



Development of A Science Literacy Test Instrument Based on Local Wisdom in Kediri on The Topic of Energy For Fourth Grade Elementary School Students

Firdalia Lisa Anggraini¹, Lilik Bintartik², Rika Mellyaning Khoiriya³

State University of Malang, Jl. Semarang No. 5 Malang, East Java, Indonesia

*Author of correspondence, Email: lilik.bintartik.fip@um.ac.id

Abstract

This study aimed to design and develop a science literacy test instrument based on local wisdom in Kediri for fourth-grade elementary students, specifically on energy topics. The development followed the ADDIE model and assessed the instrument's content validity, practicality, and psychometric properties, including reliability, difficulty index, discrimination power, and distractor effectiveness. Expert validation indicated strong content validity, with scores of 100% from subject matter experts and 94.44% from evaluation experts. In a small-scale trial, 33 items were valid and 17 invalid. Six items showed high reliability, and 27 demonstrated very high reliability. The average difficulty index was 0.718 (easy), and the discrimination index was 0.418 (good). Practicality ratings were 89.29% from teachers (very practical) and 78.5% from students (practical). A largescale trial confirmed 32 valid items and one invalid. Eight items showed very high reliability and 25 high reliability. The average difficulty index was 0.6546 (moderate), and the discrimination index was 0.5571 (good). Practicality scores were 79.6% from students (practical) and 86.9% from teachers (very practical). These results suggest that the instrument is valid, reliable, and practical for assessing science literacy in elementary education.

Keywords: Test Instrument; Science Literacy; Local Wisdom; Elementary School

1. Introduction

Literacy is a skill that everyone should master. According to Indonesian Law No. 3/2017, literacy is the ability to interpret information critically. Literate people can access This science and technology to improve their quality of life. Ali, Setiawan, Joko, Ulumuddin, & Julizar (2019) also said that literacy is the axis of efforts to improve the quality of human life. The literacy that 21st century humans need to master is called basic literacy. The Ministry of Education and Culture (2019) also said that basic literacy such as reading and writing, numeracy, understanding science, digital, financial, cultural and citizenship are important elements of 21st century skills. Science literacy is the ability of individuals to use knowledge to identify questions, build new understanding, explain phenomena scientifically, and develop reflective thinking in order to participate in solving problems and responding to various issues or ideas related to science (OECD, 2019).

In science learning, science literacy is integrated in the planning and implementation of learning. Science learning is carried out with the orientation of forming science literacy in students. So with this process, it is expected to form science literacy in students, so that science learning contributes to producing students who are literate in science. With science literacy, individuals are encouraged to develop scientific mindsets, responsible behavior, and characters who care about themselves, the environment, social, nature, and are ready to answer various challenges in the modern era which is strongly influenced by dependence on technology

(Fananta et al., 2017). Through science literacy learning, students are expected to gain critical thinking skills, logical, creative, motivational, build character, and be globally competitive (Irsan, 2020). According to Pertiwi, Atanti, and Ismawati (2018), the aims of science literacy instruction include: (1) enhancing students' conceptual understanding and investigative abilities in the field of natural sciences; (2) expanding both oral and written scientific vocabulary essential for interpreting and communicating scientific ideas; and (3) reinforcing the interrelationship between science, technology, and society. Judging from the importance of science literacy which is one of the keys in dealing with the dynamics and problems of the 21st century, then learning science literacy is also important and needs to be mastered and understood by ab students.

Science literacy learning provided to students in schools has a connection between science, technology and the surrounding environment. So that the presentation of local wisdom can provide innovation in science learning. Apriyulianti, Istiningsih, Rahmatih, & Astria (2023) said that integrating local wisdom into science learning can provide an innovative approach, where the science curriculum is designed by incorporating local wisdom values. It is important to integrate local wisdom into the education curriculum (Anggraisa, Nurlidiya, Sativa, Kholiza, & Putri, 2024). This is because it can increase awareness and appreciation of cultural diversity, and strengthen national identity. Integrating local culture into the curriculum contributes to helping maintain the integrity and preservation of a custom and culture in society. This is in accordance with the recommendation of the Ministry of Education and Culture's independent learning curriculum that learning should start from the environment around students (Ikhsan & Rachmadyanti, 2024). Local wisdom is certain principles and methods that are believed, understood, and applied by local communities in building their relationships with the surrounding environment and the results of traditions that live and develop for generations in the midst of indigenous communities. (CHAIRUL, 2019). Therefore, it is important to implement science literacy based on local wisdom in elementary school.

Quoted from the Kemedikbud website, science material in elementary school (SD) covers several chapters, one of which is the chapter on changing the form of energy (Fitri et al., 2023). The basic human need is energy, humans will not live without energy (Soeparman, 2015). To meet energy needs, it is necessary to implement energy sources into a form of energy that can help humans in their daily activities. Energy is the ability to do effort or work (Afriyanti, 2018). Science literacy based on local wisdom in the field of energy needs to be known so that the abilities and competencies of students become better at a higher level. This research will integrate local wisdom from the Kediri region. Because in the Kediri region there is a variety of local wisdom that can be analyzed concepts on energy material.

At the international level, scientific literacy assessments in education are primarily represented by two major programs: TIMSS (Trends in International Mathematics and Science Study) and PISA (Programme for International Student Assessment). The results of the TIMSS survey in 2015 involving 4th grade elementary school students, stated that Indonesia ranked 44th out of a total of 49 participating countries with an average score of 397, which is still below the average international score of 500 (Hadi & Novaliyosi, 2019). The results of the 2022 PISA survey also showed that Indonesia's science literacy score in 2018 was 396 and the science literacy score in 2022 was 383. Although the score is low, the ranking comparison has increased by 5-6 positions compared to 2018 (Kemendikbudristek, 2023; OECD, 2023).

The government in Indonesia has thought about the science literacy of students. This can be seen from the government's efforts to transform the curriculum (Rahman, 2024). At the end of 2021, the independent curriculum emerged. Schools that are selected to implement the independent curriculum are called driving schools. The existence of this driving school is expected to improve students' scientific literacy (Aisah et al., 2021). Where curriculum adjustments are made to local needs and the characteristics of students (Murtadlo, Oktafiani, & Faizah, 2024). One of the Ministry of Education and Culture's strategies for conducting assessments is the global assessment standard using AKM. However, in the implementation at school, there is no science literacy test instrument that measures students' science literacy.

One of the factors that cause low science literacy skills in Indonesia is the limited assessment instruments that refer to aspects of science literacy and do not support the development of higher order thinking skills (Murti & Sunarti, 2021). In addition, effective efforts in helping students understand science principles more deeply are carried out by developing science literacy test instruments (Astuti, Prasetyo, & Rahayu, 2014). Measurement of science literacy has an important role in knowing the extent to which students master science literacy, so that efforts to improve the quality of education in Indonesia can be carried out effectively and encourage the nation's competitiveness in the international arena (Umi Roufatuz Zahro, 2020). Therefore, it is important to develop a science literacy test instrument.

Scientific literacy instrument development has been explored in a study conducted by Putra, et al (2021) who conducted the Development of Critical Thinking Ability Instruments and Science Literacy Skills for Grade IV Elementary School Students, Research by Mardhiyyah et al (2016) who conducted the Development of Science Literacy Assessment Instruments on Energy Themes, Murti, et al (2021) who conducted the Development of Science Literacy Test Instruments Based on Local Wisdom in Trenggalek, and Maulida et al (2022) who conducted the Development of Science Literacy Test Instruments Based on Local Wisdom in Lamongan Regency. Some research results show that the level of science literacy of students is still relatively low. Therefore, it is necessary to develop a science literacy instrument that aims to motivate and familiarize students in solving science literacy-based questions, so that it is hoped that their abilities can increase.

