Meta-analysis of authentic assessment instrument development to measure learning outcomes of learners SMA/MA

Imelda Afriana and Festiyed*
Department of Physics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Padang, Jalan Prof Hamka, Padang, 25131, Indonesia

*festiyed@gmail.com

Abstract. This Meta-analysis aims to summarize the research results related to the development of authentic assessment instruments in measuring the learning outcomes that have been done before. Based on the analysis from the background of several journals that have been selected, the biggest problem lies in the less active learning process from the communication to the assumption of the learners on the learning of physics that only study Count with definite results without knowing the usefulness of science in technology development. In addition, the teacher's difficulties in creating, executing, and analyzing authentic assessments led to the teacher being more likely to assess cognitive aspect assessment and override other aspects of judgment. Based on the problem, one of them is to develop an authentic assessment instrument that able to measure learning outcomes according to the needs of learners. Summary of meta-analyses using a descriptive survey method. Based on the analysis done, the assessment aspect of attitudes and skills is more dominant than the cognitive assessment aspect. In addition, analysis results show that not all researchers have tested the level of reliability, effectiveness, practicality, assessment implementation, and student response. It depends on the development method used. Based on all the final results of the journal, the instrument is developed in criteria worthy of use in the school in measuring the outcome of SMA/MA physics.

1. Introduction
Assessment is one of the activities undertaken to measure and assess the achievement level of a curriculum and the process of learning[1]. The assessment of the 2013 curriculum is based on the education assessment standards aimed at ensuring the assessment planning of learners based on the principles of judgment and in accordance with the competencies that will be achieved, the implementation of Student assessments are professionally, educative, open, effective, efficient and in accordance with social and cultural contexts, and learners learning outcomes are reported objectively, informative and accountable[2]. The assessment of the 2013 curriculum emphasizes on authentic assessments.

Authentic assessment is an important component of educational reform since the 1990s[3]. An authentic appraisal is an assessment that engages learners into an authentic, useful and meaningful task. Authentic assessments evaluate the skills that learners have to make the real world an object. An authentic assessment of learners learns how to apply the knowledge and know-how he has in authentic tasks. An authentic assessment should consist of three mutually sustainable and balanced domains that
include three domains, namely cognitive (knowledge), affective (attitude), and psychomotor (skill). Authentic assessment arrangements can be made by your own teacher, team teacher, or involve learners in their arrangement\[4\].

Although authentic assessments include one of the most important categories of learning, there are still many educators who have not carried out an authentic assessment to the fullest. Based on the analysis that has been done in several journals discussing the development of authentic assessment instruments, 25.92% of the problems lie in the less active learners in the learning process, from the interactive teacher and learners to the thought of learners that physics is a science that only performs definite calculations without knowing that physics can be useful in technological developments. In addition, teachers have difficulty in arranging assessment instruments in accordance with authentic judgment with a 22.22% percentage. The percentage of 11.11% lies in the lack of time required in conducting and analyzing assessment instruments making some authentic judgments gained from the instantaneous view of educators to complete school assignments without any use. The tendency of teachers who only conduct cognitive assessment and override other aspects has a percentage of 11.11%. Similarly, the teacher's desire and interest in developing an authentic assessment instrument has a 11.11%. Other problems such as less the understanding of learning models, applying to some skills, and other problems have a percentage of 18.83%.

One solution that can be used in addressing the problem is to develop an authentic assessment instrument. Develop an authentic assessment instrument capable of producing a product that can measure what is to be measured, both from the realm of knowledge, attitude, and ability. In addition to developing authentic assessment instruments can also train educators to understand the creation of authentic assessment instruments and analyse the learning outcomes that have been obtained, so the assessment is able to describe as much as possible development that happens to learners. The importance of authentic instrument development makes it attractive enough for researchers who want to develop assessment instruments to gauge the specific capabilities of learners.

Various research conducted on the development of authentic assessment instruments can make it easier to obtain relevant references and research when other prospective researchers are interested in the authentic assessment of measuring results Learn learners. To facilitate prospective researchers and readers in gaining knowledge on the development of authentic assessments, there are two methods that can be used to create a summary of the research that contains various types of instrument development. An authentic assessment of meta-ethnographic and analytical methods \[5\]. The analysis Meta method is a form of a summary with qualitative data that presents statistical research data. While the ethnographic meta method is a method that presents a summary of the qualitative data. The method of analysis meta is considered more objective so that the results and data obtained obtained more accurate and quality. The purpose of this meta analysis is to provide a methodology equation in a literature review in experimental research \[6\].

