Development of performance test instrument in the experiment of law of conservation mass using self and peer assessment’s technique

W Siswaningsih*, Nahadi and D R Firmansyah
Departemen Pendidikan Kimia, Universitas Pendidikan Indonesia, Bandung, Indonesia.

*Corresponding author’s e-mail : wiwi2450@gmail.com

Abstract. The purpose of this research is to develop the instrument of performance assessment of law of mass conservation using self and peer assessment technique that meet valid and reliable criteria. The instrument components consist of task and rubric. The method used is development and validation. Value of the instrument reliability obtained from twice observations that are at four and six students every group with three same observers. Cronbach alpha value for four and six students every group consecutively are 0.94 and 0.76, indicating that value shows that the instrument is reliable. Optimum amount of the students that can be observed are four students. The implementation of the instrument to limited group of students showed that All of the students give positive responses to the instrument used with the interpretation of questionnaire scores >90% that categorized as good.

1. Introduction.
Assessment must include the whole competency’s aspects which are cognitive, skill, and attitude [1]. Assessing for skill competency is able to be done while learning process within practicum method. Assessment that able to measure skill’s aspect in practicum is performance assessment [2]. Performance assessment defines as one of non-test assessment which is done to assess skill’s aspect [3]. The components that compose performance assessment divides into two parts, they are task and rubric. Task is the jobs that make students for doing something, and rubric is the guide for assessing student’s performance [4].

The steps for developing task in the performance assessments instrument are: determining the focus which includes the determination KD and subject matter; determining the context; determining the clearrness of task; determining the object who will be assessed; determining the guide for assessing; review and revise the task [5]. On the other hand, the steps for developing rubric are: identifying the performance; adjusting level of performance; making description of performance [6].

According to the observation in one of secondary school in Bandung, shows that teachers had applied instrument of performance assessment, however it still limits only for assessing performance in groups. Literature review states that performance assessment is better to be conducted for assessing individual in order to reveal student’s strengths and weaknesses [7,8,9,10]. To cope with the weakness of performance assessment, self and peer assessment technique are able to assess skill aspect [11]. Self-assessment defines as an assessment process that involves students to assess themselves [12]. Moreover, peer assessment is kind of assessing friends of the same age [13]. An advantage of self-assessment is
decreasing teacher’s burden in doing learning evaluation [12]. The advantage of peer assessment is that students able to learn from the strengths and weaknesses of others [13]. Therefore, self and peer assessment can be used as a technique for doing performance assessment.

The subject matter that is able in development of performance assessment’s instrument is the Law of Conservation Mass. There are several skills aspect that can be assessed in that subject matter which are skill in mixing of substance and skill in weighing. According to the description above, then the study entitled “Development of Performance Assessment’s Instrument Secondary School Student in the Law of Conservation Mass Practicum with Self and Peer Assessment Technique” will be conducted. The formulation problems are: (1) How the development process of performance assessment’s instrument secondary school student in the law of conservation mass practicum with self and peer assessment technique?; (2) How the quality of performance assessment’s instrument secondary school student in the law of conservation mass practicum with self and peer assessment technique?; (3) How the implementation performance assessment’s instrument secondary school student in the law of conservation mass practicum with self and peer assessment technique?

2. Methods

This study using development and validation method which is developed by Adams and Wieman (2010) [14]. According to that and the development steps performance assessment’s instrument by Cody (1997) [5], obtained the modification of research steps, there are; (1) Planning which covers the suitability of KI and KD Syllabus Kimia 2016 for subject matter the law of conservation mass, field survey, analysis for subject matter the law of conservation mass; (2) Development of instrument which covers organizing performance assessment’s instrument and production of initial draft performance assessment’s instrument secondary school student in the law of conservation mass practicum with self and peer assessment technique; (3) Validation which covers validity test and reliability test performance assessment’s instrument secondary school student in the law of conservation mass practicum with self and peer assessment technique; (4) Trial which covers implementation test performance assessment’s instrument secondary school student in the law of conservation mass practicum with self and peer assessment technique.

The participants in reliability test are 10 students in 10th grade and the participants in implementation test are 32 students in 10th grade. Students in 10th grade is chosen because they are studying the subject matter the law of conservation mass.

