Authentic-Peer Assessment Instrument To Measure The Ability Of A Chemistry Teacher Candidate Evaluation On Basic Chemistry Small Skill Laboratory Works

Fachri Hakim¹, Muhammad Zammi²

¹,² Department of Chemistry Education, State Islamic University Walisongo Semarang,
Email: fachrihakim@walisongo.ac.id

Abstract. The dominance of paper and pencil assessment tests in chemistry small skill laboratory works make assessments in the affective and psychomotor less attention. This research aims to develop an authentic-peer assessment instrument to measure the ability of a prospective evaluation of chemistry teachers on basic chemistry small skill laboratory works. The method used research and development. The subjects were students of the second semester of chemistry education 2018 at Walisongo State Islamic University Semarang. Data collection instrument in the form of sheets validation instrument to measure the validity and reliability of peer assessment and questionnaire responses sheet learners. The results of authentic-peer assessment developed very validly in the category with a value of 31. The internal consistency reliability Spearman-Brown gained 0.78 means indicates that the authentic assessment instruments developed have a strong level of consistency. The measurement results evaluation capabilities candidate chemicals are known educator excellent ability group with a percentage of 31.45%, a good ability with the percentage of 47.48%, enough ability with the percentage of 15.79%, and less ability with the percentage of 5.28%. The questionnaire results obtained by 82 percentage points. The results of that study concluded that authentic assessment developed meet the category of valid, reliable, and practical.

Keywords: Authentic assessment, peer assessment, small skill laboratory works

1. Introduction
Evaluation as part of the overall evaluation process makes the lecture should be part of the competencies required of educators are able to evaluate correctly their students well. As stipulated in Government Regulation No.19 of 2005 on National Education Standards are set to carry out the mandate of Law No. 14 of 2005 on teachers and lecturers, which states that a teacher / lecturer should have 4 competencies are: pedagogical, personality, social, and professional. Therefore, the national education minister regulation No.16 of 2007 on standard academic qualifications and competence of the teacher considers it important for teachers to have the ability in conducting the evaluation, operation process,
Stanford University developed teaching ability is known as STCAG (Stanford Teacher Competence Appraisal Guide). Teaching abilities are classified into 4 groups include: 1) plan instruction, 2) appearance of teaching, 3) evaluate learning outcomes, and 4) professionalism and community [1]. The Government has given the determination of the assessment in education. This refers to the Education Assessment Standards contained in the National Education Minister Regulation No.20 of 2007 on the assessment standards are comprehensive and continuous education. State Islamic University Walisongo Semarang is one of the Education Workforce Education Institutions (LPTK) that play a role in preparing candidates for professional educators, demanded to give lectures to equip prospective teachers become professional educators.

Tuition chemistry lab base at State Islamic University Walisongo, they often found that the scoring system used in college chemistry lab base is still dominated by the assessment paper and pencil tests or assessment are glued to the report, while the performance of learners as well as the self-assessment by learners do not often conducted by a lecturer. When the small skill laboratory works, the students are only required results of lab reports, without ever considering the results of the performance of learners when carrying out small skill laboratory works. The tendency in the field shows that the assessment of learning outcomes is more focused on cognitive aspects [2]. In fact, psychomotor and affective aspects it is very important to assess, without the data collected in the assessment becomes incomplete and meaningless [3].

In the course of chemical in essence a field of science that is not just a collection of facts but also a series of scientific process that requires the activity of acting or hands-on [4]. Measurements carried out on the psychomotor aspects of learning outcomes in the form of appearance [5]. The one way to assess the competence of psychomotor aspect is through direct observation and behavior assessment (performance) of students during the learning activities (practice takes place) [6].

Valuation techniques used may be authentic assessment to monitor the progress of learners directly during practicum and not something contrived [3]. Thus, authentic assessment directed at the process of observing, analyzing, and interpreting the data that has been collected during the practicum takes place and not solely on learning outcomes. Learners can conduct research carefully, think critically, able to solve problems, and is capable of transferring the mastery of science acquired in school into everyday life [7].

Authentic assessment of the psychomotor student can be done through peer assessment techniques. Peer assessment is an assessment conducted by a learner to other learners. In one group, students often get the same award with other students when their contribution to the group is sometimes different. Moreover, it is possible that the problem can sometimes be taken and discussed outside the classroom. To resolve the administration of the same value on each member of the group because of the contribution of each member of different groups can use peer assessment techniques [8].

