOTS). This aligns with research findings in Sibolga City, which demonstrated the need to increase the proportion of HOTS (Higher Order Thinking Skills) questions so that learning evaluations measure not only memory and comprehension skills but also analytical, evaluation, and problem-solving abilities.

Another study by (Wulandari, R., & Mardapi, 2021) which used a Rasch model approach in thematic item analysis, found that more complex analysis techniques provided a more accurate picture of item quality, including item difficulty and discrimination power. These findings reinforce research in Sibolga City, which recommended using more comprehensive analysis techniques than conventional methods to objectively and measurably improve the quality of evaluation instruments. Furthermore, research by (Setiawan, A., 2023) on item analysis for elementary school final exams concluded that item reliability is significantly influenced by the consistency of indicator development, the quality of distractors, and the balance of difficulty levels. These findings are relevant to research in Sibolga City, which indicated that item reliability still needs to be improved through improvements in framework design and instrument testing before use in evaluations.

Based on several relevant studies, it can be concluded that the problems found in the analysis of thematic exam questions for Grade 5, Theme 6 in Sibolga City are not isolated cases, but rather a common phenomenon in the practice of thematic learning evaluation in elementary schools. Therefore, systematic efforts are needed, such as developing a structured grid, conducting regular item analysis, using more accurate analysis techniques, and increasing the proportion of HOTS (Higher Order Thinking Skills) questions so that evaluation instruments can truly measure student competencies validly, reliably, and meaningfully.

D. Conclusion

Based on the research results and discussion of the analysis of even-semester exam items, the following conclusions were reached: There is a gap between thematic exam questions and subject-based exam questions. For thematic exam questions at Public Elementary School 081234 and Public Elementary School 084081, more valid questions were found than at Public Elementary School 084087 and Public Elementary School 081224, which had more subject-based questions. Thematic questions had higher reliability scores than subject-based questions. Similarly, for the difficulty level test, more themed questions fell into the difficult category than subject-based questions. Thematic questions, however, were more likely than subject-based questions to be classified as poor for the discrimination power test. The correspondence between the provided question outline and the tested questions can be considered appropriate. Based on Bloom's Taxonomy, thematic and subject-based exam questions, many questions had low characteristics, and not a single question fell into the high characteristics category.

References

Alfani, M. F. (2022). Analisis butir soal ulangan akhir semester pada Kurikulum Merdeka di SD/MI. Sasangga. *Journal of Education and Learning*, *1*(1), 1–10.

An-Vo, D. A., Mushtaq, S., Nguyen-Ky, T., Bundschuh, J., Tran-Cong, T., Maraseni, T. N., & Reardon-Smith, K. (2015). Nonlinear Optimisation Using Production Functions to Estimate Economic Benefit of Conjunctive Water Use for Multicrop Production. *Water Resources Management*, *29*(7), 2153–2170. https://doi.org/https://doi.org/10.1007/s11269-015-0933-y

Arifin, Z., & Retnawati, H. (2024). Analisis kualitas butir soal penilaian hasil belajar berbasis HOTS di sekolah dasar. *Jurnal Evaluasi Pendidikan*, *15*(1), 1–14.

Asri, S. A., Syahrul, & Ishak, S. (2025). Analisis kualitas butir soal ulangan akhir semester genap matematika. *Pedagogy: Jurnal Pendidikan Matematika*, *10*(1), 55–66.

Darmawan, M., Sudarsono, Riyanti, D., Yuliana, Y. G. S., & S. (2022). Test-items analysis of English teacher-made test. *Journal of English Education and Teaching*, *6*(3), 412–426.

Hartati, N., & Suprayogi, H. P. (2021). Item analysis for a better quality test. *English Language in Focus (ELIF)*, *3*(2), 93–104.

Karo-Karo, M. S., Maulana, I., Nasution, S. K., & N. (2023). Analisis butir soal sumatif akhir semester genap PAI berdasarkan Taksonomi Bloom di kelas VI SD Islam Al-Abid. BADA'A. *Jurnal Ilmiah Pendidikan Dasar*, *5*(2), 89–100.

Kurniawan, D., & S. (2024). Analisis instrumen penilaian pembelajaran tematik berbasis kompetensi. *Jurnal Pendidikan Dasar Nusantara*, *10*(2), 95–107.