The quality of instrument is determined through validity and reliability. Validity used is the content validity. In general, content validity is done by experts with observing and asking to give judgment about a whole items which is validated [15]. Content validity is measured by Eq.1 [16]:

$$CVR = \left( ne - \frac{N}{2} \right) \left( \frac{N}{2} \right)^{-1}$$

(1)

The CVR value=1 shows that the instrument is valid. The reliability used is inter-rater reliability. Inter-rater reliability is a technique of measuring reliability which based on the agreement between raters while doing the observation [17]. The data processing helped with the application of IBM-SPSS 20.0 version for calculating cronbach alpha coefficient. The criteria of Cronbach alpha are $\alpha > 0.9$ very good, $0.7 \leq \alpha < 0.9$ good, $0.6 \leq \alpha < 0.7$ accepted, $0.5 \leq \alpha < 0.6$ bad, $\alpha < 0.5$ not accepted [18].

The implementation of instrument analysed by viewing the suitability between assessment by observer and assessment by students (self and peer assessment), also seeing the student responses after using the instrument. Data processing in suitability between assessment by observer and assessment by student helped with the application IBM-SPSS 20.0 version for obtaining correlation pearson value. Limit value of correlation pearson is 0.25 [19]. Questionnaire processing using the Eq. 2.

$$\text{Percentage of student response} = \frac{\text{Score}}{\text{number of respondents}} \times 100\%$$

(2)

The criteria of questionnaire’s interpretation used are 81% - 100% for very good criteria, 61% - 80% for good, 41% - 60% for adequate, 21% - 40% for deficient, and 0% - 20% for poor [20].
3. Results and Discussion

The processes of development performance assessment’s instrument secondary school student in the law of conservation mass practicum with self and peer assessment technique are: (1) Viewing the suitability KI and KD in Chemistry syllabus 2016 in subject matter the law of conservation mass. (2) Field survey/interview with chemistry teacher in secondary school. (3) Analysing the subject matter the law of conservation mass. (4) Organizing the indicators performance assessment’s instrument secondary school student in the law of conservation mass practicum with self and peer assessment technique consist of indicator and task. (5) Producing initial draft performance assessment’s instrument secondary school student in the law of conservation mass practicum with self and peer assessment technique consist of indicator, task, and rubric.

There are also revision in rubric. Rubric revised are rubric from task 1.2, task 1.33, task 2.1, task 2.2, task 3.1, task 3.2, task 3.6, task 3.7, task 3.9, task 3.10, task 6.1, and task 6.2. There is no revision of word editorial in rubric from task 3.2, 3.7, and 3.10, but there exists punctuation change from plus (+) to coma (,). Instrument declared as valid then trial tested into 10 students in 10th grade for obtaining reliability value. The test reliability was held twice of observation. First observation consists of 4 students and second observation consists of 6 students. Reliability tests’ pattern can be seen in Figure 1 and 2.

**Figure 1.** Reliability test’s pattern of 4 students  
**Figure 2.** Reliability test’s pattern of 6 students

Based on reliability test, obtained the alpha cronbach that can be seen in Figure 3 and 4.

**Figure 3.** Reliability of task 1.1 – 3.5  
**Figure 4.** Reliability of task 3.6 – 7.1

Based on Figure 3 and Figure 4, in group of 4 students there are 18 task items with very good category, 5 task items with good category, 1 task item with accepted category. Average of reliability value in group of 4 student is 0.945 with very good category. According to average of reliability value,
can be concluded that group of 4 students is kind of reliably instrument so that can be used for assessing student’s performance.

In group of 6 students, there are 13 task items with very good category, 1 task item with good category, 3 task items with accepted category, 3 task items with deficient category, and 4 task items with poor category. Average of reliability value in group of 6 student is 0.768 with good category.

Factors that caused low reliability in group of 6 students; low ability of observer for observing 6 students, differences of observer’s accuracy specifically in observing student’s performance of washing and drying tools, differences of distance between observer so that the observer who is closer to students will be more thoroughly, differences of performance speed of student so that some students unreachable by observer’s views.

Implementation test of performance assessment’s instrument secondary school student in the law of conservation mass practicum with self and peer assessment technique is for viewing the similarity of assessment by students (Self and peer assessment) with assessment by observer. Also viewing student response on the used instrument. Students roles as a rater. Rater task is to make an assessment (self and peer assessment) and also implements the law of conservation mass practicum. The result from implementation test is processed with IBM SPSS 20.0 version for obtaining correlation pearson value. The category of correlation pearson based on the category that is developed by Sarwono, J (2009). Result of correlation pearson with observer can be seen in Figure 5 and 6.

**Figure 5.** Self-assessment with observer 1  **Figure 6.** Self-assessment with observer 2

Based on Figure 5 and Figure 6, there are 6 correlation pearsons with perfect category (1.00), 11 correlation pearsons with very strong category ( > 0.75 – 0.99), 10 correlation pearsons with strong category (>0.50 – 0.75), and 5 correlation pearsons with adequate category (>0.25 – 0.50). Average of correlation value between observer 1 and rater is 0.674 with strong category and the average of correlation value between observer 2 and rater is 0.845 with very strong category. So, can be concluded that there are similarity between value by observer (1 and 2) and self-assessment value. Result of correlation pearson peer assessment by observer 1 can be seen in Figure 7, 8, 9, 10, 11, 12, 13, and 14.