Based on the results of interviews conducted with a number of homeroom teachers in Kediri, information was obtained that (1) homeroom teachers have never applied local wisdom-based science literacy test instruments to the test questions they use, and (2) the test questions used to measure student learning outcomes have not been based on science literacy. So it can be said that there is no assessment instrument used to measure students' science literacy. In making questions, classroom teachers usually use teacher's book guides and package books. Making questions is adjusted to the learning outcomes and indicators in each class. According to the opinion of the fourth grade teacher with the researcher's plan to develop a science literacy test instrument is very supportive. Even from one of the fourth grade teachers with the initials PT later wanted to learn and wanted to be taught more about the development of science literacy test instruments. Because with this test instrument can be used to measure the competence of students as well as train them to get used to facing various types of questions. That way it will improve the competence of students at the next level.

Based on the results of the learner needs questionnaire given by fourth grade students from three schools including SDN Pagu 1, SDN Pagu 2, and SDN Jabon 1. The results showed

that 96% of students knew several examples of local wisdom in Kediri. As many as 74% of students have learned related to Kediri local wisdom. A total of 98% of students thought that science lessons based on Kediri local wisdom could help them understand the material better. As many as 93% of students felt that local wisdom-based science learning was an interesting topic, and as many as 92% of students felt that working on local wisdom-based questions would be interesting.

Based on this needs analysis, the development of a science literacy test instrument entitled “Development of Science Literacy Test Instruments Based on Local Wisdom in Kediri Energy Material for Grade IV Elementary School” was carried out. This study aims to produce a science literacy test for grade IV elementary school students that is validated logically (by material experts and evaluation experts), empirically (through validity, reliability, difficulty level, differentiator power, and exemption), and practically according to teachers and students.

2. Method

This study employed the ADDIE development model in constructing the test instrument, encompassing five essential phases: needs analysis, design, development, implementation, and evaluation. This research was conducted for grade 4 at SDN Pagu 1, SDN Pagu 2, and SDN Jabon 1 on May 3 and May 10, 2025. Data collection was carried out using interviews, questionnaires, documentation and literature studies. Data sources in this study were obtained through expert validation (materials and evaluation), small and large scale tests and teacher and student responses.

There are 2 types of data analysis conducted in this study, namely qualitative and quantitative. Qualitative data analysis is obtained from interview results, conclusions from material expert questionnaire results, conclusions from evaluation expert questionnaire results, and conclusions from user practicality questionnaire results. For quantitative data analysis, it is obtained from the analysis of the validity and practicality of the instrument. The validity test is obtained from the validation questionnaire of material experts and evaluation experts and, the analysis of the test items includes validity tests, reliability, difficulty level analysis, discriminatory power analysis, and distractor analysis. While the practicality of the instrument is obtained from the practicality questionnaire of teachers and students. After the data is collected, the validity data is analyzed using the formula proposed by Akbar (2013), as follows.

$$Vah = \frac{Tse}{Tsh} \times 100\%$$

Description: Vah = Expert validation results

Tse = total score obtained

Tsh = maximum score that can be achieved

The data obtained were analyzed and classified into the following categories.

Table 1 Guidelines for scoring the validity assessment of test instruments

Assessment	Assessment Criteria
86,00-100,00	Very valid
71,00-85,00	Valid
56,00-70,00	Quite valid
41,00-55,00	Less valid
25,00-40,00	Not valid

Source: Akbar (2013)

Based on table 1, it can be concluded that if the validity percentage is 71% or above, a minor revision will be made, but if the validity percentage is less than 71%, a major revision will be made. This is adjusted to the criticism, suggestions, and input from the team of material experts and evaluation experts.

The item analysis process was carried out using MS Excel 2010, and the analysis criteria can be seen in Table 2 below.

Table 2 Question Analysis Criteria Guidelines

Characteristics	Category	Criteria
Item validity	Valid	$r_{xy(\text{hitung})} > r_{xy(\text{tabel})}$
	Not valid	$r_{xy(\text{hitung})} < r_{xy(\text{tabel})}$
Reliability	Very low	$< 0,20$
	Low	$0,21-0,40$
	Currently	$0,41-$
		$0,60$
	Tall	$0,61-$
		$0,80$
Difficulty level	Very high	$0,81-1,0$
	Difficult	$0,00-$
		$0,30$
	Medium	$0,31-$
		$0,70$
	Easy	$0,71-$
		$1,00$
Discriminating power	Very good	$0,71-$
		$1,00$
	Good	$0,41-$
		$0,70$
	Fair	$0,21-$
		$0,40$
	Poor	$0,00-$
		$0,20$
	Very bad	Negative
Distractor	Accepted because it is good	$D \geq 5\%$
	Revised by rewriting because it is not good	$5\% >$
		$D > 0$
	Rejected because it is not good	$D =$
		0

The next test is the practicality obtained from the data collected through the teacher and student response questionnaire interpreted using a formula referring to Parmin (2012) as follows.

Table 3 Practicality Criteria

Criteria	Score
$80\% < P \leq 100\%$	Very practical
$60\% < P \leq 80\%$	Practical
$40\% < P \leq 60\%$	Quite practical
$20\% < P \leq 40\%$	Less practical

$0\% \leq P \leq 20\%$

Not practical

Theoretically, a question is said to be practical if the practical percentage is ≥ 61 .

3. Results and Discussion

3.1 The results of the validity of the local wisdom-based science literacy test instrument in Kediri for energy material for grade IV elementary school. The following is the recapitulation of validation results from all aspects validated by material experts and evaluation experts. Both experts are lecturers at Malang State University. The recapitulation of the overall logical validity is presented as follows Table 4 recapitulation of expert validation results

Aspects	Assessment items	Validation by	
		Subject matter expert	Evaluation expert
Presentation	Use of language that is easy for students to understand	100	-
	Presentation of questions in accordance with local wisdom in Kediri	100	-
	Suitability of question categories	100	-
	As a practical and efficient evaluation instrument	100	-
Content quality	Conceptual accuracy	100	-
	Question accuracy	100	-
	Measuring local wisdom-based test instruments	100	-
Construction	Suitability of questions according to local wisdom-based test instruments	100	-
	Order of presentation of questions	100	-
Question grid	Providing learning motivation	100	-
	Can be used as a guideline for designing the writing of test items	-	100
Question items	Questions are made according to local wisdom-based test instruments.	-	100
	Questions made cover local wisdom of Kediri	-	100
	Questions made do not contain misconceptions	-	75
	Sentence formulation in the form of questions or commands that require answers	-	100

Proceedings Series of Educational Studies

Assessment rubric and answer key	Pictures or similar are meaningful (clear description or related to the problem being asked)	-	75
	Matching of answer key with questions	-	100
	Appropriate weighting	-	100
	Final scoring is clear and detailed	-	100
Amount		1000	850
Average		100	94,44
		Very valid	Very valid

Based on the data in Table 4, the results of the validation by material experts show a score of 100% with a very valid category, and the instrument can be used after improvements are made according to suggestions and input. The suggestions and input of the material expert validator are that the questions can be included in the grid at the same time, and the question indicators can be written in a directed manner for each question. The suggestions from the material experts are in line with the opinion of Putra, Adzim, & Hilmiyati, (2025) who stated that in the grid, questions can be designed according to the indicators and question formats that have been determined.

