

Development of Sciences Generic Skills Assessment (SGSA) Instrument: Basic Chemistry Practicum

Dinda Habba Kamaliya^{1*}, Anita Fibonacci², Zidni Azizati³

^{1,2,3}Chemistry Education UIN Walisongo Semarang

*Correspondence: pendidikankimiauinws@gmail.com

Abstract

Keyword:

Assessment
Generic Science
Skills Instruments
Basic Chemistry

Performance assessment instruments so far were still general, not specifically assessing expected performance in each practicum. This study was aimed to develop instruments to assess generic science skills in basic chemistry practicum. ADDIE development model was used as a methodology in this research with data collection methods through observation and questionnaires. The subjects of this study were 3rd-semester students taking basic chemistry practicum courses at the University. Instrument validation was done by content analysis and reliability test between the rater as the methods to reliability test of the rubrics. The results of validity and reliability tests showed that the instruments meet valid and reliable criteria for the assessment instruments of generic science skills, especially direct observation skills, symbolic language, and logic inference in basic chemistry practicum. The reliability test showed the average value of interclass correlation coefficient (ICC) in the redox and electrochemical practicum was 0.780; in chemical kinetics practicum 0.660; and in the functional group introduction practicum 0.639. The results of the feasibility test showed that the percentage of positive responses was 92.08% and negative responses were 7.92%, so it can be stated that the product received a positive response. Based on the results of the study, the assessment instruments developed were declared to be suitable as assessment instruments.

To cite this article:

Kamaliya, D.H., Fibonacci, A., Azizati, Z. (2020). Development of Sciences Generic Skills Assessment (SGSA) Instrument: Basic Chemistry Practicum. *Thabiea: Journal of Natural Science Teaching*, 3(1), 31-40.

Introduction

Assessment is an important step in a learning process because an assessment has a relationship related to the goals and learning process. A comprehensive assessment is an assessment that involves the process and learning outcomes (Basuki&Hariyanto, 2014). The learning assessment process is carried out on three aspects namely, cognitive aspect, psychomotor aspect, and affective aspects (Sudijono, 2011). Assessment in the learning process is developed based on the principle of an overall, sustainable, oriented towards achievement indicators, and by following per under the learning experience (Abdullah, 2012).

The assessment conducted in learning should be able to describe the overall process of the situation of students in learning activities undertaken (Muchtar, 2010; Kunandar, 2014). Overall assessment can be conducted through performance assessments, another name for performance assessments is authentic assessments or alternative assessments (Muchtar, 2010). Performance assessments are performed when students carry out activities or activities during the learning process.

Learning activities in the form of the practicum is a learning activity that supports the understanding of theories and concepts obtained from theoretical subjects (Susanti, 2013; Winarti, 2014). Besides, the practicum is a learning activity to combine all the skills that students have (Winarti, 2014), one of which is generic skills. Generic science skills can be stimulated through practical activities (Muspiroh, 2012; Yuliyanti, Hasan, and Syukri, 2016), in practical activities students are required to think critically and creatively in solving a problem. This opinion is in line with what was expressed by Redhana and Liliyasi (2008) that learning is not only a transfer of knowledge and skills but rather builds critical thinking skills through scientific work experience.

The results of observations explored by interviews of 64 students of chemistry education in 2018 UIN Walisongo, revealed that the students' skills in interpreting the observations were still in the low category. Observation is not only done on direct observation skills, but also symbolic language skills. The results showed that students of chemistry education were also still low in writing the equation of reaction that occurs in each experiment. Besides, for logical inference skills, the results showed that students were also not able to infer experimental data based on macroscopic phenomena observed with the underlying chemical theory in basic chemistry practicum.

The results of interviews with lecturers found that the assessment process for basic chemistry practicum at the Department of Chemistry Education on Walisongo State Islamic University conducted at the time was obtained through the results of performance assessments, pre-tests, practicum journals, practicum reports, post-tests, and final exam.

