

## Analysis of Even Semester Final Assessment Questions in Thematic Learning of Grade III Elementary Schools

Nelly Wedyawati<sup>1</sup>, WiputraCendana<sup>2</sup>

<sup>1</sup>345STKIP Persada Khatulistiwa Sintang, Indonesia, [nellywedyawati@gmail.com](mailto:nellywedyawati@gmail.com)

<sup>2</sup>Pelita Harapan University Lippo Village, Indonesia, [wiputra.cendana@uph.edu](mailto:wiputra.cendana@uph.edu)

**Abstract:** This study aims to find out the quality of the questions in the even Final Semester Assessment (FSA) of grade III Public Elementary School 01 Sepauk, West Kalimantan. Descriptive quantitative research model type document analysis. The population is 25 people, the sampling technique is saturated with 25 people. Documentation data collecting technique and question document data collecting questions. Analysis using the Anates V4 program. (1) Validity: theme 5 (20 valid, 15 not), theme 6 (12 valid, 23 not). (2) Reliability: theme 5 (PG 0.86, Description 0.78), theme 6 (PG 0.11, Description -0.43). (3) Difficulty level: theme 5 (1 Difficult, 8 Medium, 26 easy), theme 6 (3 difficult, 9 medium, 23 easy). (4) Distinction: theme 5 (13 is very good, 5 is good, 17 is bad), theme 6 (5 is very good, 5 is enough, 25 is bad). (5) Quality of the questioners: Theme 5 (9 is very good, 11 is good, 10 is bad), Theme 6: (8 is very good, 11 is good, 11 very bad). (6) The cognitive level distribution is dominated by C<sup>1</sup>. The conclusion is that the quality of the questions in the category is not good and needs to be revised.

### 1. Introduction

Education is a series of the entire learning process in which there is a learning and learning activity carried out by students and teachers with the aim of acquiring knowledge<sup>(1)</sup>. According to <sup>(2)</sup>National education functions to develop capabilities and shape the character and civilization of the nation and with dignity in order to educate the nation's life. According to <sup>(3)</sup>Elementary School is a level of basic education that manages the education of children aged around 7 to 11 years. Learning is an activity that is carried out consciously, therefore learning must have a purpose. The purpose of learning is to help students to gain various experiences so that this experience can make students better and improve learning outcomes for the better. In terms of the process, each learning consists of three stages, namely the planning stage, the implementation stage and the evaluation stage. Evaluation is an important part of the teaching and learning process, because evaluation can determine the success of learning. Evaluation is used to measure and assess the extent to which the learning objectives that have been set can be achieved. According to <sup>(4)</sup>Evaluation is an assessment process by comparing the learning objectives to be achieved with the results of the learning process that have been achieved by students.

The test is an evaluation tool in learning that has an important role in measuring student achievement. Test questions need to pay attention to validity, reliability, level of difficulty, distinguishing power and distracting quality. In accordance with the advancement of the world of education, evaluation tools must be appropriate and in accordance with the curriculum. Assessment is an important part of the series of learning processes in education, it is because the assessment of learning outcomes can determine whether or not education is good so that the accuracy of learning outcomes assessment has a very significant impact on efforts to improve the quality of education in schools. According to <sup>(5)</sup>Learning outcomes as a measurement of the assessment of learning activities or the learning process are expressed in symbols, letters or sentences that tell the results that each child has achieved in a certain period.

One part of the test is a question that contains questions that will be answered by students. According to <sup>(6)</sup>Learning outcomes can be measured by learning outcomes tests. The learning outcome test is intended to measure students' mastery of certain material or topics as a result of the learning process, so that it can be seen the success of achieving these learning objectives. According to <sup>(7)</sup>explains that questions analysis or questions analysis is the assessment of test questions in order to obtain a set of questions that have adequate quality. Questions analysis aims to identify good, poorly and bad questions. Problems that are not good (bad) should not be used anymore so that the test is truly copied from quality questions to measure student learning outcomes.

