

The Development of Performance Assessment Instrument Based on Scientific on Thematic Integrated Learning in Fourth Grade in Elementary School

Minarni^{a*}, Rochmiyati^b, Irawan Suntoro^c, Darsono^d

^aLampung University, Jl. Prof. Dr. Soemantri Brojonegoro No.1 Bandar Lampung, 35145, Indonesia

^aEmail: minarni8206@gmail.com

^bEmail: rochmiyatiwardi@yahoo.com

^cEmail: irawan.suntoro.dr.@gmail.com

^dEmail: darsono.1954@fkip.unila.ac.id

Abstract

This study aims to develop performance assessment instrument based on scientific on thematic integrated learning in fourth grade in Elementary School. The type of research used is research and development that refers to the theory of Borg & Gall (1983). The population in this study were all teachers and fourth grade students in Elementary School 2 in Langkapura , Bandar Lampung City. The samples were taken by using purposive random sampling technique as many as 2 teachers and 58 students in Elementary School 2 in Langkapura. The data is collected through questionnaire sheets and observations. The results of this study indicate that the performance assessment instrument developed based on the feasibility test was stated very good. The results of the validity test stated that the performance assessment instrument was declared valid. Reliability tests on performance assessment instrument were also stated reliability. It was concluded that the performance assessment instrument was valid and reliable for measuring and assessing student performance in both process and product aspects in fourth grade students.

Keywords: Performance Assessment Instrument; Scientific; Thematic.

1. Introduction

Assessment in the curriculum 2013 (K13) or better known as authentic assessment has strong relevance to the scientific approach to learning. The assessment illustrates the skills of observing, reasoning, trying, building networks, and others.

* Corresponding author.

The implementation of assessment of learning outcomes is carried out starting from the determination of the instruments, preparation of the instruments, review of the instruments, implementation of assessments, analysis of assessment results, and follow-up programs on assessment results. The assessment of student learning outcomes in basic education includes aspects of attitude (affective), knowledge (cognitive), and skills (psychomotor).

The attitude assessment can be done by observation, assessment of knowledge through written tests, oral tests, and assignments, as well as skills assessment through performance assessment. Performance assessment is an assessment that requires students to demonstrate a particular competency. Daniela Tuparova's [1] study reveals that performance assessment is a measure of assessment based on original tasks such as activities, exercises, or problems that require students to show what they can.

Performance assessment is part of authentic assessment. In performance assessment, emphasis can be made on the process or product. Performance assessment involves students in activities that require demonstration for certain skills in terms of creating a product as a result. Assessment of process performance by observing students when they are performing or assessing the level of demonstration skills of the students. Research from Marzanu. R.J. Pickering [2] revealed that performance assessment consists of authentic tasks that require students to apply what they know in the real world context. The students play an active role in the class because they are aware of the problems around them, provide different solutions and organize activities that help them to maintain their learning process efficiently. Effective teaching is structuring information by assigning performance tasks to students. Research by Rufina C. Rosaroso and Nelson A. Rosaroso [3] shows that students are highly motivated to learn in special classes when lessons are integrated with performance-based tasks. Students become independent because they work individually and in groups.

Students' performance instruments use the rubric, through observation in the learning process. W. I. Griffith and Hye Yeon Lim [4] explained that to make an assessment based on performance requires a good rubric to focus on specific skills or tasks. The formulation of performance assessments uses clear rubrics, and introduces technology into the classroom. This turned out to increase the motivation and performance of students. Understanding the relationship between performance assessment, rubrics, and learning activities can build the conditions of the learning class students can learn and apply to their needs in the real world, in the classroom, and outside the classroom.

According to Hosnan [5] explains that assessment is a teacher's activity which is intended to measure certain competencies or abilities for activities that have been carried out in learning activities. According to Mangiante [6] assessment is a tool to measure the extent to which students have improved their learning based on the standard. Whereas according to Harlen [7] Assessment is defined as the process of gathering and using facts for a particular purpose about the learning outcomes of the students.

Hosnan [8] argues that: scientific is a learning process that is arranged so that active students build the concepts, laws or principles through the stages of observing (to identify or find the problems), formulate problems, submit or formulate hypotheses, collect data with various techniques, analyze data, draw the conclusions and find

concepts, laws or principles found. The scientific approach according to Faiq [9] is essentially a golden sign of the development of the students' attitudes, skills and knowledge.