Based on the explanation above, researchers are interested in conducting research with the analysis meta method. This study was conducted with the aim to analyse the development of authentic assessment instruments to measure the outcomes of learners ' physics in SMA/MA. Hopefully research can be helpful and beneficial to prospective researchers and readers in developing authentic assessment instruments to measure the learning outcomes of learners physics.

2. Research Methods
The research method used in this research is a descriptive method of surveying. The population consists of a journal published in 2015-2018 on the development of authentic assessment instruments. The samples taken in this study consisted of 10 research student physics education from several universities in Indonesia. Samples are taken homogeneously that share the same characteristics so that they can be further researched. Samples consist of various types of journals, ranging from international, national, or seminar proceedings. Journals are downloaded from a variety of sources, ranging from the college journal Publication website, Google Scholar and other journal publication websites.
The research procedure in the adaptation of the steps in conducting the meta analysis is: (1) to determine the problems to be researched, (2) Determining the time span of the journals and articles to be researched, (3) looking for various journals and research articles in connection with the issues to be researched, (4) read the title and abstract of the journals and articles to determine the suitability of the problem with the topic of the problem chosen, (5) Focus the research on the Problem, research methods Used, data analysis, and research results, (6) categorical each journal and article in accordance with the paradigm, (7), comparing the entire research results of journals and articles based on predefined categories, (8) Analyzing the entire conclusion of the journal and article studies to determine the strengths and weaknesses of each study, and (9) Draw conclusions [5]

3. Result
The following is a type of development, the model of development chosen, as well as the publication year of the journal.

| Journal Number | Title                                                                 | Types Of Development               | Development Model                       | Publication Year |
|----------------|----------------------------------------------------------------------|-------------------------------------|-----------------------------------------|------------------|
| 1              | Authentic Assessment Instrument Development To Measure The Social     | Research and Development (R&D)     | ADDIE (Analyze, Design, Development,    | 2015             |
|                | Attitude Of Learners SMA Class X On Hysical Learning[7]               |                                     | Implementation and Evaluation).         |                  |
| 2              | Authentic Instrument Development Of Performance Based Assessment In  | Research and Development (R&D)     | Borg & Gall’s                           | 2015             |
|                | Physics Learning In Sub Point Discussion Of Grade Xi Students SMA   |                                     |                                         |                  |
|                | Negeri 1 Taman[8]                                                    |                                     |                                         |                  |
| 3              | Development Of Authentic Assessment Instruments Based On Scientific  | Research and Development (R&D)     | 4D(Define, Design, Develop, and        | 2016             |
|                | Literacy In The Study Of Physics In SMA As Curriculum Implementation |                                     | Disseminate)                            |                  |
|                | 2013[9]                                                              |                                     |                                         |                  |
| 4              | Development Of Authentic Instrument Assessment To Measure Higher     | Research and Development (R&D)     | Borg & Gall’s                           | 2016             |
|                | Order Thinking Skills Learners[3]                                    |                                     |                                         |                  |
| 5              | Instrument Development To Measure The Science Communication Skills   | Research and Development (R&D)     | 4D(Define, Design, Develop, and        | 2016             |
|                | Of High School Students[10]                                          |                                     | Disseminate)                            |                  |
| 6              | Development Of Project Valuation Instruments Based On E Portfolio   | Research and Development (R&D)     | ADDIE (Analyze, Design, Development,    | 2017             |
|                | Temperature, Heat And Transfer In High School[11]                   |                                     | Implementation and Evaluation).         |                  |
| 7              | The Development Of Authentic Assessment Instrument To Measure Science | Research and Development (R&D)     | Borg & Gall’s                           | 2018             |
|                | Process Skill And Achievement Based On Students’ Performance[12]     |                                     |                                         |                  |
| 8              | Student Performance Assessment Instrument Development Based On        | Research and Development (R&D)     | Borg & Gall’s                           | 2018             |
|                | Scientific Approaches To Grade X                                     |                                     |                                         |                  |
Based on the above data is obtained that all types of development used are Research and Development (R&D). Based on the analysis conducted there are several types of development models used in developing assessment instruments namely ADDIE (Analyze, Design, Development, Implementation and Evaluation), 4D (Define, Design, Develop, and Disseminate), Borg & Gallu's, and Descriptive Qualitative Research each have different steps in accordance with the supply and solution chosen so that the instrument developed can be measured. The selected journal has a vulnerable publication time from 2015-2018. Here is an analysis based on the background.