Based on the results of interviews with grade III teachers of Public Elementary School 01 Sepauk, it was explained that the evaluation that had been carried out had not paid attention to the assessment of the questions, so that the quality of the questions that met the requirements was a good measuring tool or not. Knowing whether a question is good is still in a simple way, namely the number of students who can answer or cannot answer the question. Even though an analysis of these questions needs to be done to obtain information about the quality of the test so that the quality of the test can be improved in the future. Based on the above, Public Elementary School 01Sepauk has not taken steps to develop questions according to standards. End of semester assessment questions

which given to students were not previously analyzed so it was not known how the quality of the questions was. Therefore, the final semester assessment questions need to be validated, if the 5 questions are invalid then they cannot measure students' abilities accurately. Analysis of questions for thematic learning needs to be done to determine the quality of the questions and the success of elementary school students in mastering the subject materials.

Aspects that must be followed in the question's analysis are validity, reliability, difficulty level, distinguishing power and distracting quality. The quality of the questions will help the teacher to get evaluation results that are in accordance with the real situation of the students. The results of the evaluation will provide the teacher with information about student feedback on learning and student learning progress. Seeing the importance of evaluation, educators are required to conduct questions analysis to determine the quality level of each question used. Based on the description above, it is necessary to conduct research on "Analysis of the Questions for Final Assessment of Even Semester themes 5 and 6 on Thematic Learning in Grade III of Public Elementary School 01 Sepauk".

## 2. Methods

The approach used is a quantitative approach. According to <sup>(8)</sup>, the quantitative approach is research based on phenomena or symptoms that have a cause and effect obtained from a process of collecting data from certain populations and samples. The method used in this research is quantitative research. According to <sup>(9)</sup> quantitative analysis is an analysis that is carried out by testing instruments or test questions that have been analyzed qualitatively on several students who have the same characteristics as the students who will be tested using these test questions. The form of research used is descriptive quantitative. Quantitative descriptive means that the data analyzed does not accept and reject the hypothesis, but the results of the analysis are descriptions of the observed symptoms. The population was all the grade III students of Public Elementary School 01 Sepaukin Sepauk, Sintang, with 25 students. According to <sup>(10)</sup> saturated sampling is a sampling technique when all members of the population are used as samples. So, in this study the sample was 25 students of grade III of Public Elementary School 01 Sepauk who were subjected to FSA questions on themes 5 and 6.

The technique used is the documentation technique. The data collection tool is documentation. The document sheets used are in the form of school documents such as syllabus, lesson plans, final even semester assessment questions and student answer sheets. The data analysis technique in this study used descriptive statistics. The data to be analyzed and the steps in this study are as follows:

1. Analyze the test questions and determine the reliability, level of difficulty, distinguishing and distracting power using the ANATES Version 4 questions analysis program and determine validity using the excel program.

- a. Validity

The following results of the validity analysis in this study can be seen through the results of point biserial using excel. The results of point biserial with  $r_{table}$  with a significant level of 5% adjusted for the number of students, namely 0,396. If the biserial point is greater than 0,369 then the question is valid.

- b. Reliability

**Table 1. Reliability Criteria**

Correlation Coefficient	Qualification
0.91 - 1.00	Very high
0.71 - 0.90	High
0.41 - 0.70	Enough
0.21 - 0.40	Low
Negative-0.20	Very low

- c. Level of Difficulty

**Table 2. Difficulty Index**

No.	Difficulty Index	Category
1.	0.00 - 0.30	Hard
2.	0.31 - 0.70	Moderate
3.	0.71 - 1.00	Easy

d. Discernment

**Table 3. Discriminatory Power Criteria**

No.	DP Range	Category	Decision
1	40.00 - 100.00	Very well	Be accepted
2	30.00 - 39.00	Good	Accepted and revised
3	20.00 - 29.00	Enough	Revised
4	Negative - 0.19	Ugly	Rejected/thrown away

e. Tricking Quality

The analysis of the tricking answers is carried out on the questions in the form of multiple choice. Tricking's function well if the alternative answers are selected by at least 5% of all test takers <sup>(11)</sup>The quality criteria for tricking in multiple choice form test questions are described by the symbols contained in the Anates version 4 program, which are as follows:

- |                    |                   |
|--------------------|-------------------|
| a. ** : Answer key | d. - : Not good   |
| b. ++: Very good   | e. -: Bad         |
| c. +: Fine         | f. --- : Very bad |

The quality of trickings on each questions, the researcher used criteria adapted from a Likert scale.

**Table 4. The criterion for assessing the quality of the fraud**

Trickers that works	Criteria
4	Very good
3	Good
2	Enough / moderate
1	Not good / bad
0	Very Bad / Very Bad

f. Questions Quality.

The criteria used to interpret the quality of the questions are adapted from *Likert scale* The results of <sup>(12)</sup>Werdiningsih's researchentitled "Analysis of the Quality of Odd Semester Final Test Questions in Economic Grade XII Social Grade of Public Junior High School 2 Banguntapan Academic Year 2014/2015" as follows.

**Table 5. Classification of Questions Criteria**

Number of Criteria Fulfilled	Questions Quality
4	Very good
3	Good
2	Enough / moderate
1	Not good / bad
0	Very Bad / Very Bad

2. After the test questions are analyzed, the next step is to analyze each questions of the 6 criteria that are fulfilled in each questions namely the criteria for remembering (C<sup>1</sup>), understanding (C<sup>2</sup>), applying (C<sup>3</sup>), analyzing (C<sup>4</sup>), evaluating (C<sup>5</sup>), creating (C<sup>6</sup>). After that the data is calculated according to the percentage of the question criteria.

### 3. Result

The final assessment questions of even semester on theme 5 and theme 6 consist of 60 multiple choice questions and 10 essay questions. The purpose of this study was to determine the quality of the questions in the Final Assessment of Even Semester themes 5 and 6 in thematic learning grade III Public Elementary School 01 Sepauk. The quality of the questions analyzed included validity, reliability, level of difficulty, distinguishing power, quality of trickings and analyzing the distribution of cognitive levels in the final assessment even Semester questions.

#### 3.1 Validity

Validity is a measure or standard that shows the level of accuracy of an instrument. According toWidoyoko<sup>(13)</sup>states that the measuring instrument is said to be valid if the measuring instrument can accurately measure what is being measured. Based on the results of the analysis of the validity of the questions in the Final Assessment of Even Semester on theme 5 grade III in Public Elementary School 01 Sepauk, there were 20 questions or 57% of the questions were declared valid and 15 or 43% were declared invalid. Then the results of the analysis of the validity of the questions for the Final Assessment of Even Semester on the theme of 6 grade III in Public Elementary School 01 Sepauk,

there were 12 questions or 30% of the questions were declared valid and 23 questions or 70% were declared invalid.

In accordance with AnasSudijono's theory <sup>(12)</sup> "the validity of an questions of a test is the accuracy of measuring what should be measured through these questions". A valid question in theme 5 (57%) and theme 6 (30%) means that the questions has been able to carry out its function, which is to measure what should be measured. Then for invalid questions on theme 5 (43%) and (70%) it can be caused by various factors. This is in line with the theory stated by Grounlund in<sup>(14)</sup>which states that there are three factors that affect the validity of the test results, namely the instrument factor used for the test, administration and scoring factors, and factors from student answers. Based on the description above, it can be concluded that the validity of the questions in theme 5 Final Assessment of Grade III even Semester in Public Elementary School 01 Sepauk has sufficient quality, while in theme 6 Grade III Final Assessment of Grade III in Elementary School 01 Sepauk has poor quality in terms of its validity.