Performance assessment at the basic chemical practicum is carried out at the time of the practicum. However, the performance assessment carried out in the implementation of basic chemical practicum has not been emphasized on the substance of basic chemical substances, or in other words, the existing performance assessment is not specific to each practicum, while the performance assessment at each practicum must be different in its assessment. Educators as an educational component, of course, must be responsible for the quality of learning in the classroom (Fibonacci, 2012), one of which is related to the quality of assessment. Performance assessment is also carried out during the final exam (response) using performance instruments. However, the performance assessment rubric is only limited to the practicum of making reagents and substance separation techniques, whereas for other practicums the assessment is only in the cognitive aspect. If such assessment continues to be done then it will not be measured the level of student competency in the application of basic chemical material, so it will reduce the quality of student's awareness learning in performing practicum (Urios, 2015).

One of the solutions that can be done so that practicum activities can run well is by making improvements to the performance appraisal that emphasizes the understanding of basic skills, chemical concepts, thinking skills, communicating the process and learning outcomes of chemistry (Jusniar, Side and Anwar, 2014). So, from the results of these observations and interviews, it is necessary to develop a performance assessment instrument specifically in assessing basic or generic science skills in basic chemistry practicum.

Based on the explanation, this research aims to develop a product for the assessment of generic science skills and to determine the quality of the products that have been developed.

Method

This research was a research and development method that develops an instrument for evaluating generic science skills in basic chemistry practicum. This research uses the ADDIE model adopted from Branch (2009). The research was held at the Department of Chemistry Education of UIN Walisongo Semarang. Techniques of data collection in research include interviews, observation, and documentation. The following stages of research:

1. Analyze

The purpose of the Analyze phase was to identify the problem in the learning process. The analysis phase begins with a preliminary study at the UIN Walisongo Chemistry Laboratory in Semarang, which includes 2 stages, namely observation, and interviews. The research subjects are all students of semester 3 at the Department of Chemistry Education UIN Walisongo Semarang.

2. Design

The design phase was known for making designs or making product designs that will be developed (Branch, 2009). The design phase of this research includes: determining aspects of generic science skills, making grids assessment of generic science skills, and assessment rubrics of generic science skills.

3. Development

At the development stage, the product that has been designed must be realized to form a product that is developed (Branch, 2009). The instrument developed in this study was a non-test assessment instrument in the form of an observation sheet and performance assessment guidelines.

Product designs that have been compiled then developed based on the results of the validity test by expert validators in their respective fields. The formula used in calculating the percentage is as follows:

$$\%X_{in} = \frac{\sum S}{S_{maks}} \times 100\%$$

$\%X_{in}$: Percentage score score of the 1st question on the instrument validation sheet

$\sum S$: Total Answer Scores

S_{maks} : Maximum Score

(Akbar, 2013)

The results of calculating are interpreted in Table 1.

Table 1. Validity Criteria for PKGS Instruments

Percentage (%)	Interpretation
81 – 100	Very Valid
61 – 80	Valid
41 – 60	Less Valid
21 – 40	Invalid
0,0 – 20	Very Invalid

(Akbar, 2013)

4. Implementation

The implementation phase aims to apply generic science assessment instrument products on basic chemistry practice to obtain good product effectiveness data. Subjects at the implementation stage amounted to 9 people from the 3rd-semester students of Chemistry Education at UIN Walisongo, who were selected based on high, medium, and low categories from the observations of generic science skills.

At this stage, reliability testing and product feasibility tests are carried out to determine the extent to which the products developed are reliable and suitable.