The regulation of the minister of education and culture Number 24 in 2016 concerning the basic framework and Elementary School / Madrasa Ibtida'iyah curriculum structure. The curriculum 2013 demands a complex assessment that includes all competencies namely Core Competencies (KI-1), (KI-2), (KI-3) and (KI-4). Performance assessment is used to assess students' skills through, among others, assignments (task), student interaction. Assessing the performance of students requires a comprehensive criterion called the rubric. Forms of performance assessments compiled include tasks (tasks) and rubrics (assessment criteria). Assessment is carried out on skills in performance and student interaction. The tasks are specifically designed to produce responses (oral or written), produce works (products), or show the application of knowledge. The tasks that given to students must be in accordance with the competencies to be achieved and meaningful for the students.

Performance assessments not only assess results but to see how students are actively in the learning process. Teachers will be more objective in assessing students with performance assessment. As expressed by (Popham [10] that in evaluating student performance, the teacher requires a response that is "authentic" or original in the form of observable activities. The tasks that given can take the form of oral or written, the type of assignment is adjusted to the learning objectives.

Based on the results of the needs analysis shows that the teacher in Elementary School 2 in Langkapura, Bandar Lampung did not understand the aspects of performance assessment. When the teacher conducting a performance assessment, the teacher only provides an assessment of the results that obtained by the students, without conducting a skills assessment based on scientific. In the implementation of the assessment, it turns out that the teacher experienced a lot of obstacles and problems that resulted in the teacher experiencing difficulties in conducting the assessment. Performance assessments in teacher books only assess globally / publicly, not yet in detail, especially in evaluating the performance of students, have not included instructions for use, predicates and descriptions. Difficulties in compiling and using assessment so that performance assessment has not been carried out optimally, because the assessment is dominated by written assessment.

Based on the description, the teacher is required to creatively develop performance assessment, so that it can support the development of the Learning Implementation Plan, in accordance with the material and learning approach. The goal is to improve students' performance skills in conducting experimental activities and communicating in learning. It is proper for teachers to understand and have skills in evaluating student learning outcomes, making teachers able to develop assessment instruments that are in accordance with certain rules. The assessments that are prepared in accordance with the drafting rules will produce a valid and reliable assessment. So that it will produce data and information about the level of achievement of student competencies in a valid and accurate manner.

In accordance with the problems above, the purpose of the research and development is to develop a performance assessment instrument based on scientific in thematic learning with theme 2 Always Save Energy and the sub-theme (3) Alternatif energy in fourth grade in Elementary School that valid and reliable.

2. Method

2.1 Types of Research and Procedures

Research is Research and Development (R & D). Research and development is research that aims to produce certain products. The products that produced are then tested for validity and reliability. The research and development used is the design model Borg & Gall [11] which consists of 10 steps. Steps that must be followed to produce the product, namely: research and initial information gathering, planning, development of the initial product format, initial trials, product revisions, field trials, product revisions, operational field trials, final product revisions, implementation. This study only carries out steps one through the seventh step, namely the preliminary study steps up to the field trials.

2.2 Population and Samples

The population of this research and development is all fourth grade teachers in Elementary School Langkapura, Bandar Lampung City who have implemented the curriculum 2013. In addition to teachers, this research and development also uses students as the sample. Samples were taken using purposive random sampling techniques. This sampling technique is used by considering the elements or categories in the study population. This technique is used to determine the number of teachers who assess the instruments that developed. The sample of students in this study were 58 fourth grade students of Public Elementary School 2 in Langkapura as product trial samples.

The instrument of data collection is a tool that used to collect data in research. Data is obtained from the preliminary study phase, product development and product testing. The preliminary study phase, the instrument in the form of interview guide sheets, observation sheets, and questionnaire sheets for expert validation and teacher validation.