The used of evaluation peers (peer assessment) in higher education has been advocated by many academics [9]. Peer assessment method has been tested on a different level, across disciplines and with different types of tasks including writing portfolios, oral presentations, test performance, and other skilled behaviors. A colleague or peer assessment can change the attitude of students to be better. Learners become more relaxed, did not find it difficult and awkward when judged by his colleagues and educators [10].

Rate peer assessment has been carried out by McGourty, et al. to develop the ability of engineering students in basic competencies and competency of cooperation, to provide feedback [11]. A peer assessment instrument can be used to evaluate teamwork [12]. A peer assessment instrument in the evaluation of training programs to improve the performance of fellow lecturer at the Institute Shapiro for education and research at Harvard Medical School (HMS) and Beth Israel Deaconess Medical Center (BIDMC) [13]. The importance of pedagogy ability of educators, it is appropriate if the prospective educator practicing responsible for the learning process so that students can be independent and able to train the ability of the evaluation as a preparation to be educators. Authentic assessment include through by peer-assessment techniques.
Based on the above, the importance of evaluating the ability of small skill laboratory works and learning process, it is fitting an educator can evaluate properly. Ability evaluation is not merely a theory, but it is important to be trained since becoming a candidate educator. Based on this background, the purpose of this research is the development of an authentic assessment Instruments State Islamic University Walisongo peer assessment in the basic chemistry lab.

2. Methods
This study is a research and development in the field of education that refers to the 4-D model of the development of the modified 3-D models, they are defined, design, and develop [14]. The subjects were students of the second semester of chemistry education State Islamic University Walisongo Semarang. Data collection instruments used the validation sheet instruments to measure the validity and reliability of peer assessment and questionnaire responses sheet learners. Data were analyzed with descriptive data analysis of qualitative and quantitative. This study was conducted in April 2018 - August 2018 in Semarang Walisongo State Islamic University. The population of this research is chemistry teaching students in the second semester of the academic year 2017/2018. Small class trials conducted on 9 student candidate chemistry teacher. Large classroom trials conducted on the chemistry teacher prospective students the second semester consists of 32 learners that aim to measure the ability of prospective evaluation of chemistry teachers on basic chemistry lab through the use of authentic assessment instruments peer assessment. After the completion of the large-scale test, then proceed with student responses to the use of authentic assessment instruments peer assessment. [15,16].

3. Results and Discussion
Authentic assessment by peer assessment developed in this study refers to the model of the development of 4D consists of four main stages, namely define, design, develop, disseminate [14]. However, in practice the 4-D model developed by S. Thiagarajan. It has been modified by researchers into 3D. That is because the products that have been developed are not distributed and used on a large scale it is only used in the chemical education courses the faculty of science and technology State Islamic University Walisongo Semarang [17]. The third stage is described as follows:

1. Definition phase
The definition phase is the initial stage that must be passed before researchers designing products that will be developed. This stage also a step to bring up the problems in the lecture. The definition phase includes 5 stages: the front end analysis, analysis of learners, material analysis, task analysis, and the formulation of the purpose of college. Steps in the definition phase as follows:

a. Analysis of the front end
Analysis of the front end of a preliminary analysis conducted by researchers in identifying and establishing the basic problems encountered in chemical education courses the faculty of science and technology State Islamic University Walisongo Semarang. Activities are undertaken in the observation process in the form of curriculum analysis and valuation analysis used. Observations obtained in the form of a problem that requires attention, as follows:
- Still lack an authentic assessment by lecturers;
- integration of authentic assessment in the lecture has not received serious attention;
- still fixated on the assessment of the final results of learners in the form of pencil and paper test;
- not many authentic assessment instruments developed by peer assessment for learners;
- lecturers still consider that the instrument of authentic assessment by peer assessment is not practical because it adds a burden to the teachers so the willingness of faculty to implement the instrument in any materials is still lacking;
- Limited knowledge of lecturers to variations in assessment techniques.
- The process of evaluation of the activity in lab inadequate
Assessment of performance that used during the practicum takes place only in the form of sheets of observation and assessment process is still fixated on the assessment the only group and not individually.

