Development of Authentic Assessment Tools in Physics Science Learning

Enos Taruh* and Mursalin
Department of Physics, Gorontalo State University, Gorontalo, 96128, Indonesia
*enos.taruh@gmail.com

Abstract: This study is funded by The 2016 Indonesian Budget in the scheme of Graduate School Team Research. This study aims to develop an authentic assessment tool in Physics subject of Class X high school and to produce an authentic assessment tool that is feasible, effective, and practical to be used in the Physics subject of high school. The method used in this research is research development, which is developing authentic assessment tool in physics subject learning of high school. This research procedure refers to the development model by Thiagarajan, Semmel, and Semmel known as the four-D model ie define, design, develop, and disseminate. The result of this research is the product of authentic assessment tool of Class X high school Semester 1 on the cognitive, affective and psychomotor domains that were categorized as feasible, practical and effective. It is indicated by percentage of average score of feasibility 91.35%, reliability 0.94, and percentage of average score of practicality 83.85%, and learning result completeness of 80%.

1. Introduction
Ministry of Education and Culture Year 2015 in its discretion has determined that any educational institution can reapply The School-Based Curriculum (KTSP) and also for units of primary and secondary education that has two (2) years of implementing the curriculum of 2013 should continue its implementation in order to be used as a trial for the improvement.

In connection with that policy, the school imposes two (2) types of curriculum. There are class that implement curriculum the school-based curriculum (KTSP) and there are class that implement curriculum of 2013. It makes the teachers have difficulty in making assessment based on the applied assessment standard.

Many teachers experience confusion in the implementation of the curriculum 2013, especially in the assessment of student learning outcomes. The assessment of student learning outcomes required the authentic assessment, but the assessment guidelines have not been socialized and mastered by the teachers as a whole. The guideline is unequally distributed in schools. Based on these problems, developing authentic assessment tools in learning is necessary. The authentic assessment is a comprehensive assessment conducted to assess the input, the process, and the output of learning of the Education Assessment Standards) [1].

Another problem faced by students of Educational Physic in the completion of thesis is limitations in terms of understanding the research methodology. The results of the study of Enos Taruh in 2012 about Mapping the Quality of Scientific Work (Thesis) of students in Faculty of Mathematics and
Natural Sciences State University of Gorontalo shows that there are three (3) main difficulties students preparing thesis. The three main difficulties include preparation of background problem, preparation of literature and frame of mind, and discussion of the results. On the other hand, the school Headmaster where the student on duty wants them to complete their studies as soon as possible because of their presence is needed in schools. It related to fact that numbers of physics teachers in each school are categorized as relatively less [2].

Substantively the students in addition of faced with the urge to finish his studies, they also can realize the goals of entering the Graduate Program, which is in addition can develop learning in the form of innovations in learning, can also develop authentic assessments in Physics learning. To achieve these objectives, it is necessary to process programatically to facilitate students in completing their studies through the joint research of Graduate School Team with the title: "Development of Authentic Assessment Tools in Learning Physics Class X Senior High School". The student involved in this research will be able to add insight and experience in completing the writing of his thesis.

Based on the background, the problem is examined in this study are as follows:

How to develop the 2013 curriculum based of authentic assessment tool in physics learning for first grade of senior high school using Four-D models?

How is the feasibility, effectiveness, and the practicality of authentic assessment tools were developed, so it can be used as a guide for teachers in 2013 curriculum-based assessment?

In particular, this study aims to: (1) developing authentic assessment tools in physics learning for first grade of senior high school with Four-D model (Define, Design, Develop, and Disseminate); (2) to generate authentic assessment tools that are effective and appropriate for use in Physics learning in Senior High School.

The Curriculum 2013 requires the use of authentic assessment. Paradigmatically, authentic assessment requires the realization of authentic instruction and authentic learning. It is believed that authentic judgment is better able to provide learners with a holistic and valid ability when compared to traditional judgments.

Comparison between authentic and traditional assessments is expressed in table [4].