There are three factors included in the material expert criteria, namely presentation aspects, content quality, and construction. The presentation aspect contains 4 statements, namely (1) the use of language that is easy for students to understand, (2) presentation of questions in accordance with local wisdom in Kediri, (3) suitability of question categories, and (4) as a practical and efficient evaluation instrument. In this aspect, the material expert team gave a percentage of 100% with a very valid category. One of the requirements for a good science literacy test is to adjust the science literacy indicators with learning indicators and also adjust them to everyday life (Mustofa, Kuswanti, Siti, & Hidayati, 2017).

In terms of content quality, there are 3 statements, namely (1) conceptual accuracy, (2) question accuracy, and (3) measuring local wisdom-based test instruments. In this aspect, the material expert team gave a percentage of 100% with a very valid category. This is in line with the opinion of Janah & Hidayati (2020) that understanding science concepts has an important role in science learning in the independent curriculum, where students are not only required to understand scientific facts, ideas, and theories, but also the processes used to find them.

In the construction aspect, there are 3 statements, namely (1) the suitability of the questions according to the local wisdom-based test instrument, (2) the order of presentation of the questions, and (3) providing learning motivation. In this aspect, the material expert team gave a percentage of 100% with a very valid category. In line with the opinion of Maya

Ningsetyo & Titin Sunarti (2024) that the test instrument that was prepared was adjusted to the indicators of scientific literacy competencies and the material used as a reference. Integrating local wisdom in learning has benefits such as strengthening the relevance of the material to students' daily learning experiences, strengthening students' cultural identities, and increasing their motivation in learning (Nahak, 2024).

The results of the evaluation expert's calculations obtained a percentage of 94.44% with a very valid category. The evaluation expert's suggestions and input are, namely (1) if possible add images, pay attention to HOTS (c4-c6), (2) make 2 x from the target, for example target 10 for question 20, (3) emphasize local wisdom in Kediri (image). This is in line with Karnali's opinion (2019) that HOTS-based questions can train students to think critically and analytically to solve existing problems (Kafi & Ilma, 2022).

There are three factors included in the material expert criteria, namely the question grid, test questions, and assessment rubrics and answer keys. The question grid aspect contains a statement that it can be used as a guideline for designing the writing of question items. In the question grid aspect, the percentage value is 100% and is in the very valid category. According to Putra, Adzim, & Hilmiyati, (2025) the grid is made so that the questions that are compiled are in line with the material being taught, are able to measure abilities according to learning objectives, and ensure that every aspect of the learning objectives is covered in the evaluation instrument.

In the second aspect, the question items have statements, namely (1) questions made in accordance with local wisdom-based test instruments, (2) questions made cover Kediri local wisdom, (3) questions made do not contain misconceptions, (4) sentence formulation in the form of questions or commands that require answers, and (5) images or similar are meaningful (clear explanations or are related to the problem being asked). In this aspect, a percentage of 90% is obtained, categorized as very valid. This is in line with the opinion of Safitri, Sarwanto, & Harjunowibowo (2023) that the images used need to be adjusted to the theme of local wisdom raised and relevant to events related to learning materials. Also in line with the research of Khoiriya et al. (2024) that the importance of developing questions is adjusted to the local wisdom raised in learning.

In the aspect of the assessment rubric and answer key, there are 3 statements, namely (1) the suitability of the answer key to the question, (2) appropriate weighting, and (3) the final scoring is clear and detailed. In this aspect, the percentage value is 100%. This is in line with Yulyantari's opinion (2017) that the process of giving a score can be adjusted to the existing answer key. The validator's suggestions and input are (1) if possible, add pictures, (2) pay attention to HOTS (c4-c6), (3) emphasize local wisdom in Kediri (pictures). (4) make 2 times from the target, for example, target 10 for question 20. This is in accordance with Karnali's

view (2019) which states that HOTS-based questions can train students in critical analytical thinking to solve existing problems (Kafi & Ilma, 2022).

3.2 Recapitulation of the results of the analysis of questions on the local wisdom based science literacy test instrument in Kediri on energy material for grade

IV elementary school

Small and large scale tests are used to evaluate test instruments that have been refined by experts. To produce a valid science literacy test instrument, 50 questions in the form of multiple choice, complex multiple choice, matching, short form, and description were analyzed and went through a series of trials, including validity, reliability, difficulty level, differentiating power, and exemption. The results of item analysis from small and large scale trials are summarized as follows.

3.2.1 Small Scale Test

The small-scale test was conducted at SDN Pagu 2 with 15 grade 4 students as subjects. Small trials are conducted to minimize the occurrence of errors in test instrument products and student answers to determine whether there are valid questions that can be used and which need to be discarded. The results of the small-scale test recap are 33 valid questions and 17 invalid questions, where invalid questions will not be used and will be discarded. The reliability results on multiple choice questions get a value of 0.823476 concluding very high reliability, the reliability results of complex multiple choice questions get a value of

0.75136068 concluding high reliability, the reliability results of matching questions get a value of 0.824046 concluding very high reliability, the reliability results of short fill questions get a value of 0.902840406 concluding very high reliability, and the reliability results of description questions get a value of 0.859332626 concluding very high reliability.

Furthermore, the analysis of the level of difficulty of 15 multiple choice questions in Roman I section resulted in 2 difficult questions, 6 medium questions, and 7 easy questions. The level of difficulty of complex multiple choice questions in Roman II out of 10 complex multiple choice questions, there are 2 medium questions, and 8 easy questions. In Roman III with matching questions out of 10 questions, 5 medium questions and 5 easy questions were obtained. In Roman IV with short fill-in questions out of 7 questions, 2 medium questions and 5 easy questions are obtained. In Roman V with description questions out of 8 questions, 4 easy questions, 3 medium questions, and 1 difficult question were obtained.

In the test of the differential power of Roman I multiple choice questions, 5 questions were obtained very good, 2 good questions, 5 fair questions, and 3 bad questions. In Roman II complex multiple choice questions, 1 question is very good, 4 questions are good, 2 questions are sufficient, and 3 questions are bad. In Roman III, matching questions obtained 5 good questions, 2 sufficient questions, and 3 bad questions. In Roman IV short fill-in questions, 2

questions are very good, 4 questions are sufficient, and 1 question. In Roman V, the description question obtained 1 question with excellent differentiating power, 3 questions with good differentiating power, 1 question is sufficient, 2 bad questions, and 1 bad question. Each item analysis calculation result in the small-scale test will be explained in detail below.

The empirical validity of test items is analyzed to ensure that the questions that have been compiled and tested are truly valid. The validity test was carried out using MS Excel using Pearson Correlation Validity whose results were compared with the r-table at the 5% significance level. Thus the r-table for 15 students in the small group trial was 0.514.

A question is declared valid if the value of r count is greater than r table. Conversely, if r count is smaller than r table, then the question is considered invalid and must be eliminated. The results of the calculation of empirical validity for 50 items resulted in 33 valid questions and can be continued to the field tester (large-scale test), while as many as 17 items were declared invalid so they had to be eliminated or not used. Questions that are valid and can be used for large-scale testing are number 1, 4, 5, 7, 9, 10, 11, 12, 14, 16, 18, 20, 21, 22, 23, 26, 27, 29, 30, 31, 33, 34, 35, 36, 37, 38, 39, 41, 44, 45, 46, 48, and 49. Question number 30 has the highest coefficient of 0.87 and question number 6 has the lowest coefficient of 0.11. The science literacy component in the small-scale test includes valid questions with indicators of explaining phenomena scientifically as many as 17 questions, evaluating and designing scientific investigations as many as 5 questions and interpreting data and scientific evidence as many as 11 questions.