Data analysis that used in this research activity is descriptive qualitative and quantitative descriptive data analysis. Qualitative descriptive data analysis in this study was used to process data sourced from comments and suggestions obtained from material experts, linguists and evaluation experts that contained in the validation questionnaire, an initial trial to determine the validity and usefulness of the instrument. The results of qualitative descriptive data analysis are used as a condition to meet valid theoretical criteria. Quantitative descriptive data analysis was used to analyze the data obtained in the form of a validation score assessment for material experts, linguists and evaluators to assess the validity of the content of the instruments developed, the results of the teacher's response questionnaire to measure the usefulness and feasibility of the product. As well as test results of students to measure the level of reliability of the instrument.

A data or information can be stated to be valid if it is in accordance with the actual situation [12]. Then the instrument is stated to be valid if the instrument used can measure what should be measured. Data from the validation sheet provides an overview or exposure to the quality of the assessment instruments that developed. While usability level analysis is used to see the feasibility level of the product being developed. The data from the validation sheet provides an overview or exposure to the quality of the instruments taken from the results of

the validation of the assessment experts, linguists, material experts and teachers, which are developed using the following formula.

$$Vah = \frac{T_{se}}{T_{sh}} \times 100\%$$

Information:

Vah = Expert validation

Tse = Total empirical score

Tsh = Maximum total score [13]

The validity criteria are as follows

Table 1: Criteria for Instrument Validity

Final Score	Criteria
80 % - 100%	Very valid, very complete, can be used
61% - 80%	Valid enough, quite effective, can be used
41% - 60%	Less valid, less effective, not used
21% - 40%	Invalid, ineffective, cannot be used
00% - 20 %	Very invalid, cannot be used

Empirical criteria are used to analyze items. The item analysis was used to test the quality of the questions tested in the main field test. Quantitative data in testing empirical criteria using reliability testing. Instruments are said to be reliable if they provide consistent or consistent results when repeatedly tested. Instrument reliability testing in this development research was conducted to test the reliability of measuring instruments. The reliability test of these heterogeneous performance assessment instrument using the alpha coefficient formula. The formula used to calculate the reliability of heterogeneous learning outcomes [14] is:

$$r_{11} = \left(\frac{k}{k-1} \right) \left(1 - \frac{\sum \sigma_b^2}{\sigma_1^2} \right)$$

Information:

r_{11} = reliability sought

$\sum \sigma_b^2$ = number of item variants

σ_1^2 = total variant

K = number of questions

The reliability category is used as follows

Table 2: Instrument Reliability Categories

Coefficientr	Reliability
0,80 - 1,00	Very Strong
0,60 – 0,79	Strong
0,40 – 0,59	Normal
0,20 – 0,39	Low
0,00 – 0,19	Very low

Source: Sugiyono (2010: 257)

3. Results

The results of this research and development are performance assessment instrument of students who meet valid and reliable criteria, for fourth grade students of elementary school in the odd semester, Theme 2 Always Save Energy Sub-theme 3 Alternative Energy. The procedure used in this study refers to the development model Borg & Gall which contains the main steps of the study as follows:

3.1 Preliminary Research and Data Collection

The initial information collection was carried out by observing, interviewing and distributing questionnaires in the form of questionnaires to teachers in Langkapura who had implemented the curriculum 2013. After that the researchers conducted a literature study to examine the theories and results of previous studies related to the products to be developed. Field surveys are conducted to collect data relating to needs of the researchers.

The results of the initial information collection through literature study and needs analysis obtained data that, the real conditions that exist in Langkapura, Bandar Lampung City are: 1) the assessment is still dominated by written assessment; 2) the Curriculum 2013 target teachers are currently having difficulty developing performance assessment instrument and assessment methods, so that performance assessments cannot be optimally performed; 3) the assessment process so far only emphasizes the mastery of concepts (knowledge) carried out with objective paper and pencil tests as a measurement tool; 4) performance assessments contained in teacher books are almost all the same, not varied; 5) assessment of student learning outcomes focusing on results not on the process so that students are passive.

Based on the results of the analysis, the potential for developing performance assessment is the solution to the problems that faced. The researchers develop a performance assessment rubric that can assist teachers in conducting the assessment process when students perform the performance.

3.2 Planning

This stage is carried out by designing a prototype of a performance assessment instrument based on the suitability of Core Competence, Basic competence, and Indicators in the teacher's book used in the study.