**Table 1.** Comparison Between Authentic And Traditional Assessments

| Traditional       | Authentic       |
|-------------------|-----------------|
| Selecting a Response | Performing a Task |
| Contrived         | Real-life       |
| Recall/Recognition | Construction/Application |
| Teacher-structured | Student-structured |
| Indirect Evidence  | Direct Evidence  |

Pophan state, authentic assessment is engaging and worthy problems or questions of importance which students must use knowledge to fashion performances effectively and creatively [5].

Thus, authentic assessment is a form of assessment that requires learners to display attitudes, using knowledge and skills gained from learning in performing tasks in real situations. In other words, the authentic assessment is an approach, procedure, and instrument of process assessment and learning achievement of learners in the application of attitude (spiritual and social), knowledge, and skills acquired in the form of assignment of real behavior or behavior with similarity level with the real world, or learning independence [6].
2. Research Methods
This research was conducted at the Graduate School Program, State University of Gorontalo for authentic assessment tool development and at SMAN 1 Kwandang for the tests of development kit authentic assessment product.

The method used in this research is the research development that is developing the authentic assessment tools of Physics Learning in Senior High School. In addition, these methods have to be able to achieve the goals and objectives of the research. The goal and objective is to facilitate and guide master students of prospective Master of Physical Education in completing their thesis that related to the problem of Development Authentic assessment tools in the areas of cognitive, affective, and psychomotor in physics learning to be used as a guideline assessment thorough, objective, and fair.

The procedure of this study refers to the development model by Thiagarajan, Semmel, and Semmel known as 4-D (four-D models) that are define, design, develop, and disseminate [7].

3. Results and Discussion

3.1. The Definition Phase
Generally, in this definition phase, analysis of development needs activities is performed, requirements development of products that fit the needs of users, as well as models of research and development that is suitable to develop products. This stage has five steps, namely front-end analysis, learner analysis, task analysis, concept analysis, and the formulation of special purpose assessment. The step in the process of defining this activity is shown as follows.

(1) Front End Analysis
At this stage, researchers examined the curriculum of 2013. There are core competencies and the basic competency to be achieved in the curriculum. Curriculum analysis is useful to define the competencies where such assessments are developed. Based on the analysis of this curriculum, an authentic assessment tool for Odd Semester of 1st grade in Senior High school is developed.

(2) Analysis of Students’ Characteristics
Analysis of the learners’ characteristics includes individual academic ability, physical characteristics, the ability of work group, learning motivation, economic and social background, and previous learning experience.

(3) Task Analysis,
The task analysis in the development of assessment tools includes analysis of tasks (basic competencies) that must be mastered by learners in learning physics, so that learners can achieve minimal competence.

(4) Concept analysis
Analysis of concepts/materials is performed by identifying the concepts/main materials that need assessment, collect and select relevant material, and reconstruct systematically in the form of concept maps that incorporate the concept/content and assessment of physics learning in odd semester of 1st grade at Senior High School.

(5) Analysis of the Assessment Objective
In this analysis, the assessment objectives and competencies to be achieved need to be formulated in advance. It is useful to restrict researchers in order not to deviate from the purpose of the assessment. The formulation of assessment purpose is based on analysis of the concept and analysis of tasks, thus becoming operational and expressed with behavior that can be observed.

3.2. Design Stage
The purpose of design is to prepare the prototype of authentic assessment tools in physics learning. This stage includes: (a) determine the core competencies and the basic competency, as the first act to
be a reference for teachers in developing techniques and instruments in the authentic assessment of learning physics; (b) arrange lattice that contains indicators and assessment strategies; (c) selecting methods and assessment techniques, as well as the selection of the assessment instruments used; and (d) develop guidelines for scoring. In addition, this stage focused on writing and adoption of assessment tools, consultation with experts, and discussions with colleagues.