Reliability in this study was calculated using MC Excel. The results of the analysis of small group trials conducted by 15 students resulted in reliability in Roman I of 0.823476 getting a very high reliability conclusion. In Roman II getting a value of 0.75136068 getting a high reliability conclusion. In Roman III getting a value of 0.824046 getting a very high reliability conclusion. Roman IV gets a value of 0.902840406 which gets a very high reliability conclusion. In Roman V getting a value of 0.859332626 getting a very high reliability conclusion.

The analysis of the level of difficulty aims to find out how difficult a question is. There are three categories of difficulty levels, namely difficult with an index of 0.00-0.30; moderate with an index of 0.31-0.70; and easy with an index of 0.71-1.00. Suharsimi (2005: 207) "questions are considered good if they are included in the medium category, which has a difficulty index between 0.30 and 0.70. This category indicates that the question is at a level of difficulty that is not too complicated and not too simple". Based on a small-scale test involving 15 students, of the 50 items analyzed, the following results were obtained in the Roman I section, there were 2 difficult questions, 6 medium questions, and 7 easy questions. For Roman II, 2 medium questions were found, and 8 easy questions. In the Roman III section, there are 5 medium questions and 5 easy questions. In Roman IV, there are 2 medium questions and 5 easy

questions. While in Roman V, there are 4 easy questions, 3 medium questions, and 1 difficult question.

The results of the exception test, in roman I complex multiple choice questions there are exceptions whose values are below 5% and above 5%. Eligibility that is below 5% is found in question number 13 and there are 2 questions that have an exception that is above 5% In Roman II complex multiple choice questions there are eligibility that is below 5% and above 5%. Eligibility criteria that are below 5% are 4 questions and questions that have accepted and above 5% are 6 questions. In Roman III matching questions, it is known that there are 2 questions that have 1 active exception, 4 questions that have 2 active exceptions, 2 questions that have 3 active exceptions, and 2 questions that have 4 active exceptions. In the short form question in Roman IV and the description question in Roman V, the exemptions were not analyzed because the description question did not have any exemptions. An exception is an alternative answer that is not the correct answer, which serves to test participants' understanding by trapping those who do not master the material. (Sari, Utomo, & Sumarwati, 2022). Each exception must be designed in such a way that it is interesting enough to be chosen by students as an answer. While the short form and description are questions that ask students to answer with their thoughts, and no answer choices are used. Therefore, the short form and description are not analyzed for the exemptions.

Based on the results of the small-scale trial analysis, 33 test items were found to be valid, while 17 were invalid. The valid items were retained for further use, and the invalid ones were discarded. In Section I, nine items were deemed usable, covering the topic of energy transformation (items 1, 4, 5, 7, 9, 10, 11, and 12) and potential energy (item 14). In Section II, six items were usable, including energy transformation (items 16 and 18), kinetic energy (items 20 and 23), and potential energy (items 21 and 22). In Section III, eight items were valid, consisting of energy transformation (items 26, 27, 30, and 35), potential energy (items 29, 31, and 33), and kinetic energy (item 34). In Section IV, five items were usable, comprising potential energy (item 41) and kinetic energy (items 36, 37, 38, and 39). In Section V, five items were valid, including energy transformation (items 44 and 46), kinetic energy (item 45), and potential energy (items 48 and 49). The components of science literacy measured in the small-scale trial were: explaining phenomena scientifically (17 items), evaluating and designing scientific inquiry (5 items), and interpreting data and scientific evidence (11 items). These valid items were used in the subsequent large-scale trial phase.

3.2.2 Large Scale Test

Large-scale tests were conducted at SDN Pagu 1 and SDN Jabon 1 with 70 fourth-grade students as subjects. The test items were analyzed based on validity, reliability, difficulty level, discriminating power, and distractors.

The results of the large-scale test recapitulation contained 32 valid questions and 1 invalid question, invalid questions will not be used and discarded. Valid questions are at numbers 2, 3, 4, 5, 6, 7, 8, and 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, and 33. Question number 21 has the highest coefficient of 0.93 and question number 1 has the lowest coefficient of 0.09. The components of scientific literacy in the small-scale test include valid questions with indicators of explaining phenomena scientifically as many as 16 questions, evaluating and designing scientific investigations as many as 5 questions and interpreting data and scientific evidence as many as 11 questions.

Reliability is calculated using MC Excel. The results of the analysis of large group trials conducted by 70 students produced reliability in Roman I of 0.771776 getting a conclusion of high reliability, in Roman II getting a value of 0.6362 getting a conclusion of high reliability, in Roman III getting a value of 0.939953 getting a conclusion of very high reliability, in Roman IV getting a value of 0.662352 getting a conclusion of high reliability, and in Roman V getting a value of 0.79625344 getting a conclusion of high reliability. Heale & Twycross, (2015) stated that an acceptable reliability score is > 0.6 . So it can be said that the reliability of the large group test is acceptable.

Difficulty level analysis is used to determine the level of difficulty of a question. There are 3 types of difficulty level criteria, namely difficult with a difficulty index of 0.00-0.30, moderate with a difficulty index of 0.31 -0.70, and easy with a difficulty index of 0.71-1.00. The results of the analysis of questions in the large group test by 70 students in Roman I obtained 3 questions in the easy category, 4 questions in the medium category, and 2 questions in the difficult category. In Roman II, complex multiple choice questions obtained 3 easy questions and 3 medium questions. In Roman III, matching questions obtained 1 question in the medium category and 7 questions in the easy category. In Roman IV, short answer questions obtained 1 easy question, 2 medium questions, and 2 difficult questions. In Roman V, essay questions obtained 1 medium question and 4 difficult questions. The average level of difficulty in the large-scale test, namely 0.6546, is in the medium category. According to Purniasari, Masykuri, & Ariani (2021), the level of difficulty that is close to 0 will be increasingly difficult and the one that is close to 1 will be increasingly easy. According to Nana Sujana (2002), the range of difficulty levels that are considered good ranges from 0.25 to 0.75 (Purniasari et al., 2021). The conclusion of this scientific literacy test instrument is at a moderate level so it is categorized as good and suitable for use.

In the discriminatory power test, there are 5 criteria, namely very good with criteria 0.71-1.00, good with criteria 0.41-0.70, sufficient with criteria 0.21-0.40, poor with criteria 0.21-0.40, and very poor with negative criteria. The discriminatory power test was conducted using MS Excel. The results of the discriminatory power test from a large group test

by 70 students in Roman I produced 4 good questions, 4 sufficient questions and 1 poor question. In Roman II, complex multiple choice questions, 3 very good questions, 2 good questions, and 1 sufficient question were obtained. In Roman III, matching questions, 8 good questions were obtained. In Roman IV, short answer questions, 3 very good questions and 2 good questions were obtained. In Roman V, essay questions, 3 good questions and 2 very good questions were obtained. The average discriminatory power on science literacy test questions is 0.5571 in the good category. This is in line with Arikunto's opinion (2018) that discriminatory power functions to assess the extent to which a question item is able to distinguish between students with very high and very low abilities. A question is categorized as having good discriminatory power if its discriminatory power value is in the range of 0.41 to 0.70.

In the distractor analysis of Roman I multiple-choice questions, there were 7 questions with distractors of less than 5% and 2 questions with a discriminating power of more than 5%. There were 7 questions with a discriminating power of less than 5%. The discriminating power that was not good would be replaced and revised. There were 2 questions with a discriminating power of more than 5%. In Roman II complex multiple-choice questions, 38 questions all had acceptable distractors because all distractors had a percentage of more than 5%. There were 6 acceptable distractors. In Roman III matching questions, it was found that question number 4 had 4 active distractors, question number 21 had 5 active distractors, questions number 17, 19, 20, and 23 had 6 active distractors, and questions number 22 and 24 had 7 active distractors. In short answer questions in Roman IV and essay questions in Roman V, the distractors were not analyzed because the essay questions did not have distractors. Distractors are answer options other than the answer key that are designed to trick test takers who have not mastered the material being tested. (Sari, Utomo, & Sumarwati, 2022). Each distractor must have enough appeal for participants to consider it as a possible correct answer. Meanwhile, short answer questions and essay questions are questions that ask students to answer with their thoughts, and no answer choices are used. So in short answer questions and essay questions, the distractors are not analyzed.