### **3.2 Reliability**

Reliability is the level or degree of consistency of an instrument <sup>(14)</sup>. The main purpose of calculating the reliability of test scores is to determine the level of accuracy and consistency of test scores. Based on the results of the analysis of multiple choice questions in theme 5, the results obtained that of the total 30 multiple choice questions were 0.86 which were in the "High" criteria and 5 in the essay questions, namely 0.78 which were in the "High" criteria. Meanwhile the results of the reliability test on theme 6 with a total of 30 multiple choice questions obtained a result of 0.11 which was in the "Very Low" criterion and 5 test questions obtained the result, namely -0.43 which was in the "Very Low" criterion, so that far of the reliability criteria of a test that is 0.70. So that it can be stated that the questions of the Final Even Semester Assessment on theme 6 are not reliable. A test instrument that has good validity on each questions will also have a high level of reliability as well.

### **3.3 Level of Difficulty**

Rusilowati<sup>(15)</sup>explains that the level of problem difficulty is the opportunity to correctly answer a question at a certain ability level which is usually expressed in the form of an index. Based on the results of the analysis of the difficulty level of the final assessment even Semester questions on theme 5 grade III in Public Elementary School 01Sepauk, there is 1 questions or 3% which is included in the difficult questions, 8 questions or 23% which is included in the medium questions and 26 questions or 74% are included in the easy questions. Then the results of the analysis of the difficulty level of the final assessment questions of even semester on the theme of 6 grade III in Public Elementary School 01 Sepauk, there are 3 questions or 8% which are included in the difficult questions, 9 questions or 26% included in medium questions and 23 questions or 66% included in easy questions. According to <sup>(16)</sup>if the questions are too difficult, no student will answer correctly, whereas if the questions are too easy then all students can answer all correctly so that it is considered not stimulating students to think and solve problems. So, from that the questions that have a medium level of difficulty can be used to make learning outcomes tests.

### **3.4 Discernment**

Explains that the calculation of distinguishing power is a measurement of the extent to which a question can distinguish students who have not or lacked competency based on certain criteria<sup>(14)</sup>. The higher the coefficient of distinguishing power of a question, the more able the questions are to distinguish between students who understand the material and students who do not / do not understand the material. Based on the results of the analysis of the distinguishing power of the final assessment questions of Even Semester on theme 5 grade III in Elementary School 01 Sepauk there are 13 questions or 37% have very good distinguishing power, there are no questions that have good distinguishing power, there are 5 questions 14% of questions or 14% had sufficient distinguishing power, and 17 questions or 49% had poor differentiation.

Then the results of the analysis of the distinguishing power of the final assessment questions of Even Semester on the theme of 6 grade III in Public Elementary School 01 Sepauk, there are 5 questions or 14% which have excellent distinguishing power, there are no questions that have good distinguishing power, there are 5 questions. 14% of questions or 14% had sufficient distinguishing power, and 25 questions or 72% had poor differentiation. Sudijono, who states that with the distinguishing power of each questions it is very important to know because one of the guidelines in arranging questions, question compilers must realize that the abilities of students are different<sup>(12)</sup>.

### 3.5 Tricking Quality

Based on the results of the tricking analysis of the final assessment questions of Even Semester on theme 5 grade III in Public Elementary School 01 Sepauk, it was found that out of a total of 30 multiple choice questions based on the quality of the tricking 9 questions or 30% were in the very good category, 11 questions or 37 % in good category, and 10 questions or 33% in very bad category. Then the results of the tricking analysis of the final assessment questions of Even Semester on theme 6 grade III in PublicElementary School 01 Sepauk obtained the results that from a total of 30 multiple choice questions based on the quality of the tricking 8 questions or 26% in the very good category, 11 questions or 37 % in good category, and 11 questions or 37% in very bad category.