Analysis of the design objectives undertaken to formulate the goals to be achieved. In this study, the analysis of the design objectives is based on results of field surveys and school physics learning objectives related to the subject matter of physics that are considered in accordance with the characteristics of the affective domain that are developed. Analysis of interest also adapted to SK Dirjen Mandikdasmen No.12/C/Kep/TU/2008 on the form and procedure for preparation of the learning outcomes of students unit of elementary and secondary education, which states that the affective aspects are dominant in the subjects of Mathematics, Physics, Chemistry and Biology includes thoroughness, perseverance, and ability to solve problems logically and systematically. Analysis of core competencies, basic competencies and learning objectives on Physic for 1st grade of Senior High School in the affective domain can be seen in the Table 2 [9].

| Developing honest behavior, discipline, responsibility, caring, polite, friendly, caring environment, mutual aid, cooperation, peace-loving, responsive and proactive, and demonstrate behaviors as part of a solution to the various problems of the nation in interacting effectively with the social environment and nature and put themselves as a reflection of the nation in the association world | 1. Shows a behavioral science (curious, objective, fair, thorough, meticulous, diligent, careful - careful, responsible, open, critical, creative, innovative, and caring environment) in the daily activities - day as a form of implementation of the attitude of doing trial and discussion
2. appreciates the work of individuals and groups in the daily activities - day as a form of implementation |

Table 2. Core Competence and Basic Competence Physics In Senior High School

| Core Competence | Basic Competence |
|-----------------|------------------|
| Developing honest behavior | 1. Shows a behavioral科学 (curious, objective, fair, thorough, meticulous, diligent, careful - careful, responsible, open, critical, creative, innovative, and caring environment) in the daily activities - day as a form of implementation of the attitude of doing trial and discussion |
| Discipline, responsibility, caring, polite, friendly, caring environment | 2. Appreciates the work of individuals and groups in the daily activities - day as a form of implementation |
| Mutual aid, cooperation, peace-loving, responsive and proactive | |
| Demonstrate behaviors as part of a solution to the various problems of the nation in interacting effectively with the social environment and nature and put themselves as a reflection of the nation in the association world | |

For the authenticity of the authentic device ratings in the cognitive domain validated by 4 expert and user validators can be seen in the Table 3.

Table 3. Recapitulation of Scores of Results From Expert and User Experts on Devices Cognitive Area Assessment

| Skor total (%) | Percentage average score (%) |
|----------------|-----------------------------|
| Expert 1 | 90.94 |
| Expert 2 | 92.59 |
| Expert 3 | 95.53 |
| Expert 4 | 94.23 |
| User 1 | 87.69 |
| User 2 | 87.69 |
| User 3 | 80.00 |
| User 4 | 80.00 |
| Total | 88.97 |
| Average | 83.85 |

Based on the result of the expert validator assessment it was found that the authentic assessment tools of the multiple choice form and the developed descriptions were valid based on the assessment by four experts gained the average percentage score of 93.32% with valid qualification. Similarly, based on the results of the questionnaire, it was found that the authentic choice tools of multiple choice
form and the description developed practically based on the assessment by four senior teachers as the users obtained an average percentage score of 83.85% with excellent qualification.

When the expert judgment results in the conversion in the form of percentage then the result count is summarized in the Table 4.

**Table 4. Percentage of Affective Assessment Device By Expert Validator**

| Eligibility Criteria (%) | Eligibility Feasibility |
|--------------------------|-------------------------|
| Graphic Design and Layout | 83.33                   |
| Construction Instrument  | 84.38                   |
| Language                 | 87.50                   |
| Content                  | 90.63                   |
| Feasibility Percentage   | 86.46                   |

The feasibility of the developed affective appraisal instrument is measured by four main indicators includes the feasibility of graphic design, construction, language and content feasibility. Each of these indicators is described with sub indicators. The feasibility of graphic design is obtained a percentage of 83.33%, which is supported by the clarity of letters, layouts and the location of sentences and words in the paper. The feasibility of the construction is obtained score of 84.38%, in this indicator is supported by three sub indicators includes the suitability of the curriculum, clarity division of indicators, conformity of the statement with indicators and conformity indicators with aspects measured. The language feasibility indicator, 87.5% is obtained, in the language eligibility, which is an assessment item is the use of the correct language code, language simplicity, ease of interpretation and full meaning. Based on the percentage obtained, then this instrument is considered feasible in terms of the language used. The last indicator i.e the content feasibility, this indicator becomes the main indicator in validation of instrument development. It gets percentage of 90.6%. This value indicates the feasibility of the instrument is high enough. The average feasibility of this instrument is 86.46% with this percentage, then the developed valuation instrument is said to be feasible to use.