The reliability test for the generic science skills assessment instrument uses inter-rater reliability, which was calculated used the intraclass correlation coefficient (ICC), with the following formula:

$$\rho = \frac{MSRrs - MSe}{MSr + (k - 1)MSe}$$

Keterangan:

MSRrs : mean square

MSe: mean square of error

k: raters number

(Mardapi, 2012)

Widhiarso(2004) has stated the measuring instrument is said to have adequate stability if intraclass correlation measurements are > 0.50, high stability if the ICC between measurements is ≥ 0.80 and the average value of α at each KGS practicum is 70 0.70 (Widoyoko, 2014). Feasibility testing is measured by the ease of use of the assessment instruments. The data obtained can be analyzed with the following steps:

- a. Counted the number of respondents who gave positive responses according to the aspects in question.
- b. Calculated the percentage, using the following equation:

$$\% = \frac{Score}{Maximum\ Score} \times 100\%$$

(Ali, 2015)

- c. The criteria were set to determine that observers have a positive response to the assessment instruments if 50% of them respond positively to 70% of the aspects in question (Ali, 2015).

5. Evaluation

This evaluation process was aimed to process the revision of the product being developed. In the evaluation process can be seen whether the product developed successfully or not.

Results and Discussion

The results of this research were instruments for evaluating generic science skills in basic chemistry practicum. The development of a generic science skills assessment instrument was based on the unavailability of performance assessment instruments for each topic in the

basic chemistry practicum course. Based on the research conducted, a generic science assessment skills instrument was developed in a valid and reliable basic chemistry practicum.

The results were that it can help the assessment of student performance in the implementation of basic chemistry practicum especially on redox and electrochemical materials, chemical kinetics, and introduction of functional groups. Besides, the results of the development can improve direct observation skills, symbolic language, and logic inference. Jusniar (2014) stated that the implementation of chemical practicums accompanied by authentic assessment tools could improve student generic skills. Also, according to Urios (2015) assessment instruments that use assessment rubrics, both holistic and analytical assessment rubrics can provide fair and accurate assessments and foster understanding in the learning process.

The rubric developed in this study was the analytical rubric which is a rubric that is bound to the content of a particular field of study so that its use is very specific only to a particular field of study or material (Zainul, 2001). The rubric developed aimed to assess the skills of direct observation, symbolic language, and logical inference on 3 types of practicum titles namely redox and electrochemical practicum in chemical kinetics practicum and functional group recognition practicum.

The choice of analytic rubrics was based on the specific needs of the types of direct observation skills, logical inference, and symbolic language. Nitko (2001) was stated to get detailed and specific response results, the type of rubric that is suitable is the analytic rubric when compared to the holistic rubric.

**RUBRIK PENILAIAN KETERAMPILAN GENERIK SAINS
DALAM PRAKTIKUM REDOKS DAN ELEKTROKIMIA**

No	Aspek Keterampilan Generik Sains	Indikator KGS	Aspek Penilaian	Indikator Penilaian	Skor
1	Pengamatan Langsung	a. Menggunakan sebanyak mungkin indera dalam mengamati percobaan	Praktikan mampu mengamati dan mencatat perubahan yang terjadi pada semua percobaan yang dilakukan.	1. Praktikan mengamati proses percobaan dan mencatat hasil pengamatan pada reaksi pembentukan gas.	Skor 3 jika melakukan 3 indikator Skor 2 jika melakukan 2 indikator Skor 1 jika melakukan 1 indikator
				2. Praktikan mengamati proses percobaan dan mencatat hasil pengamatan pada reaksi kekuatan asam.	
				3. Praktikan mengamati proses percobaan dan mencatat hasil pengamatan reaksi perubahan bilangan oksidasi	
	b. Mengumpulkan fakta-fakta hasil percobaan	Praktikan menemukan dan mengungkapkan fakta tentang percobaan yang dilakukan.	1. Praktikan menyebutkan 1 fakta pada percobaan pembentukan gas	Skor 3 jika 3 indikator terpenuhi Skor 2 jika 2 indikator terpenuhi Skor 1 jika 1 indikator terpenuhi	
			2. Praktikan menyebutkan 1 fakta pada percobaan kekuatan asam		
			3. Praktikan menyebutkan 1 fakta pada percobaan reaksi bilangan oksidasi		

Figure 1. SGSAExample

The process for testing the validity and reliability was described as follows:

1. Validity Test

A validity test has aimed to see the strengths and weaknesses of the product being developed. The validation test stage is also used to find out whether the product being developed was feasible or not. Widoyoko (2011) stated that the validity test is used to test whether a particular instrument can be said to be valid so that the instrument is not ambiguous when used.