3.3 Initial Product Format Design

The drafting of the initial product was carried out as a follow-up to the planning step. The initial product drafting steps are as follows: 1) mapping the instrument, (b) compiling the grid, (c) compiling the instrument, (d) analyzing the quality of the instrument qualitatively, (e) testing the measuring instrument, (f) implementing measurements.

3.4 Early Test

The preliminary trial to determine the feasibility of performance assessment instrument based on scientific in thematic learning based on the response of teachers and fourth grade students of Public Elementary School 2 Langkapura. The initial trial used fourth grade teacher respondents as many as 2 people and fourth grade students numbered 6 people, students in the high, medium and low categories. The results obtained are as follows:

Table 3: Results of Questionnaire Usability Value in the Early Test.

No	Rated aspect	Teacher 1	Teacher 2
1.	Construction	90 %	83 %
2.	Language	90 %	87 %
3.	Writing Rules	100 %	100 %
4.	Total score	69	67
5.	Maximum score	75	75
6.	Percentage of Scores	92 %	89%
7.	Score Range	84%-100%	84%-100%
8.	Average Score in every Teacher	4,6	4,5
9.	Average overall Teacher Score	4,6	
	Percentage average	90,5%	
	Criteria	Very Worthy	

Source: Results of Primary Data Processing

Table 4: Questionnaire Results of Readabilty Values in the Preliminary Test

No	Rated aspect	Students					
		1	2	3	4	5	6
1.	Instrument value	100%	96%	84%	80%	80%	76%
2.	Language	100%	96%	88%	80%	76%	72%
3.	Total score	50	48	43	41	40	37
4.	Maximum score	50	50	50	50	50	50
5.	Percentage of Scores	100%	96%	86%	82%	80%	74%
6.	Score Range	84%-100%		84%-100%		67%-83%	
7.	Average Per Score	5,0	4,8	4,3	4,1	4,0	3,8
8.	Average All Student Scores	4,3					
9.	Percentage average	86,3%					
10.	Criteria	Very Worthy					

Source: Results of Primary Data Processing

3.5 Product Revision

Product revisions are based on suggestions from the validator), as well as the results of the initial trial. Valuation is carried out by material expert lecturers, linguists, and evaluation experts. Suggestions obtained from the validation process from material experts are: a) adjusting instrument indicators with operational verbs, b) indicators must be adjusted to the statements in Basic Competence, c) indicators must cover the contents of existing basic competencies, d) revise material adapted to performance assignments.

The suggestions from linguist lecturers, namely: a) revise the punctuation on the performance assessment instrument, b) revise the use of capital letters for each aspect of the assessment, c) Add description on the rubric of the performance assessment instrument so that the assessment is more specific to the experimental material and tools, and work steps experiment, d) revise the writing of the performance assessment instrument, adapted to the Indonesian Spelling General Guidelines, and e) revise the instructions and instructions on the performance assessment instrument, so that the scientific process is more visible.

The suggestions from expert evaluation lecturers are: a) revising instrument indicators with operational verbs, b) revising the instrument usage instructions must be clear in the performance assessment instruction to make it easier for users, c) revise the instrument by creating a scoring rubric, and d) revise the casting items made from complete to very complete.

3.6 Main Field Test

The main field trials were held at Public Elementary School 2 Langkapura, Bandar Lampung City, in 2 teachers and 2 classes, that are fourth grade in A class with 28 students and fourth grade in B class with 30 students. The

main field trials were conducted to measure reliability levels product instrument. Measurement of reliability using the alpha coefficient formula.

Based on the data analysis shows that the instruments developed in the study obtained reliability test results of 0.740 in fourth grade in A class with the criteria of consistency in the reliability level in the strong category, and 0.470 in the fourth grade in B class with the criteria of consistency in the reliability level in the medium category. The results of the validity and reliability tests show results that are in accordance with the testing criteria, and state that the performance assessment instrument of students who are developed are valid as empirical and reliable. This is in accordance with the results of Suwaibah's research (2015: 8) that the reliability test results were obtained from the high coefficients on the performance instrument assessment process.

2.7 Product Revision

Based on the results of the usability trial by teacher and readability by students who were subject to small trials and main field trials, the instrument products developed were not revised and were feasible to be implemented.

