An assessment of performance appraisal in enthalpy change laboratory activity: a comparative of scoring by peer and teacher

Nahadi*, W Siswaningsih and H A Fatah

Departement of Chemical Education, Universitas Pendidikan Indonesia, Bandung, Indonesia

*Corresponding author’s email: nahadi@upi.edu

Abstract. The purpose of this study was to develop performance assessment instruments and their application to high school practicum for enthalpy change. The instrument used is a validation sheet to see the validity. The instrument that has been analyzed is a performance assessment about students’ laboratory skills in enthalpy change (∆H) laboratory activity. The method used in the research was development and validation. The instrument was evaluated by 8 teachers and 32 students of year XI from one of senior high school in Bandung. The result shows that there is a similarity of the score given by students and that the teacher. This similarity is supported by the pearson correlation score r=0.89 have been categorized as very good. This result implies that performance skills can be administered by peers using this instrument since it provide the similar result with the teachers.

1. Introduction

Assessment that can be used to measure the skill of performance in laboratory work is performance appraisal. Performance appraisal is the assessment that ask the student to demonstrate and apply the knowledge into the contexts that are in accordance with the established criteria. Furthermore, students’ performance can be assessed by performance assessment because this assessment is appropriate to be applied as an assessment tool in a laboratory that can assess the process and results, but the clear criteria are needed to describe the performance that is assessed.

Based on the result of an interview with one of the chemistry teacher in one of the Senior Secondary Schools in Bandung, it was obtained that the performance appraisal that was done in the school was limited by evaluating group performance or assessing the results of the worksheet from each group. The limitation of performance appraisal was caused by an obstacles in the number of observer during the implementation of performance appraisal. There is an alternative form of performance appraisal, called peer assessment which is facilitated by the teacher. This is in accordance with the statement that in implementing the student-centered in the learning process, the teacher has a role as a facilitator for students so that the student can assess their own learning outcomes and work within the groups.

In line with the statement, another researcher revealed that peer assessment is an assessment technique in which students are asked to assess their own classmate related to the status, process, and achievement level of competencies that is learned in certain subjects. Those statements show that the use of peer assessment can be applied in assessing student performance. In line with that statement, another researcher
affirms that peer assessment can be used to assess the students’ cognitive, affective, and psychomotor.

Based on the explanation above, the researcher is interested in conducting research to develop an instrument of performance appraisal in the practical work of enthalpy change. The purpose of this study was to develop performance assessment instruments and their application to high school practicum for enthalpy change.

2. Methods
The assessment was developed by using method as developed by development and validation. The outline of the step of development and validation method that is conducted in this research consist of (1) planning stage, (2) instrument development stage, (3) validation stage, (4) testing stage.

The validation stage that was done in this study is validation content test toward the instrument that is already developed. To analyze the result of the expert consideration, the CVR technique is used. The equation to calculate the CVR of each performance aspect and rubric can be shown in the Equation 1 below.

\[
CVR = \frac{(n - N/2)}{N/2}
\]

Description:
- \(CVR\) : ratio of content validity
- \(n\) : number of validators declared “valid”
- \(N\) : number of validators

The CVR value that is obtained, then compared with the minimum CVR score based on the number of validator. The performance aspects and rubric is accepted if the CVR score is above or equal with the minimum CVR score. In other hand, the performance aspect is rejected if the CVR score is under the minimum CVR score. The instrument that is avowed valid by expert validator, then the realibility test was conducted to see the assessment that was done by researcher is reliable and determining the number of group members that is effective in laboratory work. SPSS program version 20.0 was used to analyzed the realibility of all the test items by using Cronbach’a Alpha as the index. The score that is obtained then analyzed by using SPSS version 20.0 and interpreted through the assessment criteria according to Cronbach’s alpha criteria. The instrument that is avowed valid and already have the optimal number of group member, then tested to 32 students of XI grade in one of Senior Secondary School that is located in Bandung that already learned thermochemistry topic. The test is conducted to measure the correlation of the assessment that was done by researcher and the peer assessment in assessing the performance of Senior Secondary School students in laboratory work about the determination of enthalpy change using developed instrument. The score of students is analyzed by using IBM SPSS version 20 to obtain the pearson correlation. The value of pearson correlation is used to identify the correlation of the score from observer and peer assessment score. The pearson correlation value that was analyzed is interpreted based on its criteria.