Poor and bad tricking indicate that the cheaters are too conspicuous, so that the tricking do not have the appeal of being selected by test takers who do not understand the material. The quality of the tricking can be identified by analyzing the effectiveness of the tricker. Trickers are made not only to complement but to mislead students who don't understand the material, so they don't choose answer keys. Analysis of the quality of these questions was assisted using the Anates V4 application. Trickers function well if the alternative answers are selected by at least 5% of all test takers(11)

### 3.6 Questions Quality.

The results of quantitative analysis of multiple choice questions Final Assessment of Even Semester Themes 5 and 6 grade III on thematic learning which includes analysis in terms of validity, reliability, difficulty level, distinguishing power and distracting quality (except questions in description form) as a whole determines the percentage of questions good questions in the question pack. The questions used were questions in the "Very Good" and "Good" categories.

Based on the results of the analysis of the quality of the questions in theme 5, there is 1 question which has very good quality. There are 12 good quality questions. There are 8 questions that have sifficent criteria. There are 9 questions with poor quality, and 5 questions with very poor quality. the results of the analysis of the quality of the questions in theme 6, there were 2 questions which had very good quality. There are 5 questions of good quality. There are 5 questions that have sifficent criteria. There are 17 questions with poor quality, and 6 questions with very poor quality. The following is a breakdown of the causes for the poor quality of questions in grade III Public Elementary School 01 Sepauk:

**Table 6. Causes of Bad Questions Points**

Theme	Cause Failure	of	Number of questions	total	Percenta ge
<b>Theme 5</b>	Validity		1, 4, 6, 9, 10, 13, 14, 15, 16, 19, 22, 27, 28, 29, 30	15	43%
	Level Difficulty	of	1, 2, 3, 4, 5, 7, 8, 9, 11, 12, 14, 17, 18, 19, 20, 21, 22, 24, 25, 26, 27, 28, 29, 30, 31, 32, 34	27	77%
	Discernment		1, 2, 3, 4, 6, 7, 8, 9, 11, 14, 17, 19, 22, 24, 25, 26, 27, 29, 30, 32, 33, 35	22	63%
	Tricking Quality		3, 4, 9, 11, 17, 22, 25, 26, 29, 30	10	33%
<b>Theme 6</b>	Validity		1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 13, 15, 17, 19, 21, 23, 25, 26, 27, 28, 29, 32, 34	23	66%
	Level Difficulty	of	1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 20, 21, 23, 24, 26, 27, 30, 32, 34, 35	26	74%
	Discernment		1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34.	30	86%
	Tricking Quality		4, 5, 6, 7, 8, 14, 16, 18, 21, 24, 26,	11	37%

The quality of a questions is influenced by four factors (except for the form of description), namely validity, level of difficulty, distinguishing power and deceiving effectiveness. The low quality of the questions on theme 5 and theme 6 were analyzed based on the distribution of students' answers. The results of the analysis show that the factors that dominate the causes of the low quality of the questions are the difficulty level and distinguishing power. This shows that the questions are still at a difficult and easy level so that they cannot distinguish the abilities of high-ability students (understood the material) and low-ability students (not / not yet / not really understand the material). More than 50% of the questions in both themes had a level of difficulty and differentiation with poor quality.

### 3.7 Cognitive Domain Analysis.

Analysis of the distribution of the cognitive domains was carried out by matching the questions. Bloom's Taxonomy of cognitive domains that has been revised by Anderson and Krathwohl is: remembering, understanding, implementing, analyzing, evaluating, and creating.

**Table 7. Percentage of Themes 5 and 6 Based on Anderson's Cognitive Level**

Theme	Cognitive Level	Questions	total	%
Theme 5	C1 (Given)	2, 5, 8, 14, 15, 21, 23, 25, 26, 27, 34	11	31%
	C2 (Understand)	1, 7, 9, 10, 28, 31	6	17%
	C3 (Apply)	3, 4, 6, 11, 13, 16, 17, 19, 30, 33	10	29%
	C4 (Analyze)	12, 18, 20, 22, 24, 29, 35	7	20%
	C5 (Evaluate)	-	-	-
	C6 (Invent)	32	1	3%
Theme 6	C1 (Given)	4, 7, 8, 11, 12, 14, 15, 19, 23, 26, 30, 34	12	34%
	C2 (Understand)	1, 2, 3, 5, 13, 22, 29, 31, 32	9	26%
	C3 (Apply)	6, 16, 17, 18, 20, 24, 27, 33	8	23%
	C4 (Analyze)	9, 10, 21, 25, 28, 35	6	17%
	C5 (Evaluate)	-	-	-
	C6 (Invent)	-	-	-