Table 2. The Instrument of Assessment Developed

| No | The instrument of assessment developed | Assessment techniques | Journal |
|----|----------------------------------------|-----------------------|---------|
| 1  | Knowledge                              | Written Multiple Choice Tests | √       |
| 2  | Attitude                               | Oral Tests            | √       |
|    |                                        | Assignment            | √       |
| 3  | Skills                                 | Observation           | √       |
|    |                                        | Self-Assessment       | √       |
|    |                                        | Peer Assessment       | √       |
|    |                                        | Journal Interview     | √       |
|    |                                        | Performance Project   | √       |
|    |                                        | Potopolio Products    | √       |

Source: Secondary Data 2019

Based on the table above, there are some assessment techniques done in conducting authentic assessments. The knowledge sphere consists of three assessment techniques: Written tests, oral tests, and assignments. Written tests can be done in a variety of ways, namely multiple choice, short answer, stuffing, correct, wrong, and description. The test chooses the correct answer, short stuffing, is an assessment technique that only able to assess the ability of thinking at low levels, namely only remember or memorize, but still can use the material terms and demands Competence in accordance with these
tests [5]. Based on the data above it appears that in the knowledge assessment, the more dominant technique used is the description. The description can assess the different types of skills that learners have, such as critical thinking, creative, and solving problems. The unraveling test is also able to present a free answer to the measurement of skills and ideas for learners or the arithmetic of specific learning materials or concepts such as those found in mathematical learning, Physics, Chemistry, and others in writing [17]. In addition, assignments are also suitable for physics learning because it can help learners deepen the mastery of knowledge learned in the learning process [16].

Attitude assessment has several techniques, namely observation, self-assessment, peer assessment, journal and interviews. Based on the analysis that has been done in the assessment of self-assessment attitude is more dominant than other assessment techniques. Self-assessment is an assessment technique by how learners assess themselves with the shortcomings and advantages that are held in the constraints of achieving spiritual, social attitude competency. Self-assessment has several positive impacts, including providing confidence to learners, realizing the pros and cons they have, and can train learners to be honest, because in student self-assessments are required to be honest in conducting assessments. The observation technique is also able to assess the attitude of learners, so teachers can observe the learners under their construction [17].

Based on the data analysis above, there are several techniques that can be done in obtaining the value of skills, namely through performance techniques, projects, products, and Portofolio. The results of analysts showed that performance assessment techniques were more dominant than other assessment techniques. Performance appraisal emphasizes the learner's process in implementing the knowledge he has. It is certainly very suitable in the physics learning process so that educators are able to know the extent of learners understanding of the material that has been studied. Instruments worthy of use in the learning process. Here are some of the analyses that are conducted to determine deserving or not a development of the instrument.

| Journal | Validity | Reliability | Effectiveness | Practicalities | Assessment Implementation | Student Response | Category |
|---------|----------|-------------|---------------|----------------|---------------------------|------------------|----------|
| 1       | 3,39     | 0,96        | -             | -              | 3,39                      | 3,43             | High, Perfect, Good, Very Positive |
| 2       | 4,23     | -           | -             | -              | 3,78                      | -                | Very Good, Very Good. |
| 3       | 3,48     | 0,46        | -             | -              | -                         | 3,65             | Very good, quite, very good. |
| 4       | 84,85%   | -           | -             | -              | -                         | -                | Excellent. |
| 5       | 4,10     | 0,88        | 3,72          | -              | -                         | -                | Very Good, High, Good. |
| 6       | 4,37     | 0,783       | -             | 82,37%         | -                         | -                | Very Valid, High, Very Strong. |
| 7       | 0,89     | 0,705       | -             | -              | -                         | -                | Very High, High. |
| 8       | 84,06    | 0,896       | -             | -              | -                         | -                | Very Good, Very High. |
| 9       | 83,33%   | 0,86        | 88,19%        | 86,83%         | 3,47                      | -                | Very Good, High, Very High, Very High. |
| 10      | 0,47     | 0,67        | -             | -              | -                         | -                | Quite, High. |