Assessment of the performance of students in practical activities is only done by an appraiser that due to the limited ability of the assessor in observing the overall student practicum cause the performance of each sub-optimal assessment.

b. Chemical analysis of a prospective educator

Analysis of the chemistry teacher prospective students to know the characteristics of the candidate of chemistry teachers based on the needs and development as a reference for the design development of authentic assessment peer assessment. The results of analysis are obtained by the students of the department of chemistry teaching semesters 2. According to Vygotsky, in the age range of individual cognitive development in the region of the tasks that are very difficult to overcome by the individual alone, but will be achieved when the guidance or help from an adult or colleague who is more skilled.

c. The task analysis

Based on analysis of learners and analysis of the material, then the task is given in the form of task performance through basic chemistry lab using laboratory methods. The performance task associated with the practicum manufacture of reagents and acid-base titration, which is part of the basic chemistry lab subjects. Its main activity is a task that is not hard to do, the tools/materials used are also easy to reach and able to provide new knowledge to potential chemistry teachers to engage directly perform the basic chemistry lab.

d. The formulation of practical purpose

Based on the analysis of the material and analysis of the above tasks, then formulated the practical achievement indicators. The practical achievement indicators as follows:

| No. | Material Lab | Purpose |
|-----|-------------|---------|
| 1   | Preparation of Reagent Chemicals | Practitioner can make chemical reagents from a crystalline material (solid) Practitioner can make the chemical reagents from a solution (liquid) |
| 2   | Volumetric Analysis: Acid-Base Reactions | The practitioner can perform volumetric analysis using acid-base titration to determine the concentration of acid or base. The practitioner can determine the Ka value of a weak acid |

2. The design phase

The purpose of this phase is to design the form of authentic assessment that the results of peer assessment are referred to as prototype design. Based on the above definition phase, the design of the products developed was described as follows:

a. Preparation of assessment instruments

The preparation of authentic assessment instruments peer assessment in basic chemistry lab requested the assessment to the materials and media experts. Product development authentic assessment instruments are considered valid peer assessment; student deserves to be tested to the education department of chemistry. Students were also given questionnaires about product response authentic assessment instruments that peer assessment.

b. Selection format

The criteria selected researchers in the process so that the target of assessment is when the learners do chores or doing practical activities. The criteria specified to assist in the preparation instructions as well as the use of authentic assessment rubric peer-assessment techniques. Instructions for the use of instruments need to be made because authentic assessments of the performance of basic chemical
practicums developed by researchers use peer assessment techniques or peer assessments that enable lay knowledge of prospective educator students to the assessment process.

c. The initial design of the product
An authentic assessment peer assessment instrument designed by the systematic presentation of the material by the purpose lab practicum conducted. Materials and images are collected from relevant sources. Based on the format of the selection, preliminary design of load cover products authentic assessment peer assessment, instructions for use instrument that consists of general instructions and specific instructions, lattice as a benchmark in making aspect/performance will be judged rubrics, and scoring.

3. Stage of development
The development stage is the process of preparing authentic assessment instruments tailored to the peer assessment indicators of achievement basic chemistry lab courses. At this stage, the authentic peer assessment instrument was revised based on input from validator experts that will be used in the pilot phase. The measures undertaken in this development is the validation of the product as follows.

a. Validation device
Validation of experts was conducted to determine the feasibility of the products developed. Validation is done through two stages, validation expert in materials chemistry and media education. Product feasibility assessment instruments authentic peer assessment was done by a validator material use assessment instruments that validation sheet that contains some aspects and indicators that have been determined. Validator advises from experts in the field of chemical materials used as a basis for authentic assessment instrument repair products for the better. The results of the validation expert in the field of materials and media can be seen in the following table.

| No. | Rated aspect | Evaluator |
|-----|--------------|-----------|
| A. Aspect Content (Creative Content) | 1 Conformity assessment dimensional aspects of working with the observed 4 | 4 |
| | 2 Authentic assessment peer assessment rubric clearly stated and carefully | 4 3 |
| | 3 Aspects observed assessment presented with clear and concise 3 | 4 |
| | 4 Scoring is appropriate in accordance with the employment dimension | 4 3 |
| | 5 Observation sheet presented clearly and easily understood 3 | 4 |
| | 6 The assessment criteria carefully detailed and clear | 3 3 |
| B. Aspects of Language (Language Program) | 1 The words used straightforward, concise, and easy to understand 4 4 |
| | 2 Do not give rise to ambiguity 4 | 3 |
| | 3 Compliance with the rules discussed Indonesia is good and right 4 | 4 |

The level of validity of authentic assessment peer assessment techniques derived from two validators average is 31, which is at a very valid category with a value of $30 \leq M \leq 36$. In addition to providing an assessment on the sheet validation, validator also provides advice and criticism for product improvement authentic assessment instruments peer assessment.