The results of the distribution of the cognitive domain in the question of theme 5 Final Assessment of Even Semester in grade III Public Elementary School 01 Sepauk, there are 11 questions or 31% of the questions are at the C1 cognitive level. There are 6 questions at the cognitive level C2 or 17%. There are 10 questions at the cognitive level C3 or 29%. The highest distribution of questions is at the cognitive level C4, amounting to 7 questions or 20%. There is no distribution for the C5 cognitive level and only 1 question or 3% are at the C6 cognitive level. While the results of the analysis, the results of the distribution of the cognitive domain on the question of theme 6 Final Assessment of Even Semester in grade III Public Elementary School 01 Sepauk contained 12 questions or 34% of the questions were at the C1 cognitive level. There are 9 questions at the cognitive level C2 or 26%. There are 8 questions at the cognitive level C3 or 23%. The distribution of questions is at the cognitive level C4, which is 6 questions or 17%. There is no distribution of the C5 and C6 cognitive levels. Once teachers have developed an understanding of Education for Sustainable Development, they then need to be able to translate it into their pedagogy, or in other words, make it accessible for the children in their classrooms

### 3.8 Recapitulation of Questions Quality.

This recapitulation of questions is the result of analyzing the quality of questions that have been sorted and already have a quality that is feasible to use which is included with the results of the analysis of the cognitive level domain. Based on the results of the analysis, the results of the distribution of the sub-themes on the question of theme 5 Final Assessment of Even Semester in grade III Public Elementary School 01 Sepauk have 2 questions in sub-theme 1, 4 questions in sub-theme 2, 2 questions in sub-theme 3, and 2 questions in sub-theme 4 while distribution of the cognitive domain in theme 5, there are 4 questions in the C1 category, 1 question in the C2 category, 5 questions in the C4 category and there are no categories C3, C5 and C6 in the questions. While the results of the distribution of the sub-themes on theme 6 Final Assessment of Even Semester in grade III of Public Elementary School 01 Sepauk, there is 1 question in sub-theme 1, 2 questions are in sub-theme 2.

**Table 8. Results Recapitulation of the quality of the questions**

Theme	No of Question	Sub Themes				Cognitive Domain
		ST 1	ST 2	ST 3	ST 4	
Theme 5	Question 5		✓			C1 (Remembering)
	Question 12			✓		C4 (Analyzing)
	Question 18			✓		C4 (Analyzing)
	Question 20		✓			C4 (Analyzing)
	Question 21		✓			C1 (Remembering)
	Question 23				✓	C1 (Remembering)
	Question 24			✓		C4 (Analyzing)
	Question 31				✓	C2 (Understand)
	Question 34					✓ C1 (Remembering)
	Question 35					✓ C4 (Analyzing)
Theme 6	Question 12	✓				C1 (Remembering)
	Question 20				✓	C3 (implementing)
	Question 22				✓	C2 (Understand)
	Question 30			✓		C1 (Remembering)
	Question 35			✓		C4 (Analyzing)

The questions used are the final assessment questions of even semester on theme 5 and theme 6. The number of questions analyzed is 60 multiple choice questions and 10 essay questions, which have been prepared by the Education Office. Theme 5 is entitled weather and theme 6 is titled energy and its changes. Each theme has 4 sub-themes with five subjects, namely Bahasa Indonesia, Mathematics, Pancasila and Nationality Education, Art-Culture and Crafts, and Physical Education. In order to be able to improve the ability to think in a higher direction and an attitude of caring for the environment a means is needed in the learning process. One of the facilities that can be used by teachers is to use teaching materials based on science and conservation literacy, teaching materials are one of the most important components besides students and teachers in a learning process(17). This was expressed by (18)The existence of scientific literacy will make students reflective, critical, and intelligent, students will use scientific knowledge in the learning process.