Source: Secondary Data 2019
Based on the above data, it is seen that not all aspects are done by researchers. Although some journals have the same development model, the resulting analyses are not the same. The entire journal has a validity value with criteria ranging from enough, good, high, very good, and very valid. The difference of this category is because the foundation of the theory is used differently so that the criteria used differ. The level of reliability on the entire development of the instrument has a fairly high, high, very high and flawless criterion. Based on 10 journals analyzed, only two journals perform effectiveness with the effectiveness rate of 3.72 with good criteria, and 88.19% in percent form with high criteria. Practicality is only done by two researchers with very strong criteria with percent 82.37%, and the criteria is very high with a commissioner of 86.83%. The implementation of an authentic assessment was carried out by two researchers with a value of 3.39 on good criteria, as well as 3.78 with excellent criteria. learners’ response to the development of authentic assessment instruments was conducted by 3 researchers, with grades 3.43, 3.65, and 3.47 with very positive, and excellent criteria. It can be concluded that the development of authentic assessment instruments from several journal analyses can be feasible and usable.

4. Discussion
A learning process has reached the goal of learning or not be known from the results of the assessment obtained. Assessment in the learning process plays a very important role for educators in choosing the action to take if learning objectives have not been achieved with maximum. In the 2013 curriculum the assessment is not only assessed in terms of training results or tests of learners, but the assessment on the 2013 curriculum emphasizes more on authentic (real) assessments that measure the realm of knowledge, attitudes, and skills of the participants learners.

An authentic assessment is a meaningful assessment where learners can learn by application of science that he/she has by doing tasks that are authentic. An authentic assessment has strong relevance to the scientific approach of learning in accordance with the 2013 curriculum. This assessment is able to provide an overview of developments that occur to learners, both in observing activities, reasoning,, trying, building the network, and others. Authentic assessments can also be applied to certain fields of science, such as art and science in general, with a major review of the process and learning outcomes.

Although an authentic assessment plays an important role in the learning process, based on the analysis done on 10 journals, there are still many educators who have difficulty in drafting assessment, both cognitive, affective and psychomotor. In addition, educators are still based on the cognitive outcomes of the tests the learners work on. The affective and psychomotor assessments whose applicability are difficult to observe, take a long time in assessing and analyzing the results of the assessment to be the reason why affective and psychomotor assessments are uncommon, other causes such as less An understanding of the learning model and still using a long learning model makes the authentic Judgment unapplied. For this it is necessary to develop authentic assessment development to facilitate educators in measuring student learning outcomes so that all learners development can be invaluable to the fullest extent possible.

Based on the analysis done on 10 journals on the development of authentic assessment instruments, the type of development used to develop an authentic assessment instrument is Research and Development (R&D) with Some of the development models used are ADDIE (Analyze, Design, Development, Implementation and Evaluation), 4D (Define, Design, Develop, and Disseminate), Borg & View’s, and Descriptive Qualitative Research with vulnerable publication time from 2015 to 2018. Assessment aspects of attitudes and skills are more dominant than the cognitive assessment aspect. In addition, analysis results show that not all researchers have tested the level of reliability, effectiveness, practicality, assessment implementation, and student response. It depends on the development method used. Based on all the final results of the journal, the instrument is developed in criteria worthy of use in the school in measuring the outcome of SMA/MA physics.
5. Conclusion
Based on the results of the analysis and the discussion that has been done in 10 selected journals using a summary of the survey method meta-analyst is descriptive, with the same type of development, different development methods produce a different analysis of instruments. All journals carry out the validity test according to the instruments that are developed and have the criteria ranging from enough to very good, and very high. However, it is not the whole journal of Reliability, effectiveness, practicality, assessment and responsiveness of learners. It depends on the development methods used by researchers. Overall based on the instrument analysis that has been done by each researcher that the development of authentic assessment instruments can be used to measure the outcomes of learners’ physics of knowledge, attitudes, and skills.

6. Advice
Advice for readers, when interested in doing similar research, this research is only sourced from analysis of 10 journals, need to expand the reference so that the data obtained more complete and accurate. In addition to the development of assessment instruments that are important in the learning process, many other innovations are expected to develop authentic assessment instruments to help educators and learners in the process Learning and able to improve the quality of teaching.