b. simulations
This step is done teaching the design process that aims to make it easier for researchers when performing tests on products that are already in a valid criterion of validator matter experts and media. Simulation results are referred to as the outlined steps to be performed during a test, a measure that begins with setting up the task in the form of practical activities to be carried out, and then provide an explanation regarding the products developed and began the process of peer assessment performance assessment techniques.
c. Small test
A small test, the product instruments authentic assessment peer assessment tested in practical small groups with the number 9 students were selected based on the ability of different, they are 3 students with low ability, 3 students with the ability to moderate, and 3 students with the high ability. So, the samples can be representative of the population. Meetings During the first phase students are given tasks such as basic chemistry lab activity with the manufacture of the reagent material. These small-scale trials are used to obtain a product of authentic assessment instruments better peer assessment before application in a large trial. The test is also helped by a laboratory assistant as a benchmark in authentic peer assessment. Intent does this limited test namely that researchers can get an idea related to the application of the product that will be tested on learners with real class [21, 22]. Results of a prospective educator performance value chemicals through authentic assessment obtained can be seen in the appendix.

d. Large trials
The large trial stage is an advanced stage of the simulation phase and the test phase in which little has been done actual research on products that have been developed in chemistry teacher prospective students with real class. The study which took place on the first day before going to class, researchers conducted activities related product introductions authentic assessment of performance chemistry lab basic with peer assessment of the chemical laboratory assistant. After that, based on the recommendation of the three assistants, researchers given the opportunity to conduct field trials on learners with real class. This field trial was attended by students of chemical education 2nd half with the number of students as many as 32 people. Activities that take place in the classroom followed by a product introductions delivery of materials lab activities to be carried out in assessing performance[25,26]. The materials related to the manufacture of reagents. For the second meeting, researchers conducted a peer assessment process authentic assessment with laboratory assistants and lecturers to act as assessors to the second and third. Meeting three chemistry teacher prospective students are explained that the material will be practiced material chemical equilibrium. The fourth meeting into the continued implementation of authentic assessment peer assessment with laboratory assistants and lecturers to act as assessors to the second and third. In the fifth meeting, prospective pre-service teacher chemistry students were given an explanation of chemical materials to be practiced that matter acid-base titration. Later in the meeting of the sixth implementation of authentic assessment peer assessment with laboratory assistants and lecturers to act as assessors to two and three at a time to distribute a questionnaire to prospective educators to assistant chemical and basic chemical laboratory to ascertain the practicality of the product developed.

Class test data were analyzed using the Spearman-Brown formula to test the reliability of the products' authentic assessment instruments peer assessment. Based on the total scores obtained on each item, look for correlation using the Pearson product-moment correlation formula. Once calculated, the obtained correlation coefficient of 0.77. The correlation coefficient obtained result of 0.78 which states that the reliability of the instrument is in the strong category as a value in the range of 0.60 <R ≤ 0.80.

The feasibility of authentic assessment instruments peer assessment was done by analyzing data from the questionnaire responses educator candidate chemicals that have been filled after testing a large class. The results of data analysis showed that the average value of the chemistry teacher candidate questionnaire responses is 3.02. According to the table of practicality criteria, it can be said the practicality of the instrument in the middle category with a value of 3 ≤ Va ≤ 4. While based on a percentage formula, the value of the average percentage of 82.5% of respondents giving a positive response. The percentage value, the researchers developed a product that can be said practically because of more than 80% of respondents giving a positive response.

The ability to evaluate students in conducting peer assessment can be determined by comparing the results of basic chemical practicum test results based on peer assessment results with the assessment results by teacher assessment. [30]
The figure can be seen that students who gave a rating greater than the researchers' assessment by 45.10%, students who gave the same assessment with the researchers' assessment by 25.7%, and students who gave a smaller assessment than the researchers' assessment by 29.2%.

Based on the results of data processing, the percentage of students for each category of peer assessment skills can be seen in the following figure.

![Peer Assessment Evaluation Capabilities](image)

**Figure 2. Peer Assessment Evaluation Capabilities**

The picture above shows that there are four groups of students' ability to do peer assessment, namely the excellent ability group with a percentage of 31.45%, a good ability group with a percentage of 47.48%, a sufficient ability group with a percentage of 15.79%, and the ability group lacking with a percentage of 5.28%.

4. Conclusion

Based on the analysis of the results of research and development of authentic assessment instruments peer assessment in basic chemistry lab conducted in chemistry education courses State Islamic University Walisongo conclusions can be drawn that the authentic peer assessment instruments in chemical lab gets very valid criteria basis with the level of validity is 31 (very valid), which means that this instrument can be used in the assessment of the performance of the basic chemistry lab at the stage of dissemination. The measurement results evaluation capabilities candidate chemicals are known educator excellent ability group with a percentage of 31.45%, a good ability with the percentage of 47.48%, enough ability with the percentage of 15.79%, and less ability with the percentage of 5.28%.
Reference