Furthermore, for authentic assessment tools in the form of rating performance and assessment project. Results from expert validators are then analyzed descriptively by using 5 aiken and alpha cronbach analysis. Based on the formula 5 Aiken and alpha cronbach it can be determined the feasibility of the product assessment instrument assessment and project appraisal developed can be seen in the Table 5.

**Table 5.** Feasibility Of Authentic Assessment Tool Of Psychomotor Area Assessment Reliability

| Assessment | Validity | Reliability |
|------------|----------|-------------|
| Performance| 0.79     | Interpretaion | Criteria | r | Interpertasi | Description |
| Project    | 0.81     | excellent   | 0.80     | Very high | Reliabel    |

3.3. Development Stage

At this stage, the two activities are performed, namely: expert appraisal and developmental testing. Expert appraisal is a technique to validate or assess the feasibility of product design. While developmental testing an experimental activities in the product design targets the actual subject. The measures undertaken at this stage as follows: (1). Validation of the assessment by expert/specialist; (2) Revised assessment tools based on input from experts at the time of validation; (3) The test is limited
in the classroom; (4) Revised assessment tools based on trial results; and (5). Implementation of votes in the wider region.

Before authentic assessment tools in the affective domain tested in the wider region, it first has to do research on the instrument. The study carried out by two teams of experts that is a lecturer of Physics, Gorontalo State University.

1. **Limited Test**
   - Limited testing is conducted to see how effective the first draft of authentic assessment tools in the affective domain. This limited test is conducted on 30 students in School SMA Negeri 1 Kwandang Gorontalo Utara as respondents.

2. **Revised of Limited Draft Test**
   - The next trial draft instrument revisited in point of not good statement. The draft further improved based on the factors that lead to the invalidity of the item statement. After a further revised draft instrument trials conducted in the wider region.

3. **Field Test**
   - Field test is performed to see how effective the draft affective assessment instruments [10]. The trial was conducted on 30 students of SMA Negeri 1 Kwandang as respondents. Analysis of the validity and reliability of affective assessment instruments that have been tested to show that the instrument developed an effective and feasible for use. Here are the results of analysis of the validity of the draft field test instrument.

| Statement number | $r_{\text{arithmetic}}$ | $r_{\text{critical}}$ | Criteria |
|------------------|--------------------------|------------------------|----------|
| 1, 2, 3, 4, 5, 6, 10, 11, 13, 15, 19, 21, 22, 23, 24, 28, 31, 34, 35, 36, 38, 39, 41, 44, 45, 46, 49, 50, 51, 52 | 0.36 | 0.36 | Valid |
| 8, 9, 12, 16, 20, 27, 29, 43, 48. | 0.40 | 0.36 | Valid |
| 7, 14, 17, 18, 30, 33, 37, 40, 42, 47 | 0.53 | 0.36 | Valid |
| 25, 26, 32 | 0.61 | 0.36 | Valid |

4. **Product Enhancements**
   - At this stage the instrument was completed by adding a section in the form of theoretical basis, instructions for use, and usage guidelines. The products have been completed subsequently assessed by the subject teachers of physics in school. Teacher assessment results of the obtained results that the instrument developed is good and feasible for use in schools.

3.4. **Dissemination Stage**
   - Dissemination phase is implemented in the form of workshops by involving Physics teachers incorporated in Subject Teacher Council and related institutions, as well as product development of the authentic assessment is published in scientific journals nationally and internationally accredited.