3.3 Product Test Practicality Results

The data source of practicality in this study was obtained from practitioner questionnaires distributed to teachers and students from SDN Pagu 1, SDN Pagu 2, and SDN Jabon 1 after they completed and analyzed the local wisdom-based saian literacy test instrument in Kediri energy material grade IV elementary school. The results are presented in description form as follows.

3.3.1 Practicality Test Results from Teachers

The results of the teacher practicality test assessment were evaluated using small and large scale products based on the local wisdom-based saian literacy test instrument in Kediri energy material grade IV elementary school. The results of the product practicality analysis in the small scale test at SDN Pagu 2 obtained a percentage of 89.2857% and was categorized as very practical with the test worth using without revision. The comment from teacher I is that the instrument developed is good enough, but it needs to be considered again the differences in student characters. The results of the product practicality analysis in the largescale test at SDN Pagu 1, and Jabon 1 received a percentage of 86.90% and were categorized as very practical. The comment from teacher II is that the instrument is very good. The comment from the third teacher is that the format of the test instrument is actually easy to understand, it's just that it is constrained by the lack of student interest in reading. The comment from teacher IV is that everything is appropriate but the lack of student interest in reading is an obstacle.

In the development of this local wisdom-based science literacy test instrument, the practicality aspect of the instrument is able to provide insight to students about local wisdom. This statement is in line with the views expressed by Ikhsan & Rachmadyanti (2024) that integrating local wisdom can increase awareness and appreciation of cultural diversity, and strengthen national identity. This is also in accordance with the recommendation of the Ministry of Education and Culture's independent learning curriculum that learning should start from the environment that surrounds students.

In the aspect of reviewing the ability to be achieved, the local wisdom-based science literacy test instrument developed is appropriate for use at the elementary school level. This is in line with the findings of two other researchers who stated that the indicators in the scientific literacy evaluation instrument need to be adjusted to the level of ability and education level of students. (Miftahuzzakiah, 2018; Muslihah, 2021).

In the suitability and completeness aspects of the local wisdom-based science literacy test instrument, there are grids, instructions for use, answer keys, and scoring. A good test instrument certainly consists of a complete grid, instructions for use, answer key, and scoring (Kurniawan Heru, 2021). This is also in line with the views of other researchers who mention that the grid acts as a guide in compiling items and assembling them into a test device (Putra, Adzim, & Hilmiyati, 2025). The results of testing the practicality by teachers of the science literacy test instrument based on local wisdom in Kediri energy material for grade IV elementary school show that the instrument is practical and feasible to use.

3.3.1 Practicality Test Results from Students

The results of the student practicality test in the small group test at SDN Pagu 2 received a percentage of 78.5%, which is in the practical category. In the aspect of interest, the statement gets a percentage of 80.416%, in the material aspect it gets a percentage of 76, 2499%, and in language it gets a percentage of 79.1666%. The results of the student practicality test in the large group test at SDN Pagu 1 and SDN Jabon 1 received a percentage of 79.6% in the practical category. In the interest aspect, the percentage was 78.550%, in the material aspect, the percentage was 79.235%, and in the language aspect, the percentage was 79.81%.

The student practicality test has 3 aspects, namely aspects of interest, material, and language. The interest aspect consists of the science literacy test instrument making students more enthusiastic about learning, not boring, clear instructions, and making students excited and challenged to learn. The interest aspect gets a score of 80.416%, and 78.5595%. In line with research conducted by Apriyulianti, Istiningsih, Rahmatih, & Astria (2023) that incorporating local wisdom in science learning can present an innovative approach through integrating local cultural values into the science curriculum.

The material aspect consists of science filtering test questions in accordance with the material that has been learned by students, questions are easy to understand, easy to do and the time taken is in accordance with the number of questions available. The material aspect in the small test scored 76.249908325 and in the large test scored 79.235%. In line with the research of Afrian, Islami, & Mustika (2018) who said that the preparation of questions must refer to the predetermined indicators and follow the appropriate question form.

The language aspect consists of sentences that are clear and easy to understand, questions using standard Indonesian and young to understand. In the language aspect, the scores are 79.16665 and 79.81. This is in line with the findings of Syahriandi (2017) which states that every test needs to be prepared using correct Indonesian and in accordance with applicable language rules.

4. Conclusion

Based on the evaluation of the science literacy test instrument product—developed using local wisdom in Kediri and focused on energy material for fourth-grade elementary students—the content validation results were highly satisfactory. Validation from subject matter experts reached 100% (very valid), and validation from evaluation experts reached 94.44% (very valid), indicating that the test instrument is suitable for use without the need for revision. In the small-scale trial, 33 items were categorized as valid, while 17 items were found to be invalid. Six items demonstrated high reliability, and 27 items showed very high reliability. The average item difficulty index was 0.718, which falls under the easy category. The discrimination index was 0.418, classified as good. Practicality, as perceived by teachers, was 89.29%, while students rated it at 78.5%. In the large-scale trial, 32 items were deemed valid and 1 item invalid. Eight items showed very high reliability, and 25 items showed high reliability. The difficulty index was 0.6546, categorized as moderate. The discrimination index was 0.5571, also falling under the good category. Practicality ratings were 79.6% (practical) from students and 86.90% (very practical) from teachers.

Based on the results of the empirical validation test, the test instrument got a very valid category, the results of the logical validity test got a valid category, and the level of practicality got a practical percentage. Therefore, this instrument can be used so that students become accustomed and trained in working on questions based on scientific literacy and can also introduce local wisdom in Kediri. Suggestions for further research are to develop a scientific literacy test instrument for science material, with educational levels, and other local wisdom.

References

- Ainissyifa, H., Nasrullah, Y. M., Fatonah, N., Indriani, S. A., Asyifiya, S. N., & Rohmah, A. (2024). *Manajemen Pendidikan Dalam Kurikulum Merdeka Di Madrasah*. Cahaya Smart Nusantara.
- Afriyanti, Yulia., Sasana, H., & Jalunggono. (2018). *Analisis Faktor-faktor Yang Mempengaruhi Konsumsi Energi Terbarukan di Indonesia*. DINAMIC: Directory Journal of Economic Volume 2 Nomor 3.
- Afrian, R., Islami, Z. R., & Mustika, F. (2018). *Pembinaan Pembuatan Tes Buatan Guru (Soal) Mata Pelajaran Geografi SMA/MA Kota Langsa*. Jurnal Vokasi - Politeknik Negeri Lhokseumawe, 1(2). <https://doi.org/10.30811/vokasi.v1i2.687>
- Ahmad, Susanto. (2013). *Teori Belajar dan Pembelajaran di Sekolah Dasar*. Jakarta: Kencana Prenada Media Group.
- Akbar, S. (2013). *Instrumen Perangkat Pembelajaran*. Remaja Rosdakarya.
- Akbar, Sa'dun. (2017). *Instrumen Perangkat Pembelajaran*. Bandung: PT Remaja Rosdakarya.
- Ali, N. B. V., Setiawan, I. M. B., Joko, B. S., Ulumuddin, I., & Julizar, K. (2019). *Evaluasi program gerakan literasi sekolah (GLS)*. RISTEK: Jurnal Riset, inovasi, dan Teknologi (Vol. 3).