The recapitulation of the quality of this question is the result of the questions that have been sorted by looking at the quality of the questions from their validity, distinguishing power and distracting quality. (19)A good question if it has valid validity, good distinguishing power and good distracting quality. The distribution of questions on theme 5 can be categorized as good because each sub-theme has a representative question but the distribution of subjects is not evenly distributed because there are more SBDP subjects compared to other subjects while in theme 6 there is no sub-theme 4 so that the distribution of questions is not evenly distributed and the spread of eyes lessons are not evenly distributed because there are no Mathematics and Pancasila and Nationality Education subjects. The results of the analysis of the cognitive domain are categorized as not good because the questions are only at C1 to C4 so that these questions cannot stimulate children in higher-order thinking. It is better if questions are made in the cognitive domains of C4 to C6 with a percentage of 40% because in the 2013 curriculum, children are guided to think at high levels. (20) the percentage of learning assessments on the cognitive aspects should be distributed to C1 as much as 5%, C2 as much as 10%, C3 as much as 45%, C4 as much as 25%, C5 as much as 10%, C6 as much as 5%.

#### 4. Discussion And Conclusion

Data were analyzed using Anates V4 and Microsoft Excel programs. The validity of theme 5; 20 categories are valid and 15 are categorized as invalid. Theme 6; 12 categories are valid and 23 are categorized as invalid. Test the reliability of theme 5; multiple choice is 0.86 and description is 0.78. Theme 6; multiple choice namely 0.11 and the description is -0.43. Difficulty level theme 5; 1 in the difficult category, 8 in the medium category and 26 in the easy category. In theme 6, there are 3 difficult categories, 9 medium categories and 23 easy categories. The distinguishing power of 5 themes; 13 are in the very good category, 5 are in the fair category and 17 are in the bad category. Theme 6; 5 are in the very good category, 5 are in the fair category, and 25 are in the bad category. Outbreak quality theme 5; 9 categories are very good, 11 are categorized as good, 10 are categorized as very bad. Theme 6: 8 in the very good category, 11 for the good category, 11 for the very bad category. The results of the distribution of cognitive domains on question 5; 11 categories C1, 6 categories C2, 10 categories C3, 7 categories C4 and 1 category C6. Theme 6; 12 items C1, 9 categories C2, 8 categories C3, 6 categories C4. The quality of the items in the Final Assessment of Even Semester in the thematic learning of class III Public Elementary School 01 Sepauk is in the poor category and needs to be revised.