**Figure 7.** Correlation pearson group 1  **Figure 8.** Correlation pearson group 2
Based on the description above, can be known that there are 17 correlation pearson with perfect category (1.00), 44 correlation pearsons with very strong category (>0.75 – 0.99), 31 correlation pearson with strong category (>0.50 – 0.75), and 4 correlation pearson with adequate category (>0.25 – 0.50). It shows that students are able to apply peer assessment in the law of conservation mass practicum because there is similarity between value by observer and value by peer assessment. According to that description about implementation test using correlation pearson, show that a whole students can do self and peer assessment so well. It causes there is similarity between value by students (self and peer assessment) and value by observer. To support obtained data from test implementation, there exists student questionnaire.
Based on the analysis of a whole questionnaires, can be concluded that almost a whole of students give good responses into implementation of performance assessment’s instrument of secondary school student in the law of conservation mass practicum with self and peer assessment technique

4. Conclusion
The process of development performance assessment’s instrument secondary school student in the law of conservation mass practicum with self and peer assessment technique consist of viewing suitability of KI and KD, field survey, matter analysis, organizing instrument’s indicator, and producing initial draft instrument. The quality of developed instrument included good category because it has validity value equals to 1 and average of the reliability for 4 students is 0.946 with categorized as good. There are similarity in assessment between observer and students (self and peer assessment) because a whole correlation pearson value is above 0.25. A whole students give good responses into the instrument used with score of questionnaire interpretation >90% with very good category.

5. References
[1] Depdikbud 2016 Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 23 tahun 2016 (Jakarta: Depdikbud)
[2] Abrahams I, Reiss M J and Sharpe R M 2013. The assessment of practical work in school science. Studies in Science Education, 49 2 209-251.
[3] Zainul A 2001 Alternatif Assessment (Jakarta: Universitas Terbuka)
[4] Susilaningisih E 2014 International Conference On Mathematics, Science, dan Education.(ICMSE 2014)
[5] Cody W S 1996 English Teacher’s Proceedings (Jerusalem: Ministry of education)
[6] Wolf K and Stevens E 2007 The role of rubrics in advancing and assessing student learning.
[7] Rita W 2014 Pengembangan Instrumen Penilaian Kinerja Untuk Praktikum Siswa SMA Pada Materi Pokok Faktor Faktor Yang Mempengaruhi Laju Reaksi. (Bandung: UPI)
[8] Khorunnisa H 2015 Pengembangan Instrumen Penilaian Kinerja Siswa SMA pada Praktikum Hidrolisis Garam. (Bandung: UPI)
[9] Novalia R 2015 Pengembangan Instrumen Asemen Kinerja pada Praktikum Pengaruh Konsentrasi terhadap Laju Reaksi (Lampung: Universitas Negeri Lampung)
[10] Rizkayanti N 2016 Pengembangan Instrumen Penilaian Kinerja Siswa SMA pada Praktikum Larutan Penyangga (Bandung: Universitas Pendidikan Indonesia)
[11] Zulharman 2007 Jurnal Ilmu Kedokteran. 11 1 1-45
[12] El-Koumy A K 2010 Paper accepted for presentation at the CPLA conference (Lebanon : Lebanese American University)
[13] Bostock S 2000 Student peer assessment. Learning Technology, 5 (Heslington: Higher Education Academy)
[14] Adams W K and Wieman C E 2011 Development and validation of instruments to measure

![Figure 15. Graph of questionnaire interpretation](image-url)
learning of expert-like thinking *International Journal of Science Education* **33** 9 1289-1312

[15] Sukardi 2009 *Metode Penelitian Pendidikan* (Jakarta: Bumi Aksara)

[16] Lawshe C H 1975 A quantitative approach to content validity *Personnel psychology* **28** 4 563-575.

[17] Jackson S L 2014 *Research methods: A modular approach* (Boston: Cengage Learning).

[18] Bhatnagar R., Kim J and Many J E 2014 Candidate surveys on program evaluation: Examining Instrument reliability, validity and program effectiveness *American Journal of Educational Research* **2** 8 683-690.

[19] Sarwono J 2009 *Statistik itu mudah: panduan lengkap untuk belajar komputasi statistik menggunakan SPSS 16.* (Yogyakarta: Universitas Atma Jaya Yogyakarta)

[20] Riduwan 2013 *Dasar–dasar statistika* (Bandung: Alfabeta)