- Alwi, I. (2015). *Kriteria Empirik Dalam Menentukan Ukuran Sampel*. Jurnal Formatif, 2(2), 140–148.
- Anggraisa, A., Nurlidiya, E., Sativa, O., Kholiza, T., & Putri, N. (2024). *Mengintegrasikan Kearifan Lokal Dalam Kurikulum Pendidikan*. I-WIN Library.
- Anita, A., Tyowati, S., & Zulfadrial, Z. (2018). *Analisis Kualitas Butir Soal Fisika Kelas X Sekolah Menengah Atas*. Edukasi: Jurnal Pendidikan, 16(1), 35. <https://doi.org/10.31571/edukasi.v16i1.780>
- Apriyulianti, S., Istiningsih, S., Rahmatih, A. N., & Astria, F. P. (2023). *Pengembangan Media Scrapbook Berbasis Kearifan Lokal Suku Sasak pada Mata Pelajaran IPA di Sekolah Dasar*. Jurnal Pendidikan MIPA, Volume 13.(Desember 2023), 723–731.
- Astuti, W. P., Prasetyo, A. P. B., & Rahayu, E. S. (2014). *Pengembangan Instrumen Asesmen Autentik Berbasis Literasi Sains Pada Materi Sistem Ekskresi*. Journal UNNES, 43(2), 94–102.
- CHAIRUL, A. (2019). *Kearifan Lokal Dalam Tradisi Mancoliak Anak Pada Masyarakat Adat Silungkang*. Jurnal Penelitian Sejarah Dan Budaya, 5(2), 172–188. <https://doi.org/10.36424/jpsb.v5i2.86>
- Dewantari, N., & Singgih, S. (2020). *Penerapan Literasi Sains Dalam Pembelajaran Ipa*. Indonesian Journal of Natural Science Education (IJNSE), 3(2), 366–371. <https://doi.org/10.31002/nse.v3i2.1085>
- Dewi, C.C.A., Erna, M., Haris, I., & Kundera, I.N. (2021). *The effect of contextual collaborative learning based ethnoscience to increase student's scientific literacy ability*. Journal of Turkish Science Education, 18(3):525-541. <https://orcid.org/0000-0001-9381-9645>.
- Fahrianur, Monica, R., Wawan, K., Misnawati, Nurachmana, A., Veniaty, S., & Ramadhan, I. Y. (2023). *Implementasi Literasi di Sekolah Dasar*. Journal of Student Research (JSR), 1(1), 102–113.
- Fananta, M. R., Widjiasih, A. E., Setiawan, R., Hanifah, N., Miftahussururi, Nento, M. N., ... Ayomi, J. M. (2017a). *Materi Pendukung Literasi Sains*. Repositori Kemdikbud. <https://Repositori.Kemdikbud.Go.Id/11631/1/Cover-Materi-Pendukung-Literasi-Sains-Gabung.Pdf>, 1–23.
- Fananta, M. R., Widjiasih, A. E., Setiawan, R., Hanifah, N., Miftahussururi, Nento, M. N., ... Ayomi, J. M. (2017b). *Gerakan Literasi Nasional*, 1–36.
- Farah, A., Hadi, W., & Soro, S. 2020. *Evaluasi Proses dan Hasil Belajar Matematika dengan Diskusi dan Simulasi*. Jakarta: Uhamka Press.
- Firdaus, M., & Asmali. (2021). *Pengembangan Instrumen Penilaian Berbasis Literasi Numerasi*. Direktorat Sekolah Menengah Atas, Direktorat Jenderal Pendidikan Anak Usia Dini, Pendidikan Dasar, Dan Pendidikan Menengah - Kementerian Pendidikan, Kebudayaan, Riset, Dan Teknologi.
- Fitri, A., Rasa, A. A., Kusumawardhani, A., Nursya'bani, K. K., Fatimah, K., & Setianingsih, N. I. (2021). *Buku Panduan Guru Ilmu Pengetahuan Alam dan Sosial*.
- Fitri, A., Rasa, A. A., Kusumawardhani, A., Nursya'bani, K. K., Fatimah, K., & Setianingsih, N. I. (2023). *Buku Siswa Ilmu Pengetahuan Alam dan Sosial*.
- Heale, R., & Twycross, A. (2015). *Validity and reliability in quantitative studies*. Evidence-Based Nursing, 18(3), 66–67. <https://doi.org/10.1136/eb-2015-102129>.
- Herawati, V. R. (2022). *Ritual Larung Sesaji Di Kawah Gunung Kelud Sarana Pengembangan Pariwisata di Kab. Kediri*.
- I'tishambillah, M. Q. (2023). *Pengembangan Instrumen Tes Untuk Mengukur Kemampuan Literasi Statistik Siswa SMA/MA*. Repository UIN SYarif Hidayatullah.
- Ida, I. M., Pritasari, A. C., & Hakim, M. L. (2013). *Analisis Muatan IPA Kelas 5 berdasarkan Aspek Literasi Sains dan Integrasi terhadap Potensi Madura*. Journal of Chemical Information and Modeling, 53(9), 1689–1699.
- Ikhsan, A. H. N., & Rachmadyanti, P. (2024). *Implementasi Model Inkuiri Berbasis Kearifan Lokal Terhadap Hasil Belajar IPA Kelas IV Sekolah Dasar*. Taksonomi: Jurnal Penelitian Pendidikan Dasar, Volume 4 N, 389–398.
- Irsan. (2020). *Implementasi Literasi Sains Dalam Pembelajaran IPA di Sekolah Dasar*. Jurnal Basicedu, 5(5), 3(2), 524–532. Retrieved from <https://journal.uui.ac.id/ajie/article/view/971>
- Janah, F. R., & Hidayati, S. N. (2020). *Analisis Pemahaman Konsep IPA Siswa SMP di Surabaya*. Jurnal Pendidikan MIPA, 14(September), 723–731.