The determination of the validity of the SGSA(science generic skills assessment) was carried out by calculating the score given by the validator and analyzing the

validator's suggestions and input on the product assessment questionnaire. Following were the results of validity tests conducted by experts on products developed:

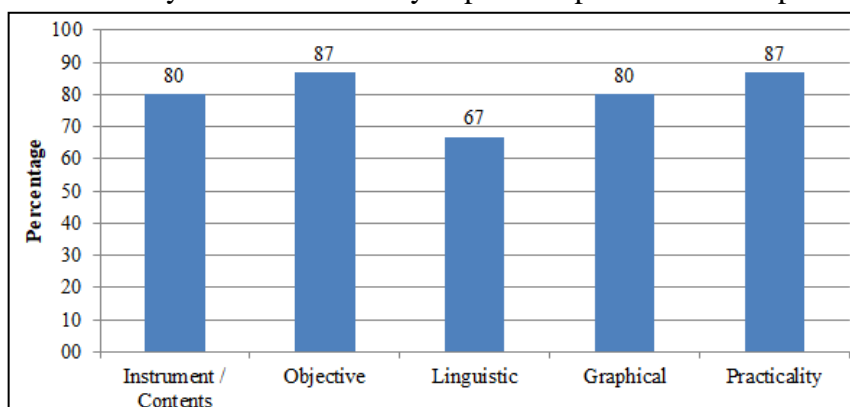


Figure 2. Validity Test Results

Based on experts' assessments, the quality of the instruments met valid criteria and very valid for each criterion. If the overall average score was calculated, an average percentage score was 80% (Valid). The results of the validation from the experts showed that SGSAI instrument was feasible so that it could be used to assess generic science skills at the practicum.

2. Reliability Test

Reliability tests for generic science skills assessment instruments used interrater reliability which was calculated using Intraclass Correlation Coefficients. The use of the intraclass correlation coefficient is preferred over the kappa coefficient because the scores generated on the generic science assessment instrument are more ordinal (Widhiarso, 2004).

Reliability testing for generic science skills assessment instruments uses interrater reliability, which is analyzed through the SPSS 16.0 program. The reliability test results for each item can be seen in the following graph:

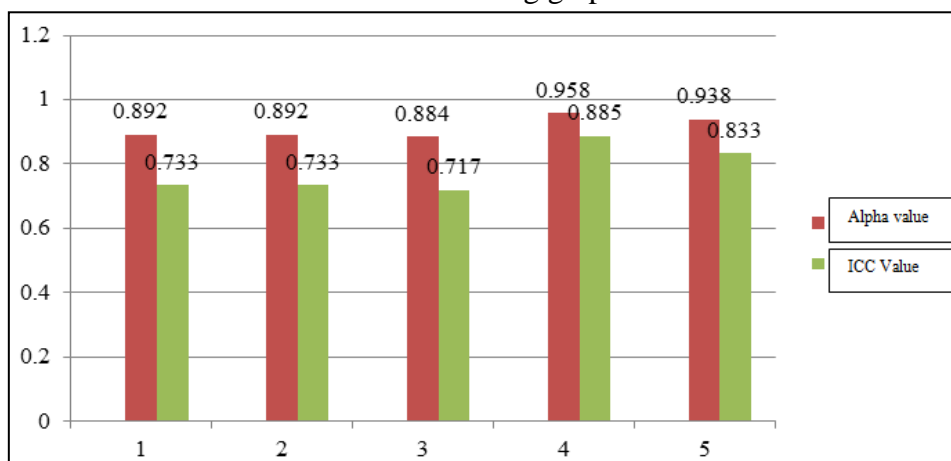


Figure 3. Reliability Results in Redox and Electrochemical Practicum

Figure 3 shows the results of the reliability test on the KGS assessment instrument for redox and electrochemical practicum, that in the assessment aspects 1 and 2 which are assessment indicators for the KGS aspect the direct observation shows the α value ≥ 0.70 which is 0.892 and the ICC value > 0.50 which is 0.733 . The assessment aspect 3 is an