Lutfina, A., Susapti, P., & Rahman, A. (2022). Test item quality analysis of final mathematic assessment

in Islamic elementary school. *MUDARRISA: Jurnal Kajian Pendidikan Islam*, *14*(1), 63–78.
Marlinda, S., & Budiman, I. (2025). Analisis Kompetensi Strategis Matematis Siswa Sekolah Menengah Kejuruan Pada Materi Barisan dan Deret Aritmatika. *Kognitif: Jurnal Riset HOTS Pendidikan Matematika*, *5*(1), 154–166.
Marlinda, S., & Budiman, I. (2025). Analisis Kompetensi Strategis Matematis Siswa Sekolah Menengah Kejuruan Pada Materi Barisan dan Deret Aritmatika. *Kognitif: Jurnal Riset HOTS Pendidikan Matematika*, *5*(1), 154–166.
Maulidya, D., Shaleh, S., & Rohmi, P. (2022). An analysis of elementary school thematic question points using Iteman software and Winsteps Rasch model. Al-Adzka. *Jurnal Ilmiah Pendidikan Guru Madrasah Ibtidaiyah*, *12*(2), 115–126.
Muchlizani, N., Mania, S., & Rasyid, M. N. A. (2022). Analisis kualitas butir soal ujian akhir semester mata pelajaran Akidah Akhlak kelas V MI. *Inspiratif Pendidikan*, *11*(2), 210–221.
Nawir, M. S., Nur, S., Abubakar, M., Asmawati, A., & M. (2023). The analysis of item difficulty and discrimination of English summative teacher-made test. *ELSTIC: English Language, Linguistics, and Culture International Journal*, *3*(1), 1–12.
Oktavia, F. T. A., Maharani, D., & Qudsiyah, K. (2023). Problematika Penerapan Kurikulum Merdeka Belajar Pada Pembelajaran Matematika Di Smk Negeri 2 Pacitan. *Jurnal Edumatic : Jurnal Pendidikan Matematika*, *4*(2), 14–23.
Pratiwi, I., & Widoyoko, E. P. (2023). Analisis tingkat kesukaran dan daya pembeda soal tematik kelas V sekolah dasar. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *27*(2), 158–170.
Putra, R. A., & Sudrajat, A. (2022). Analisis butir soal pilihan ganda menggunakan model Rasch pada evaluasi pembelajaran. *Jurnal Pendidikan Indonesia*, *11*(4), 620–630.
Rahmawati, D., & S. (2022). Validitas isi instrumen evaluasi pembelajaran tematik sekolah dasar. *Jurnal Pendidikan Dasar Indonesia*, *7*(1), 45–56.
Sari, N., & Hidayat, R. (2023). Evaluasi kualitas soal ujian semester sekolah dasar. *Jurnal Pendidikan Dan Pembelajaran*, *30*(2), 210–222.
Setiawan, A., & H. (2023). Kualitas butir soal ujian sekolah berbasis kurikulum 2013 di sekolah dasar. *Jurnal Basicedu*, *7*(3), 1789–1799.
Sijabat, M. P., Hutabarat, K., Sitorus, L., & S. (2023). Analisis soal tes hasil belajar siswa berstandar nasional Bahasa Indonesia kelas V SD. *Jurnal Basicedu*, *7*(5), 3061–3071.
Sugiyono. (2018). *Metodologi Penelitian Pendidikan (Pendidikan Kuantitatif, Kualitatif dan R&D)*. Alfabeta.
Sukmawati, D., & K. (2024). Analisis tingkat kesukaran butir soal penilaian tengah semester PPKn kelas V sekolah dasar. *PROSA: Jurnal Penelitian Pendidikan Guru Sekolah Dasar*, *8*(1), 34–43.
Surapranata, S. (2016). *Analisis, Validitas, Reliabilitas dan Interpretasi Hasil Tes.* PT. Remaja Rosdakarya.
Wulandari, R., & Mardapi, D. (2021). Analisis reliabilitas tes hasil belajar menggunakan pendekatan klasik. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *25*(1), 88–100.
Yusuf, M., & Anwar, K. (2021). Analisis daya pembeda butir soal pada evaluasi pembelajaran. *Jurnal Pendidikan Matematika*, *15*(1), 55–66.