#### References .

1. Novioleta, R, Wedyawati, N and Kurniati, A. (2020). Model Mind Mapping pada Pembelajaran Tematik kelas IV Siswa Sekolah Dasar. *Jurnal Pendidikan Dasar Perkhasa* Volume 6 Number 1 hlm. 41-54. Doi 10.31932/jpdp.v6i1.645.
2. Wedyawati, N. and Lisa, Yasinta, (2019). *Pembelajaran IPA di Sekolah Dasar*. Sleman: Deepublishing.
3. Wedyawati, N, Setyawan, A and Putri, M., (2020). *Pembelajaran SD Berbasis Problem Solving Method*. Yogyakarta: Literasi Nusantara.
4. Wedyawati, N, Setyawan A.E., Dike, D. (2020). Design of Gamikar (Game Mitigasi Kebakaran) for High-Grade Elementary Student. *Journal of Advanced Research in Dynamical and Control Systems* Volume 12. Doi: 10.5373/JARDCS/V12SP7/20202096.
5. Marius, A, Wedyawati, N and Rismawati, M., (2020). Minat Belajar dan Hasil Belajar Siswa Kelas V dengan Pembelajaran Berbantuan Power Point. *Journal Education and Techology* Vol. 1. Edisi 1. doi: <https://doi.org/10.31932/jutech.v1i1.694>.
6. Wedyawati, N. and Lisa, Y, (2017). Developing integrated natural sciences disaster mitigation model for elementary School Sintang. *Journal of Educational Science and Technology (EST)* Vol. 3.
7. Sudjana, N. (2016). *Penilaian Hasil Proses Belajar Mengajar*. Bandung: Alfabeta.
8. Afrida, (2020). Analisis Kualitas Butir Soal Ujian Semester Ganjil pada Mata Pelajaran Pendidikan Kewarganegaraan kelas V MI. *Jurnal Keilmuan dan Pendidikan* Vol. 12.
9. Arbiatin, E.M. (2020). Analisis Kelayakan Butir Soal Tes Penilaian Akhir Semester Mata Pelajaran Matematika Kelas VI di SD Negeri 19 Ampenan Tahun Pelajaran 2019/2020. *Journal El Midad*, 2020, Vol. 12 edisi 2. doi:10.20414/elmidad.v12i2.2627.
10. Sugiyono, (2016). *Metode Penelitian Kuantitatif, Kualitatif dan R&D*. Bandung: Alfabeta.
11. Muluki, A., (2020). Analisis Kualitas Butir Soal Tes Semester Ganjil Mata Pelajaran IPA Kelas IV MI Radhiatul Adawiyah. *Jurnal Ilmiah Sekolah Dasar* Vol. 4.
12. Werdiningsih, G., (2015). *Analisis Kualitas Butir Soal Penilaian Akhir Semester Gasal Mata Pelajaran Ekonomi*. Yogyakarta: eprints-Lumbung Pustaka UNY.
13. Widoyoko, E.P., (2014). *Penilaian Hasil Belajar di Sekolah*. Yogyakarta: Pustaka Pelajar.



14. Zainal, A., (2017). *Evaluasi Pembelajaran*. Bandung: RemajaRosdakarya.
15. Rusilowati, A., (2014). *Pengembangan Instrumen Penilaian*. Semarang: Unnes Press.
16. Purwanto, (2016). *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Pelajar.
17. Ruwiyatun, Rusilowati, A and Saptono, S. (2020). Improving Students' Higher Order Thinking and Environmental Attitudes through Implementation of Teaching Materials Based on Scientific Literacy and Conservation. *Journal of Innovative Science Education* Vol. 9.
18. Smith, K.V., Loughran, J., Berry, A. and Dimitrakopoulou, (2012). Developing Scientific Literacy in a Primary School. *International Journal of Science Education* Vol. 34. (1)
19. Hartadiyati, E., et al., (2020). Pedagogical Content Knowledge (PCK) of Prospective Biology Teacher on Respiratory System Material to Education for Sustainable Development. *Journal of Physics: Conference Series. International Conference on Mathematics and Science Education*. IOP Publishing Conf. Ser. 1521 042034.
20. Helmawati, (2019). *Pembelajaran dan Penilaian berbasis HOTS*. Bandung: RemajaRosdakarya.

#### BIOGRAPHIES OF AUTHORS



Nama Lengkap & Gelar : Nelly Wedyawati, S.Si., M.Pd.  
NIP/NIDN : 1111058501  
Golongan Kepangkatan : Lektor  
Alamat Email : [Nellywedyawati@gmail.com](mailto:Nellywedyawati@gmail.com)  
Nomor Kontak (Telp & HP) : 0085245229150



Nama Lengkap & Gelar : Wiputra Cendana. B.Sc., M.Pd.  
NIP/NIDN : 0308128805  
Golongan Kepangkatan : Asisten Ahli  
Alamat Email : [wiputra.cendana@uph.edu](mailto:wiputra.cendana@uph.edu)  
Nomor Kontak (Telp & HP) : 085739919999