assessment indicator for the KGS aspect of symbolic language, which produces a value of $\alpha \geq 0.70$ which is 0.884 and an ICC value > 0.50 which is 0.717. Whereas, for the aspects of evaluation 4 and 5 are assessment indicators for the KGS aspect of logical inference that shows the value of $\alpha \geq 0.70$, namely 0.958 and 0.938; ICC value > 0.50 which is 0.885 and 0.833.

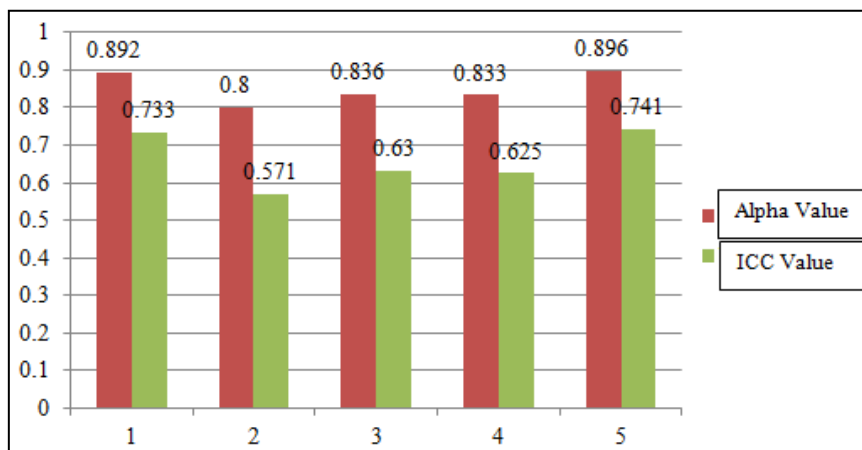


Figure 4. Reliability in chemistry kinetics practicum

Figure 4 shows the reliability on the KGS assessment Instrument for chemical kinetics practicum. Results of the reliability on the assessment indicators for a direct observation of the KGS are shown in the assessment aspects 1 and 2, where the graph shows the result of $\alpha \geq 0.70$ which is 0.892 and 0.800; ICC values > 0.50 , 0.733 and 0.571. The assessment aspect of 3 and 4 is the assessment indicator for the aspect of KGS symbolic language with the result of $\alpha \geq 0.70$ which is 0.836 and 0.833 and ICC value > 0.50 is 0.630 and 0.625. While the aspect of evaluation 5 is the assesment indicator for the KGS aspect of logical inference that shows the value of $\alpha \geq 0.70$ namely 0.896 and ICC value > 0.50 which is 0.741.

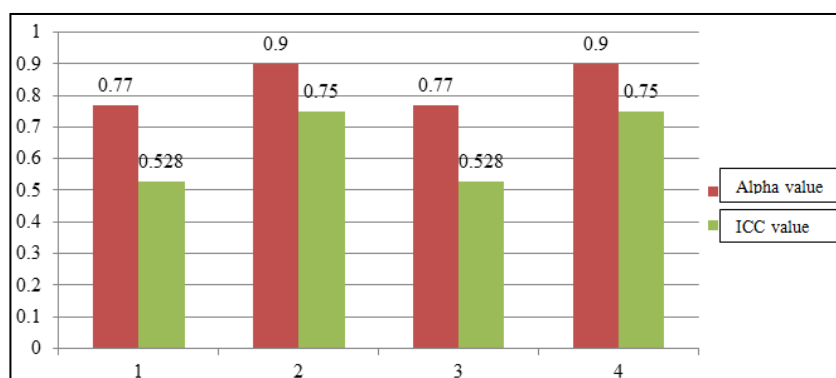


Figure 5. Reliability in functional Group Introduction Practicum

Based on Figure 5, it is known that the results of the reliability on the KGS assessment Instrument for introduction of function group practicum. On the assessment Aspect 1 and 2 is the assessment indicator for the KGS aspect of direct observation with the result of $\alpha \geq 0.70$ namely 0.770 and 0.900; ICC values > 0.50 , 0.528 and 0.750. The assessment Aspect 3 is the assessment indicator for the aspect of KGS symbolic language with the result of $\alpha \geq 0.70$ which is 0.770 and ICC value > 0.50 is 0.528. Meanwhile, for