Reference
[1] Ma'aruf., Lenny Rahim, Andi. 2015. The Development Of Authentic Assessment Devices In Physics Learning To Improve Motivation, Participation And Outcomes Of Learning Physics For Grade X Students In SMA Negeri 1 Pangkajene. Physics Education of Muhammadiyah University of Makassar. JPF | Volume I, number 3. 252-263
[2] Lindayani., Siregar, Antoni., Yuberti. 2016. Pengembangan Instrumen Authentic Assessment Untuk Mengukur Higher Order Thinking Skills Peserta Didik. National Seminar on Education, Physics Education Study Program. IAIN Bandarlampung. 140-149
[3] Festiyed. 2016 Perubahan Paradigma Pendidikan: Peningkatan Layanan Profesional Melalui Pembelajaran Autentik dan Asesmen autentik. National Seminar and Forum State of LPTK graduate in Indonesia: State University of Gorontalo
[4] Hosnan. 2014. Pendekatan Saintifik dan Konseptual dalam Pembelajaran Abad 21. (Bogor: Ghalia Indonesia)
[5] Rasti, Putri Ramadhani, et al. 2018. Meta Analisis Pengaruh Bahan Ajar Terhadap Kompetensi Fisika Peserta Didik. Prosiding PDS UNP. 201-207
[6] Surwanti Rima., Tiara et al 2018 Meta-Analisis Pengaruh Lembar Kerja Siswa Terhadap Peningkatan Hasil Belajar Fisika SMA. Prosiding PDS UNP. 227-233
[7] Nur Anggraheni, Nadya., Suryono., Ngafifah, Nur. 2015. Pengembangan Instrumen Penilaian Autentik Untuk Mengukur Sikap Sosial Peserta Didik SMA Kelas X Pada Pembelajaran Fisika. Jurnal Radiasi vol.7 No.2. 1-6
[8] Nini Yuliana, Putri., Arief, Alimufi. 2015. Pengembangan Instrumen Authentic Assessment Berbasis Kinerja Dalam Pembelajaran Fisika Pada Sub Pokok Bahasan Titik Berat Siswa Kelas XI SMA Negeri 1 Taman. JIPF Vol.4 No. 1. 47-52
[9] Widowati, Tutut., Siti Aminah, Nonoh., Cari. 2016. Pengembangan Instrumen Penilaian Otentik Berbasis Scientific Literacy Pada Pembelajaran Fisika Di SMA Sebagai Implementasi Kurikulum 2013. Jurnal Inkuiri vol. 5 No.2. 8-19
[10] Kartika, Diana., Srijono., Ngazizah, Nur. 2016. Pengembangan Instrumen Untuk Mengukur Kemampuan Komunikasi Sains Siswa SMA. Jurnal Radiasi Vol. 08 No.1. 28-32
[11] Agustin, Dewinta., Djoko Lesmono, Albertus., Subiki. 2017. Pengembangan Instrumen Penilaian Proyek Berbasis E Portofolio Pada Pokok Bahasan Suhu, Kalor Dan Perpindahannya di SMA. National Seminar on Physics education 2017. University of Jember. Vol 2. 1-6
[12] Kurniawati, Atin., Sukardiyono. 2018. The Development Of Authentic Assessment Instrument To Measure Science Process Skill And Achievement Based On Students' Performance. *JPPPF*. Vol.2 Issue. 2. 65-74

[13] Severina, Vina., Aris Santoso, MM., Rahmat, Cecep. 2018. Pengembangan Instrumen Penilaian Kinerja Peserta Didik Berbasis Pendekatan Saintifik Pada Praktikum Pengukuran Kelas X SMA. *JRKPF UAD*. Vol.5 No. 2. 87-94

[14] Rabiudin, Taruh ,Enos., Mursalin. 2018. Development of Authentic Affective Assessment Instrument in High School Physics Learning *IOP Conf. Series: Journal Of Physic: Corf Series*. 1028. 1-7

[15] Sugiarti, T., Kaniawati, I., Aviyanti, L. 2017 Development of Assessment Instrument of Critical Thinking in Physics at Senior High School. *IOP Conf. Series: Journal Of Physic: Corf Series*. 812. 1-9

[16] Kunandar. 2013. *Penilaian Autentetik (Penilaian Hasil Belajar Peserta Didik Berdasarkan Kurikulum 2013)*. (Jakarta: Rajawali Pers)