4. Discussion

4.1 Realization of Performance Assessment Instrument Based on Scientific in Valid Thematic Integrated Learning in Fourth Grade

Based on research and development, literature studies and the needs of analysis, information was obtained that the real conditions in Langkapura in Bandar Lampung City were: 1) the assessment was still dominated by written assessment; 2) the Curriculum 2013 target teachers are currently having difficulty developing performance assessment instrument and assessment methods, so that performance assessments cannot be optimally performed; 3) the assessment process this far emphasizes the mastery of concepts (knowledge) that done with objective paper and pencil tests as a measurement tool; 4) performance assessments contained in teacher books are almost all the same, not varied; 5) assessment of student learning outcomes focusing on results not on the process so that students are passive.

Development of performance assessment instrument to improve student performance skills in conducting experiments or activities to try and communicate in scientific learning by providing a number of material assistance and instructions to students. So that the teacher can develop an authentic assessment that is feasible and fill certain criteria according to the assessment of the response of students, teachers and three validators. In accordance, with the learning theory from Vygotsky, namely Scaffolding. Scaffolding means giving a large amount of assistance to a student during the early stages of learning then the student takes on the responsibility that gets bigger as soon as he can do it. The assistance is in the form of instructions, warnings, encouragement, describing the problem into the steps of solving, giving examples, or others so as to enable students to grow independently.

The assessment of the performance of fourth grade students is applied to learning in schools that implement Curriculum 2013. The curriculum uses integrated thematic learning. Integrated learning as a teaching activity by

combining the material of several subjects in one theme and at the same time. Learning in the curriculum 2013 uses a scientific approach that consists of seeing, asking, trying, reasoning, and communicating, this is in accordance with Dyers' theory in Maryanti (2015: 1) that the ability of creativity can be obtained through observing, asking, trying, reasoning, and making network.

The development of performance assessment instrument is based on the strengths of performance assessment. Performance assessment is used as an alternative test that has been widely used to measure the learning success of students in school. Therefore the use of performance assessment is important in the learning process because it can provide more information about the abilities of students in processes and products, rather than just getting information about just right or wrong answers. Students are better able to theorize, but are less skilled at doing the theory. this is in line with the opinion of Stiggins (1994: 161) revealing that there are several reasons why performance assessment needs to be done as follows: 1) Giving more opportunities to the teacher to recognize more fully because in reality not all students who are less successful in objective tests or essays automatically can be said to be unskilled or not creative. Thus, the assessment of student performance complements other assessments, 2) The ability of students who are difficult to know or detect only by looking at the final results of their work, or only through written tests in terms of skills and creativity.

This study refers to the skills competency assessment, according to Kunandar (2015: 259) in the realm of skills there are five levels of thinking processes namely: (1) imitation, (2) manipulation, (3) precision, (4) articulation, and (5) naturalization. The development of this assessment is at the level of imitation and manipulation. Imitation is the ability to carry out simple activities and exactly the same as what has been seen or noticed before, and manipulation is the ability to carry out simple activities that have never been seen, but based on instructions or guidelines.

The performance assessment instrument is used as an alternative assessment of what has been done so far to measure student learning success. Therefore, the use of performance assessment instrument is important to be implemented so that the students really know the ability not only the results but also the learning process, so that there is no subjective assessment tendency. The teacher not only assesses the right and wrong answers without any reason, but the teacher must also assess the students' abilities when practicing. In line with Cabrera's (2001) research, this study showed that the instruments developed fulfilled several conditions recommended by the assessment literature, that are: (1) meaningful to users, (2) reliable and valid, and (3) observable index behavior avoided from a subjective impression.

The instrument that developed has passed several stages, starting from fulfilling the rules of writing instruments, validating theoretically and validating empirically. The results of the validity are in line with the research conducted by Budhiwaluyo (2016) in his research that the results of field trials show that the product has a high value of validity and reliability in measuring the performance of student practicum.

The results of the content validity are in line with the research conducted by Putri (2017), the results of her research show that the performance assessment developed has fulfilled the content validity based on evaluation by 3 experts and 3 practitioners. The Reliability of all rubrics in performance assessment is categorized very

high. The same study conducted to obtain content validity from experts done by Kurniawaty (2017) results showed that each item in the performance assessment instrument based on thematic learning of fourth grade of Elementary School students that had been validated by a team of experts was deemed feasible to measure psychomotor or skill aspects of the students.