the assessment aspect 4 is assessment indicator for the KGS aspect of logical inference that shows the value of $\alpha \geq 0.70$ which is 0.900 and $tICCv\text{value} > 0.50$, i.e. 0.750.

The difference in the number of practice aspects of the practicum specific is due to the inequality of the KGS skill indicator that appears in each chemical practicum. Thus, the development of the assessment instrument is only derived from the existing KGS skill indicator in the chemical practicum.

The results of reliability tests in all practicums belong to the reliability category, Widhiarso (2004) reported that the instrument is named as reliable if ICC (Correlation Coefficients) values on each KGS indicators > 0.50 and α values in each of the KGS indicators ≥ 0.70 . The reliability test results show that a good perception created between the observer about generic science skills in the generic science skills of assessment instruments on basic chemical practicum (redox and electrochemistry, chemical kinetics and the introduction of functional group).

3. Feasibility test

Feasibility test is used to determine the user's perception or response to the instrument. Eligibility is measured from the level of ease in the use of improved assessment devices. Product feasibility Data is obtained by giving a response questionnaire to lecturers and Observer as a user. The response questionnaire contains of the questions and the evaluation column of the lecturer's response (SS, S, KS, TS) to that question. Positive response showed by providing 3 and 4 score and negative response showed by giving 1 and 2 score. The results of feasibility test can be seen in Figure 6.

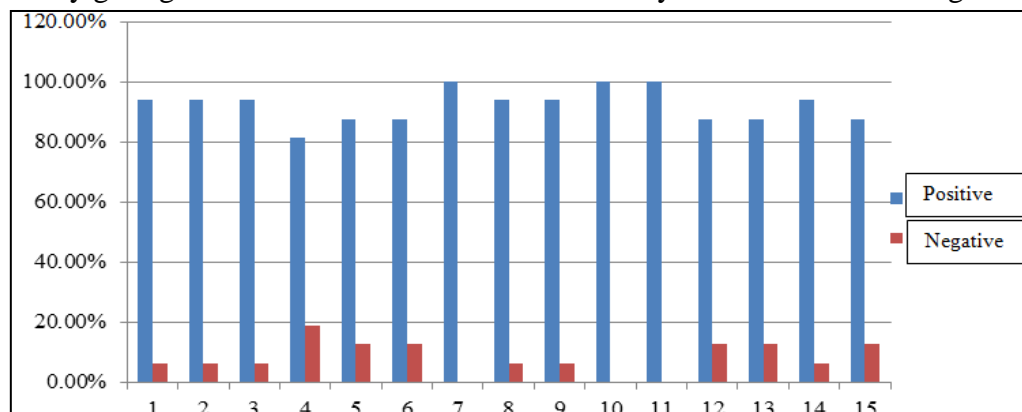


Figure 6. Feasibility test results

Based on Figure 6 it can be seen that the grades given by lecturers and observer as respondents are 92.08% for positive response and 7.92% for negative response, so that it can be stated that the product gets a positive response. This means the product of the generic science skills Assessment instruments on basic chemistry practicum is stated practically.

Conclusion

The conclusion of the research is a generic science skill assessment instrument that is developed have a decent quality to use, where the validation test results show a percentage yield of 80.00% which belongs to the valid category. The results of reliability tests on each practicum belong to the reliability category, where the average ICC values in each practicum

> 0.50 and the average value of α in each practice of KGS $\geq 0,70$. Due feasibility test results show a positive response percentage of 92.08% and a negative response percentage of 7.92%, so it can be stated that the product is getting a positive response.