The value of pearson correlation can be used as the standard of peer assessment correlation strength in assessing the students’ performance in practical work about enthalpy change determination (\(\Delta H\)) of a reaction. The greater the pearson correlation value, the relationship between the results of peer assessment and the results of the assessment did by observer is getting stronger. It shows that the implementation of peer assessment in assessing the students’ performance in practical work about enthalpy change determination (\(\Delta H\)) of a reaction is better.

3. Result and Discussion
The quality determination of each performance aspects and rubric in terms of validity is done through the content validity. In the content validity stage, all performance aspects and rubric is avowed valid by 5 validators with each performance aspect (task) and rubric that have a CVR value > 0.99. Thus, the 15 performance aspects are stated that it has good criteria in content validity.
The fifteenth performance aspects include the basic skills in practical work about the determination of enthalpy change. The instrument that have been validated in terms of content were tested to Senior Secondary School students in XI grade who have studied thermochemistry topic especially the determination of enthalpy change. The test is conducted to see the assessment that was done by researcher is reliable and determining the number of group members that is effective in laboratory work.

The test is conducted two times by observing in the classes that was recommended by senior chemistry teacher based on the homogen students’ ability. Based on the reliability test and the result of data analyzing by using SPSS version 20.0, the instrument quality for all performance aspects is obtained as 0.971 with the effective number of group member in practical work is 4 students.

The instrument that fulfilled the criteria of content validity and reliability is applied to 32 Senior Secondary School students in XI grade that was divided into 8 groups that consist of 4 students in each group. The assessment pattern is carried out as follows:

Figure 1. The Pattern of Performance Assessment by Peer Assessment

![Peer assessment pattern](image)

Description: S = Student, O = Observer

Figure 2. The Person correlation value between peer assessment and observer assessment.

Based on the data analyzing, group 1-4 is connected by two categories of pearson correlation strength. The high correlation value suggests a high degree of similarity between the assessment
given by the teacher and the peer assessment. There are 42 correlation values between the result of peer assessment and the result of observer assessment which has the very strong category (>0.75-0.99) and there are 7 correlation values which has strong category (>0.50-0.75). There is a similarity on the top of the graph of person correlation value between the red, blue and green color in all students. It happened because there is similarity of peer assessment value in all students who did the practical work. The Pearson correlation values of 11 students is lower than other students, because there is different score that was given in peer assessment and in observer assessment on several aspects of the instrument that was used. Overall, in group 1-4, in terms of the average Pearson correlation value is obtained as 0.808 which is categorized as very strong. It shows that all students in group 1-4 are able to conduct peer assessment to assess the students’ performance in the determination of enthalpy changes using the developed instrument. In other words, the implementation of peer assessment in assessing students’ performance in the practical work about the determination of enthalpy changes in group 1-4 was well-conducted.

![Figure 3. The Person Correlation value between Peer Assessment and Observer Assessment.](image)

Based on the data analyzing, group 5-8 is connected by two categories of Pearson correlation strength [12]. There are 6 correlation values between the result of peer assessment and the result of observer assessment which has the excellent category (1.00), and there are 42 correlation values which has very strong category (>0.75 – 0.99). There is a similarity on the top of the graph of person correlation value between the red, blue and green color in all students. It happened because there is similarity of peer assessment value in all students who did the practical work. It is because there is different score that was given in peer assessment and in observer assessment on several aspects of the instrument that was used. Overall, in group 5-8, in terms of the average Pearson correlation value is obtained as 0.896 which is categorized as very strong. It shows that all students in group 5-8 are able to conduct peer assessment to assess the students’ performance in the determination of enthalpy changes using the developed instrument. In other words, the correlation between peer assessment and observer assessment in assessing students’ performance in the practical work about the determination of enthalpy changes in group 5-8 has a good correlation. [16].