4. **Discussion**
   - Development The trial of Draft 1 affective assessment tools tested on a small group consist of 30 respondents who were students of the School SMA Negeri 1 Kwandang. The results of early trials were then analyzed its validity. Results of the analysis showed that 84.18% grain instruments are
stated to have good validity for $r_{\text{arithmetic}} > 12.38$, and $15.20\%$ while the grains have poor validity for $r_{\text{arithmetic}} < 0.36$. Therefore the results of early trials showed overall pretty good point statement, then performed a field test.

Field test was conducted on 39 students of SMA Negeri 1 Gorontalo and 40 students of SMA Negeri 3 Gorontalo. The result of field test showed that the whole item of insrument has good validity with good percentage $10.60\%$ and $89.40\%$ very good. The reference used to determine the quality of the instrument is the guidance of the preparation of the device by the Directorate of High School Development. Based on the results of the analysis on limited trials and extensive trials, the overall grain of cognitive, affective and psychomotor assessment instruments developed is effective for use.

5. Conclusion

Based The High School authentic assessment tools in the cognitive, affective, and psychomotor domains are feasible, effective, and practical used to measure the learning outcomes of Class X Physics High School. The criteria are feasible, based on the assessment and examination of the expert validator shown by the percentage of the average score of eligibility $91.35\%$ and reliability of $0.82$, as well as the percentage of the average score of $83.85\%$ practicality, and the learning result completeness of $80\%$.

Acknowledgement

This article is based on a research scheme funded by the State Budget of 2016-2017 It explores “develop authentic assessment in science physics learning”. The researchers extend their sincere gratitude to (1) Gorontalo state university and (2) Directorate General for Sience and Technology Institution and Higher Education for facilitating this research, as well as to (3) SMA Negeri 1 Kwandang, SMA Negeri 1 Gorontalo, and SMA Negeri 3 Gorontalo for granting the permission to conduct the research on the site.

References

[1] Supardi 2015 Authentic Assessment, Afective, cognitive and Psychomotor Learning (Penilaian Autentik. Pembelajaran Afektif, Kognitif dan Psikomotor. Konsep dan Aplikasi). Jakarta: PT Rajagrafindo Persada 1 48-49.
[2] Taruh, Enos 2012 Pemetaan dan Peningkatan Mutu Karya Tulis Mahasiswa Fakultas MIPA Universitas Negeri Gorontalo. (Penelitian Hibah Bersaing) Laporan Penelitian. Gorontalo: Lemlit UNG (2012) 147-148.
[3] Fogarty Robin. 1991. How to Integrate the Curricula. Illinois: IRI/Skylight Publishing, Inc. (1991) 134-135.
[4] Mueller J 2005. The Authentic Assessment Toolbox: Enhancing Student Learning Through Online Faculty Development. Journal of Online Learning and Teaching, 1(1) 11-17.
[5] Pophan W J 1973 Criterion Referenced Measurement. Englewood Cliffs. New Jersey: Educational Technology Publications, 3 114-115.
[6] Gulikers J T M, Bastiaens, Th J, & Kirschner, P. A. 2006. Authenti assessment student an teacher perceptions: the practica valu of the fiv dimensional framework. Journal of Vocational Education and Training, 58 337-357.
[7] Thiagarajan, S Semmel, DS Semmel, M 1974. Instructional Development for Training Teacher of Exceptional Children. A Source Book. Blomington: Central for Inovation on Teaching the Handicapped 112-113.
[8] Kubiszym T and Borich G 2013 Educational Testing and Measurement Classroom Aplication and Practice. New York: John Willey & Sons, 6 201-202.
[9] Permen Dikbud No 66 2012 Standar Penilaian Pendidikan dasar dan Menengah, 49.
[10] Arsyad N Rahman A and Ahmar A S, 2017 Developing a self-learning model based on open-ended questions to increase the students’ creativity in calculus Glob. J. Eng. Educ. 19, 2.