- Kafi, F. A., & Ilma, Z. (2022). *Analisis Soal Tematik Bahasa Arab berbasis HOTS pada Kelas XII MAN 3 Jember*. Al-Fusha : Arabic Language Education Journal, 4(1), 1–9. <https://doi.org/10.62097/alfusha.v4i1.729>
- Kemendikbud, S. G. L. S. (2019). *Desain Induk Gerakan Literasi Sekolah (Direktorat Jenderal Pendidikan Dasar dan Menengah Kementerian Pendidikan dan Kebudayaan)*. Retrieved from <https://gln.kemdikbud.go.id/glnsite/wp-content/uploads/2019/07/Desain-Induk-Gerakan-LiterasiSekolah-2019.pdf>
- Kemendikbudristek. (2023). *Pisa 2022 Dan Pemulihan Pembelajaran Di Indonesia*. Laporan Pisa Kemendikbudristek, 1–25.
- Kemendikbud. (2021). *Asesmen Nasional: Lembar Tanya Jawab. Kementerian Pendidikan Dan Kebudayaan*, 1–32. Retrieved from https://hasilun.puspendik.kemdikbud.go.id/akm/file_akm_202101_1.pdf
- Kirana, R. N., Surayanah, & Mas'ula, S. (2024). *Development of a Mathematical Literacy Test Instrument Based on Local Wisdom in Blitar Raya for Fifth-Grade Elementary Students*. Journal of Development Research, 8(2), 110–124.
- Khoiriya, R. M., Mudiono, A., Barus, Y. K., Sari, N. N., Attila, Z., & Hidayati, R. P. (2024). *Developing a Multiliteracy Learning Model Based on Malangan Masks to Enhance Cultural Literacy and Global Diversity in Elementary Students*. At-Thullab : Journal of Madrasah Ibtidaiyah Teacher Education, 8(2), 1–23.
- Kusumaningsih, N. A. R. (2019). *Analisis Perubahan Konsepsi Siswa Melalui Penerapan Model Transformative Learning Pada Materi Perubahan Lingkungan Universitas Pendidikan Indonesia* | repository.upi.edu | perpustakaan.upi.edu. UPI Repository, 1–7.
- Lestari, R., Surayanah, & Bintartik, L. (2024a). *Pengembangan Instrumen Tes IPA Berbasis Kearifan Lokal Blitar pada Siswa Sekolah Dasar*. Proceedings Series of Educational Studies, (1).
- Lestari, R., Surayanah, & Bintartik, L. (2024b). *Pengembangan Instrumen Tes IPA Berbasis Kearifan Lokal Blitar pada Siswa Sekolah Dasar*. Proceedings Series of Educational Studies Seminar Nasional Dan Prosiding PGSD FIP UM 2024, (1).
- Lumbantobing, L. R. (2021). *Kelebihan dan Keterbatasan dari Bentuk-Bentuk Penilaian Tes Tertulis*. Kemenkeu Learning Center, 1–2. Retrieved from <https://klc2.kemenkeu.go.id/kms/knowledge/kelebihanketerbatasan-penilaian-tes-tertulis-0f81178c/detail/>
- Magdalena, I., Syariah, E. N., Mahromiyati, M., & Nurkamilah, S. (2021). *Analisis Instrumen Tes Sebagai Alat Evaluasi Pada Mata Pelajaran Sbdp Siswa Kelas Ii Sdn Duri Kosambi 06 Pagi*. Jurnal Pendidikan Dan Ilmu Sosial, 3(2), 276–287. Retrieved from <https://ejournal.stitpn.ac.id/index.php/nusantara>
- Maya Ningsetyo, & Titin Sunarti. (2024). *Pengembangan Instrumen Tes Literasi Sains Berbasis Kearifan Lokal Di Probolinggo*. Jurnal Arjuna : Publikasi Ilmu Pendidikan, Bahasa Dan Matematika, 2(2), 59–68. <https://doi.org/10.61132/arjuna.v2i2.612>
- Mujiono, & Dharsono. (2015). *Keberadaan Batik Kediri Jawa Timur*. Gelar: Jurnal Seni Budaya, 13(1), 1–12. Retrieved from <https://jurnal.isi-ska.ac.id/index.php/gelar/article/view/1535/1485>
- Murtadlo, M. K. A., Oktafiani, D. A., & Faizah, H. (2024). *Evaluasi Efektivitas Kurikulum Merdeka Sebagai Langkah Kritis Menuju Implementasi yang Optimal*. Edukatif : Jurnal Ilmu Pendidikan, 6(5), 5564–5571. <https://doi.org/10.31004/edukatif.v6i5.7052>
- Murti, W. W., & Sunarti, T. (2021). *Pengembangan Instrumen Tes Literasi Sains Berbasis Kearifan Lokal Di Trenggalek*. ORBITA: Jurnal Kajian, Inovasi Dan Aplikasi Pendidikan Fisika, 7(1), 33. <https://doi.org/10.31764/orbita.v7i1.4386>
- Mustofa, A., Kuswanti, N., Siti, & Hidayati, N. (2017). *Keefektifan Lks Berbasis Model Pembelajaran Discovery Learning Untuk Meningkatkan Kemampuan Literasi Sains*. E-Jurnal Pensa, 5(1), 27–32.
- Muthrofin, M. A. F., Erwanto, D., & Yanuartanti, I. (2024). *Convolutional Neural Network untuk Klasifikasi Batik Tenun Ikat Bandar Berdasarkan Fitur Warna dan Tekstur*. Journal of Electrical Engineering and Computer (JEECOM) Dengan, Vol. 6, No(1), 120–128. <https://doi.org/10.33650/jeeecom.v6i1.8060>
- Nahak, R. L. (2024). *Pengembangan Asesmen Diagnostik Kognitif Berbasis Kearifan Lokal Masyarakat Ende Lio*. Edukatif : Jurnal Ilmu Pendidikan, 6(5), 5825–5835. <https://doi.org/10.31004/edukatif.v6i5.7562>

- Natalia, N., Dewi, I. C., & Pemasaran, S. (2024). *Strategi Pemasaran Grthuk Pisang Sebagai Oleh-Oleh Khas Kediri*. Jurnal Review Pendidikan Dan Pengajaran, Volume 7 N, 4838–4849. Retrieved from <http://journal.universitaspahlawan.ac.id/index.php/jrpp>
- Nurdiana, D., Aprijani, D. A., Amastini, F., Maulana, M. R., & Utama, M. R. P. A. (2024). *Pengembangan Sistem Informasi Pengelolaan Pembimbing Lapangan Praktik Kerja Lapangan (PKL) Prodi Sistem Informasi Universitas Terbuka*. Decode: Jurnal Pendidikan Teknologi Informasi, 4(2), 418–436. <https://doi.org/10.51454/decode.v4i2.433>
- Nurbaeti, N., Mayasari, A., & Arifudin, O. (2022). *Penerapan Metode Bercerita Dalam Meningkatkan Literasi Anak Terhadap Mata Pelajaran Bahasa Indonesia*. Jurnal Tahsinia, 3(2), 98–106. <https://doi.org/10.57171/jt.v3i2.328>
- OECD. (2019). *PISA 2018 Assessment and Analytical Framework*. OECD Publishing.
- OECD. (2023b). *Program For International Student (PISA) 2022 Assessment and Analytical Framework*. Paris: OECD Publishing. <https://doi.org/10.1787/dfe0bf9c-en>.
- OECD. (2023). *PISA 2022 Results Factsheets Indonesia*. OECD (Organisation for Economic Co-Operation and Development) Publication, 1–9. Retrieved from https://www.oecd.org/en/publications/pisa-2022results-volume-i-and-ii-country-notes_ed6fbcc5-en/indonesia_c2e1ae0e-en.html
- OECD. (2022). *Programme For International Student Assesment (PISA) Results From PISA 2022*.
- PANGASTUTI, R. L., & Sustiyatik, E. (2018). *Ibm Usaha Mikro Pembuatan Tahu Takwa Kuning Dan Stik Tahu Desa Tinalan Kecamatan Pesantren Kota Kediri*. Jurnal Abdi Masyarakat, 1(1). <https://doi.org/10.30737/jaim.v1i1.86>
- Pascasarjana Undiksha. (2020). *Taraf kesukaran, daya beda, dan efektivitas pengecoh*. Universitas Pendidikan Ganesha, 81–87.
- Pertiwi, U. D., Atanti, R. D., & Ismawati, R. (2018). *Pentingnya Literasi Sains Pada Pembelajaran Ipa Smp Abad 21*. Indonesian Journal of Natural Science Education (IJNSE), 1(1), 24–29. <https://doi.org/10.31002/nse.v1i1.173>
- Pramitasari, A. (2016). *Pengembangan Materi Ajar BIPA Bermuatan Kearifan Budaya Lokal Kota Pekalongan*. Proceeding, 1(1), 49–53.
- Prihandini, R. D., & Artono. (2021). *Eksistensi industri tahu kuning di Kediri tahun 1990-2020*. AVATARA: Journal Pendidikan Sejarah, 10(3), 1–9.
- Rahman, A. (2024). *Asesmen Kemampuan Literasi Sains Siswa Pada Mata Pelajaran IPA di SMP Kota Bandung*. Repository.Upi.Edu |, 4(1), 1–23.
- Risma, M., Rahmayani, R., & Handayani, F. (2019). *Analisis Konten Buku Teks IPA Terpadu Kelas VIII Semester 1 Ditinjau Dari Aspek Literasi Sainifik*. Jurnal Eksakta Pendidikan (Jep), 3(2), 200. <https://doi.org/10.24036/jep/vol3-iss2/396>
- Rohim, D. C., Rahmawati, S., & Ganestri, I. D. (2021). *Konsep Asesmen Kompetensi Minimum untuk Meningkatkan Kemampuan Literasi Numerasi Siswa Sekolah Dasar*. Jurnal VARIDIKA, 33(1), 54–62. <https://doi.org/10.23917/varidika.v33i1.14993>
- Safitri, A. N., Sarwanto, S., & Harjunowibowo, D. (2023). *Pengembangan Modul Pembelajaran Fisika Berbasis Kearifan Lokal Pada Materi Suhu dan Kalor*. Jurnal Materi Dan Pembelajaran Fisika, 13(1), 32. <https://doi.org/10.20961/jmpf.v13i1.60093>
- Sa'idah, N., Yulistianti, H. D., & Megawati, E. (2019). *Analisis Instrumen Tes Higher Order Thinking Matematika Smp*. Jurnal Pendidikan Matematika, 13(1), 41–54. <https://doi.org/10.22342/jpm.13.1.6619.41-54>
- Saiful, A., Rakhma, A. E., & Muhamad, J. (2019). *Peningkatan Produktivitas Industri Rumah Tangga Getuk Pisang Kediri Melalui Pemanfaatan Teknologi Pencetak Getuk Pisang Otomatis*. Prosiding Seminar Nasional Teknologi Dan Sains (SNasTekS), (September), 97–103. Retrieved from <https://www.journal.unusida.ac.id/index.php/snts/article/view/80>
- Sanusi, R. N. A., & Aziez, F. (2021). *Analisis Butir Soal Tes Objektif dan Subjektif untuk Keterampilan Membaca Pemahaman pada Kelas VII SMP N 3 Kalibagor*. Metafora: Jurnal Pembelajaran Bahasa Dan Sastra, 8(1), 99. <https://doi.org/10.30595/mtf.v8i1.8501>
- Sari, H. A., & Yuningsih, U. W. (2020). *Perkembangan Ragam Hias Dan Warna Tenun Ikat Bandar Kidul Kota Kediri Jawa Timur*. E-Jurnal, Volume 09(75), 147–154. Retrieved from