[1] Poerwanti, E. Evaluasi Pembelajaran. Modul Akta Mengajar. UMM Press, 2001.
[2] Haryono Agung. Authentic Assessment dan Pembelajaran Inovatif dalam Pengembangan Kemampuan Siswa. Jurnal JPE-Volume 2, No.1. Fakultas Ekonomi Pembangunan UM 2009.
[3] Arifin, Z. Evaluasi Pembelajaran Prinsip Teknik Prosedur. Bandung, Rosdakarya, pp. 179-181. 2009.
[4] Shore, L.M. Tetrick, L.E., Shore, T.H. “A Comparison Of Self, Peer, And Assessor Evaluation of Managerial Potential”. Journal of Social Behavior and Personality. Vol.13 No.1 Hal 85-101, 1998.
[5] Arikunto, S. Dasar-Dasar Evaluasi Pendidikan Edisi Revisi. (Jakarta: PT Bumi Aksara, 2010).
[6] Haryati, M. Model & Teknik Penilaian pada Tingkat Satuan Pendidikan. Jakarta: Gaung Persada Pers, 2008
[7] Chang, Shu-Nu dan Chiu, Mei Hung. The Development Of Authentic Assessments To Investigate Ninth Graders’ Scientific Literacy: In The Case Of Scientific Cognition Concerning The Concepts Of Chemistry And Physics. International Journal of Science and Mathematics Education 3: 117–140. Taiwan, 2005.
[8] Wiggins, G. Assessment: Authenticity, Context and Validity. Phi Delta Kappan, 75(3), 200–214, 1993.
[9] Tiew, F. Business Students View Of Peer Assessment On Classparticipation. International Education Studies, 3(3): 126-132, 2010.
[10] McGourty, J., Sebastian, C., & Reilly, R. Incorporating Student Peer Review And Feedback Into The Assessment Process. Engineering education, 2004
[11] Matthew, W. et al. Designing A Peer Evaluation Instrument That Is Simple, Reliable And Valid Proceedings Of The 2004 American Society For Engineering Education Annual Conference & Exposition, 2004.
[12] Lori, R.N., et al. Developing Peer Assessment Of Lecturing Instrument: Lesson Learned. Academic medicine. 84(8), 2009.
[13] Thiagarajan, S., Semmel, D.S. & Semmel M. Instructional Development for Teaching Teachers of Exceptional Children. Bloomington: Center of Innovation on teaching handicapped, 1974
[14] Bostock, S. Student peer assessment. www.niu.edu/assessment/toolkit/vol3_ish3.pdf (diunduh 3 Juni 2014)
[15] Mardapi, J. Pengukuran Penilaian dan Evaluasi Pendidikan. Yogyakarta: Nuha Medika. 2012
[16] Noonan, B. dan Duncan, C. Peer And Self-Assessment In High Schools Practical Assessment, Research & Evaluation. ISSN 1531=7714 vol. (10)17 Hal 1-8, 2005
[17] Qu, W. Dan Yang, S. “A peer and self assessment project implemented in practical group work”. Journal of Language Teaching And Research. Vol. 1. No. 6 Hal 776-781, 2010.
[18] Nurgiyantoro, B. 2008. Penilaian Otentik. Cakrawala Pendidikan,(Online), Th. XXVII, No.3, (http://eprints.uny.ac.id/1552/1/NOV_08_BURHAN.pdf) diakses 5 Mei 2014.Sanjaya, W. Strategi Pembelajaran Berorientasi Standar Proses Pendidikan. Jakarta: Kencana Predana Media Grup, 2009.
[19] Stiggins, R.J. Assessment For Learning: A key to student motivation and learning. Phi Delta Kappa EDGE, 2006
[25] Sugiyono. *Metode Penelitian Kuantitatif Kualitatif Dan R&D*. Bandung: Alfabeta, 2011

[26] Sugiyono. *Metode Penelitian dan Pengembangan (Research and Development/R & D)*. Bandung: Penerbit Alfabeta, 2016

[27] Sukardi. *Evaluasi Pendidikan Prinsip Dan Operasionalnya*. Jakarta: PT Bumi Aksara, 2009

[28] Surapranata, S. *Analisis Validitas Reliabilitas Dan Interpretasi Hasil Tes Implementasi Kurikulum 2004*. Bandung: PT. Remaja Rosadakarya, 2006.

[29] Wheather C. dan Philips. 2005. *Students assessing student: case studies on peer assessment*. http://www.gees.ac.uk./planet/p15/cpw.pdf, diunduh 15 Juni 2014

[30] White, E. Student Perspectives of Peer Assessment for Learning in a Public Speaking Course. *Asian EFL Journal*. 33(2):1-30. 2009.