The feasibility of developing the product is in line with the results of research conducted by Ratnami (2016), about the feasibility of performance assessment, that is the quality of the results of the development of performance assessment according to expert review, that is the content expert test is very good qualification, 90.00%; the design expert test of learning is in very good qualification which is 92.00%; the expert assessment test was in the very good learning qualification of 90.00% and the field trial was in a very good qualification of 90.76%). Usman (2014) in his research produced teacher response data in a very good category.

Mardhapi (2004) based on the articles compiled can be concluded that the assessment instrument is not always a written test form that can be an observation guide, but must have evidence of validity and reliability.

This is in accordance with the principle that must be considered by the teacher in conducting the assessment, according to Sudaryono (2012: 54-55), that are: (1) the principle of sustainability, (2) the overall principle, (3) the principle of objectivity, (4) the principle of validity and reliability (5) the principle of measuring criteria, (6) the principle of usability.

Based on the results of the study that the performance assessment instrument of students in theme learning always save energy has many advantages compared to the psychomotor aspect assessment instruments contained in the teacher's book. The following is the difference between the products developed and the existing instruments:

Table 5: The Differences in Student Performance Assessment Instrument Results of Development with Performance Assessment Instrument in Teacher Books

Skills Assessment Instrument in the Teacher's Book	Development Result Student Assessment Instrument
Instruments are presented globally / publicly but have not detailed capabilities that are assessed both by process and product.	The instruments are presented in detail in terms of the aspects and processes of the product.
Have not included instructions for use so the teacher has difficulty using it	The instructions for using the book are clear
Rubrics are always the same for each lesson	Rubrics are more varied
The criteria in the aspects assessed are unclear	The criteria that used are clear to assess student performance, making it easy to use.

Based on these differences, there are differences in the advantages of the products that developed, so that they have the potential to continue to be developed and used as instruments of performance assessment in learning. The instrument developed fulfills the requirements as a good evaluation tool so that it is worthy of being used as an assessment instrument, in line with Pratiwi's research (2016) that the performance assessment instrument developed has high validity and reliability so that it is used as an assessment instrument to assess student performance. in learning.

4.2 Realization of Performance Assessment Instrument Based on Scientific in Reliable Thematic Integrated Learning in Fourth Grade

The Instrument reliability testing in development research was conducted to test the reliability of the instrument. Testing the reliability of heterogeneous performance assessment instrument using the alpha coefficient formula. Based on data analysis shows the instrument developed in this study obtained reliability test results of 0.740 in the fourth grade in A class with the criteria for the consistency of the reliability level in the strong category, and 0.470 in the fourth grade in A with the criteria for the consistency of the reliability level in the medium category.

The results of validity and reliability tests show results that are in accordance with the testing criteria, and state that students' performance assessment instrument developed validly as empirically at a reliable level. This is in accordance with the principles that must be considered by teachers in conducting assessments, according to Sudaryono (2012: 54): Principles of validity and reliability.

5. Conclusion

This study produced valid and reliable performance assessment instrument based on Scientific, for fourth grade in Elementary School in the odd semester. Theme 2 Always Save Energy Sub-theme 3 Alternative energy, which consists of 5 learning content, that are Civic education, Indonesian Language, Natural Sciences , Culture and Practice Art and Social Sciences. Validity test results show that the level of validity in the good / valid category. While, the reliability test results show that the assessment instrument has a level of reliability in the good category.

Performance assessment gives more opportunities to teachers to recognize students more because in reality not all students who are less successful in cognitive tests are not necessarily unskilled or not creative. This performance assessment provides an opportunity for teachers to assess very objectively in accordance with the ability of students to perform performance in the learning process, and answer the assessment demands that are in accordance with the curriculum 2013, that is authentic assessment. Thus, performance assessment is a process assessment to train students' skills and support the assessment of student learning outcomes.

References

- [1] Tuparova, Daniela. Automated real-live performance-Based Assessment of ICT Skills. Procedia Social and Behavioural Science. 2010. 2: 4747-4751.