- <https://doi.org/10.1016/j.jnc.2020.125798><https://doi.org/10.1016/j.smr.2020.02.002><http://www.ncbi.nlm.nih.gov/pubmed/810049><http://doi.wiley.com/10.1002/anie.197505391><http://www.sciencedirect.com/science/article/pii/B9780857090409500205>
- Sari, V. N. I., Utomo, A. P. Y., & Sumarwati. (2022). *Kualitas Soal Bahasa Indonesia di SMP Muhammadiyah 1 Pontianak: Analisis Butir Soal*. *Jurnal Pendidikan Bahasa Dan Sastra Indonesia*, 11(2), 112–119. Retrieved from <https://journal.unnes.ac.id/sju/index.php/jpbsi/article/view/24018>
- Setiawan, A. A., & Sudigdo, A. (2019). *Penguatan Literasi Siswa Sekolah Dasar Melalui Kunjungan Perpustakaan. Prosiding Seminar Nasional PGSD*, 2015, 24–30. Retrieved from <https://core.ac.uk/download/pdf/230386992.pdf>
- Slamet, R., & Wahyuningsih, S. (2022). *Validitas Dan Reliabilitas Terhadap Instrumen Kepuasan Kerja*. *Aliansi : Jurnal Manajemen Dan Bisnis*, 17(2), 51–58. <https://doi.org/10.46975/aliansi.v17i2.428> Sugiono. (2023). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Suhelayanti, Z, S., & Rahmawati, I. (2023). *Pembelajaran Ilmu Pengetahuan Alam Sosial (IPAS)*. Penerbit Yayasan Kita Menulis.
- Sunaryati, T., Azzahra, S. S., Khasanah, F. N., Dewi, N., & Komariyah, S. (2024). *Analisis Instrumen Test Sebagai Alat Evaluasi pada Pembelajaran di Sekolah Dasar*. *Jurnal Riset Pendidikan Dasar*, XX, 316–324.
- Susanto, A. (2013). *Teori Belajar & Pembelajaran di Sekolah Dasar*.
- Sutikno, Y., Hosan, H., & Irawati, I. (2021). *Implementasi Metode Penugasan untuk Meningkatkan Kemampuan Melakukan Wawancara untuk Mahasiswa STAB Maitreyawira*. *Jurnal Maitreyawira*, 2(2), 1–8. <https://doi.org/10.69607/jm.v2i2.40>
- Syahriandi. (2017). *Kualitas Butir Soal Ujian Ditinjau Dari Segi Bahasa (Analisis Kualitatif Butir Soal)*. *Visipena Journal*, 8(1), 174–183. <https://doi.org/10.46244/visipena.v8i1.386>
- Tillah, N. F., & Subekti, H. (2025). *Analisis Kemampuan Literasi Sains Siswa SMP Berdasarkan Indikator dan Level Literasi Sains*. *Edusaintek: Jurnal Pendidikan, Sains Dan Teknologi*, 12(1), 137–154.
- Ulva, T. S., Samitra, D., & Kusnanto, R. A. B. (2021). *Pengembangan Instrumen Tes Literasi Sains pada Tema 8 Lingkungan Sahabat Kita untuk Siswa Sekolah Dasar*. *Jurnal Basicedu*, 5(5), 3781. Retrieved from <https://journal.uui.ac.id/ajie/article/view/971>
- Umami, R., Rusdi, M., & Kamid, K. (2021). *Pengembangan Instrumen Tes untuk Mengukur Higher Order Thinking Skills (HOTS) Berorientasi Programme For International Student Assessment (PISA) pada Peserta Didik*. *JP3M (Jurnal Penelitian Pendidikan Dan Pengajaran Matematika)*, 7(1), 57–68. <https://doi.org/10.37058/jp3m.v7i1.2069>
- Umi Roufatuz Zahro'. (2020). *Pengembangan Instrumen Tes Untuk Mengukur Kemampuan Literasi Sains Siswa Smp Pada Tema Pemanasan Global*. *Lib.Unnes.Ac.Id*, 04(02), 771–778.
- Umi Roufatuz Zahro. (2020). *Pengembangan Instrumen Tes Untuk Mengukur Kemampuan Literasi Sains Siswa SMP Pada Tema Pemanasan Global*. *Lib.Unnes.Ac.Id*, 04(02), 771–778.
- Undang-Undang NRI No. 3 Tahun 2017 tentang Sistem Perbukuan (UU Sisbuk)
- Warju, W., Ariyanto, S. R., Soeryanto, S., & Trisna, R. A. (2020). *Analisis Kualitas Butir Soal Tipe Hots Pada Kompetensi Sistem Rem Di Sekolah Menengah Kejuruan*. *Jurnal Pendidikan Teknologi Dan Kejuruan*, 17(1), 95. <https://doi.org/10.23887/jptk-undiksha.v17i1.22914>
- Yulyantari, L. M. (2017). *Penilaian Essai Menggunakan Rubrik Penilaian*. *Jurnal Sistem Dan Informatika*, 12(2), 97–103.
- Yusrizal, & Rahmati. 2020. *Tes Hail Belajar*. Aceh: Bandar Publishing
- Widana, I. W. (2017). *Modul Penyusunan Soal Higher Order Thinking Skill (HOTS)*. Jakarta: Departemen Pendidikan dan Kebudayaan.