4. Conclusion
The quality of the instrument that is developed has very good validity for all the task items, and rubric. The average value of instrument reliability for the group that consist of 4 students is 0.971 which is categorized as very good. It can be said that the instrument that was developed in this study has good quality. There is a good correlation between peer assessment and the assessment done by observer in assessing the practical work about the determination of enthalpy changes, because all the Pearson correlation values are obtained above the minimum limit with the categories from strong to excellent.
5. References

[1] Ningtyas F K and Agustini R 2014 Pengembangan Instrumen Penilaian Kinerja Siswa Untuk Mengasah Keterampilan Proses Dalam Praktikum Senyawa Polar dan Non Polar Kelas X SMA. UNESA. *Journal of chemical education*. **03** 3 169-175

[2] Nahadi et all 2016 Pengembangan Instrumen Penilaian Kinerja Pada Pembelajaran Titrasi Asam Basa dengan Metode Praktikum. *Jurnal Pengajaran MIPA* **21** 35-41

[3] Slater T F 1998 *Performance Assessment*. (Montana: Department of Physics Montana State University)

[4] Firman H 2013 *Evaluasi pemberlakuan kimia*. (Bandung: Jurusan Pendidikan Kimia FPMIPA UPI)

[5] Kartono 2011 *Efektifitas Penilaian Diri dan Teman Sejawat Untuk Penilaian Formatif dan Sumatif Pada Pembelajaran Mata Kuliah Analisis Kompleks*. Prosiding Seminar Metematika Universitas Muhamadiyah Surakarta (Surakarta : Universitas Muhamadiyah)

[6] Majid A 2014 *Penilaian Autentik Proses dan Hasil Belajar* (Bandung : PT Remaja Rosdakarya)

[7] Nahadi, Firman H and Ulum M 2018 Development and validation of reasoning-based multiple choice test for measuring the mastery of chemistry *Journal of Engineering Science and Technology*, **13** 6 1476-1488

[8] Zulharman 2007 Self dan Peer Assessment sebagai Penilaian Formatif dan Sumatif. *Jurnal Ilmu kedokteran*, **1** 1 1-45

[9] Adams K and Wieman C 2010 Depelovment and Validation Of Instrument to Measure Learning Of Expert-Like Thinking *International Journal Science Education*. **1** 1. 1-24

[10] Lawshe C H 1975 A quantitative approach to content validity 1 *Personnel psychology* **28** 4 563-575

[11] Gliem J H. and Gliem R R 2003 “Cronbach’s Alpha Reliability Coefficient for Likert-Type Scale”. *Midwest Research Practice Conference in Adult, Continuing, and Community Education Calculating, Interpreting, and Reporting*. p. 87

[12] Ningtyas F K and Agustini R 2014 Pengembangan Instrumen Penilaian Kinerja Siswa Untuk Mengasah Keterampilan Proses Dalam Praktikum Senyawa Polar dan Non Polar Kelas X SMA. UNESA. *Journal of chemical education*. **03** 3 169-175

[13] Nahadi, Siswaningsih W and Muhtar H K 2017 An investigation into students’ misconception on electrolyte and non electrolytes solutions with two tier diagnostic test based pictorial, *Journal of Advanced Science Letters* **23** 11 10555-10558.

[14] Nahadi, Firman H and Farina J 2015 Effect of Feedback in Formative Assessment in the Student Learning Activities on Chemical Course to the Formation of Habit of Minds, *Jurnal Pendidikan IPA Indonesia*, **4** 1 36-42

[15] Nahadi, Firman H and Kurniadi A 2018 Development and Validation of Multiple Representation-Based Virtual Test to Assess Students’ Decision-Making Competence of Buffers, *Journal of Education and Learning (EduLearn)*, **44**-51.

[16] Nahadi, Anwar S, and Kharisma D 2017 Investigating formative assessment strategy to chemistry habits of mind (Chom) of buffer solution concept in learning chemistry habits of mind (CHOM) of buffer solution concept in learning chemistry, *TheTurkish Online Journal Of Educational Technology* **276-285**.

Acknowledgements
We acknowledged UPI for supporting the present study. We also would like to thank to Department of chemical education who is already support this study.