- [2] Marzano, R, J, D. Pickering, and J McTighe. *Assessing Students Outcomes. Performance Assessment Using the Dimensions of learning Model*. ASCD Publications. Alexandria. 1993.
- [3] Rufina C. Rosaroso and Nelson A. Rosaroso. *Performance-based assessment in Selected Higer Education Institutions in Cebu City, Philippines*. *Asia Pacific Journal of Multidisciplinary Research*. 2015. 3 (4): 72-77.
- [4] W. I. Griffith, and Hye-Yeon Lim. *Performance-Based Assessment: Rubrics, Web 2.0 Tools and Language Competencies*. *MEXTESOL Journal*, 2011. 36, (1). http://mextesol.net/journal/index.php?page=journal&id_article=108. Accessed on: February 9, 2017
- [5] [8] Hosnan. *Scientific and Constitutional Approach in 21st Century Learning: The Key to the Success of the Curriculum 2013*. Jakarta: Ghalia Indonesia. 2014.
- [6] Mangiante, Elaine Silva. *Planing Science Intruction for Critical Thinking: Two Urban Elementary Teachers 'Responses to a State Science Assessment*. *Journal of Education Science*, 2013. Vol. 3: 222-258.
- [7] Harlen, W. *Assessment & Inquiry Based Science Education*. *Triestly Italy: Global Network of Science Academies (IAP) Science Education Program (SEP)*. 2013.
- [9] Popham, W.J. *Education Evaluation*. Englewood Cliffs N. J: Prentice-hall _____. 1978. *Measurement Referenced Criteria*, Englewood, Cliffs, N.J: Prentice-p: //web.ku.edu. 1995.
- [10] Borg & Gall *Educational Research: An Introduction*, 4th ed. Yew York & Londong: Longman Inc., 1983.
- [11] Arikunto, Suharsimi. *Fundamentals of Education Evaluation*. Second edition. Jakarta. Earth Literacy.. 2013.
- [12] Sudijono, Anas *Introduction to Educational Evaluation*. Rajawali Pers. Jakarta. 2013.
- [13] Suwaibah, Siti Nur. *The Development of Authentic Performance Assessment Instruments Based on Authentic Assessment with Estimated Reliability Using the Genova Program*. *Chemistry in Education*. 2015. 5 (1): 8-14.
- [14] Maryanti, EndahFebriana. *PETASAN GALAU Authentic Assessment Instrument in Business Entrepreneurship Subjects*. *Journal of the UNILA Social Study*. 2015. Vol.2 No. 4 p. 24
- [15] Stiggins, R.J. *Student-Centered Classroom Assessment*. New York. Mac Millan College Publishing Company. 1994.
- [16] Kunandar. *Authentic Assessment (Assessing Student Learning Outcomes Based on the Curriculum*

- 2013). Jakarta. Raja Grafindo Persada. 2015.
- [17] Cabrera, F. Aberto. Developing Performance Indicators for assessing Classroom Teaching Practices and Student Learning. *Research in Higher Education*. 2001.. 42. (3). PP 327-352.
- [18] Budhiwaluyo, Nugroho. The Development of Performance Assessment Instrument on Practical Cell Structure and Function in Public High School 1 Jambi. *Edu-Science Journal*. (2016) Vol 5. No.2.
- [19] Putri, Suwandi F & Istiyono, Edi. The Development of Performance Assessment of Stem-Based Critical Thinking Skills in the High School Physics Lessons. *International Journal of Environmental And Science Education*. 2017. 12 (5) pp. 1269-1281.
- [20] Kurniawaty. The Development of Performance-Based Instruments on Thematic Learning in Elementary School. *Pedagogy Journal*. 2017. Vol 6. No 3.
- [21] Ratnami, Made V. Performance assessment in Science Learning In Fifth Grade Students of Public Elementary School 1 Banyuning, Buleleng District, Buleleng Regency 2015/2016 Academic Year. University of Education Ganesha. 2016.
- [22] Mardapi, Djemari. *Technique of Preparing Non TestTest Instruments*. Yogyakarta: Cendikia Partners. 2004.
- [23] [24] Sudaryono. *Basics of Learning Evaluation*, Yogyakarta. Graha Ilmu. 2012.