References

- Akbar, Sa'dun. (2013). *Instrumen Perangkat Pembelajaran*, Bandung : PT Remaja Rosdakarya Offset.
- Ali, Ahmad, dkk. (2015). Pengembangan Perangkat Assesmen Praktikum Anatomi Fisiologi Manusia Berbasis Keterampilan Proses Sains Mahasiswa Jurusan Pendidikan Biologi UIN Alauddin Makasar. *Jurnal Program Pascasarjana UNM*. Vol 2, No 1.
- Basuki, Ismet dan Hariyanto. (2014). *Assesmen pembelajaran*, Bandung : Remaja Rosdakarya.
- Branch, R. M. (2009). *Instructional design : the ADDIE approach*. New York: Springer.
- Jusniar, Side, S. and Anwar, M. (2014). Pengembangan perangkat assesment berbasis keterampilan generik sains (KGS). *Jurnal Pendidikan kimia*, 1(1), pp. 35–42.
- Kunandar. (2014). *Penilaian Autentik (Penilaian Hasil Belajar Peserta Didik Berdasarkan Kurikulum 2013) Suatu Pendekatan Praktik*. Jakarta :Rajawali Pers.
- Mardapi, Djemari. (2012). *Pengukuran Penilaian dan Evaluasi Pendidikan*. Yogyakarta : Nuha Medika.
- Muchtar, H. (2010). Penerapan Penilaian Autentik dalam Upaya Peningkatan Mutu Pendidikan. *Jurnal Pendidikan Penabur*.
- Muspiroh, N. (2012). Analisis kemampuan generik sains mahasiswa calon guru biologi pada praktikum anatomi tumbuhan. *Jurnal Scientiae Educatia* Vol. 1 Edisi 1 Analisis, 1 (April), pp. 1–15.
- Nitko, A. J. (2001). *Educational assessment of students (3rd ed.)*. Upper Saddle River, NJ: Merrill.
- Redhana, I. W. and Liliyasi. (2008). Program pembelajaran keterampilan berpikir kritis pada topik laju reaksi untuk siswa SMA. *Forum Kependidikan*, 27(2), pp. 103–112.
- Sudijono, Anas. (2011). *Pengantar Evaluasi Pendidikan*. Jakarta: Rajawali Pers
- Sudjana, Nana. (2014). *Penilaian Hasil Proses Belajar Mengajar*. Bandung : PT Remaja Rosdakarya.
- Susanti, R. (2013). *Pengaruh penerapan pembelajaran berbasis masalah pada praktikum fotosintesis dan respirasi untuk meningkatkan kemampuan generik sains mahasiswa pendidikan biologi FKIP Unsri*.
- Urios, M. I. et al. (2015). Generic skill development and learning/assessment process: Use of rubrics and student validation, *Journal of Technology and Science Education*, 5(2), pp. 107–121.
- Widoyoko, Eko Putro. (2011). *Evaluasi Program Pembelajaran*. Yogyakarta : PUSTAKA PELAJAR.
- Widoyoko, Eko Putro. (2014). *Penilaian Hasil Pembelajaran di Sekolah*. Yogyakarta : PUSTAKA PELAJAR.
- Widhiarso, W. (2004). *Mengestimasi reliabilitas*. Yogyakarta: UGM PRESS.

- Winarti, T. (2014). Pembelajaran Praktikum Berorientasi Proyek untuk Meningkatkan Keterampilan Proses Sains dan Pemahaman Konsep. *Jurnal Inovasi Pendidikan Kimia*, 8(2), pp. 1409–1420.
- Yuliyanti, E., Hasan, M. and Syukri, M. (2016). Peningkatan Keterampilan Generik Sains dan Penguasaan Konsep Melalui Laboratorium Virtual Berbasis Inkuiri. *Jurnal Pendidikan Sains Indonesia*, 4(2), pp. 76–83.
- Zainul, A. (2001). *Alternative Assessment*. Jakarta: Dirjen Dikti.