

# THE INTERNATIONAL JOURNAL OF SCIENCE & TECHNOLEDGE

## Exploring Item Validity: Critical-Science Process Skills Integrated Assessment Instrument on Electrolyte Solution and Acid Base Matter

**Suyanta**

Senior Lecturer, Department of Chemistry Education, Faculty of Mathematics and Natural Science,  
Yogyakarta State University, Indonesia

**Endang Widjajanti LFX**

Professor, Department of Chemistry Education, Faculty of Mathematics and Natural Science,  
Yogyakarta State University, Indonesia

**Eli Rohaeti**

Senior Lecturer, Department of Chemistry Education, Faculty of Mathematics and Natural Science,  
Yogyakarta State University, Indonesia

**Anggi Ristiyana Puspita Sari**

Student, Graduate School Program of Chemistry Education, Yogyakarta State University, Indonesia

**Sri Rejeki Dwi Astuti**

Student, Graduate School Program of Chemistry Education, Yogyakarta State University, Indonesia

### **Abstract:**

*The demanding of assessment in learning process was affected by policy changes. Students should be prepared to address the challenges with higher level skills involving critical thinking and science process skills, so that they can solve problems in educational learning and other fields. Recognizing the importance of the development of critical thinking and science process skills, the instrument should give attention to the characteristics of chemistry. Therefore, constructing an accurate instrument for measuring those skills is important. However, the integrated instrument assessment is limited in number. This paper aimed how to develop integrated assessment instrument and to verify instrument's validity such as content validity and construct validity on electrolyte solution and acid base matter. The model development of the test instrument adapted McIntire mode. Process data development was acquired based on the development test step and was analyzed by qualitative analysis. To acquire content validity, initial product was observed by six peer reviewers of postgraduate students and 14 experts including ten teachers and four lecturers. This research involved 376 students of grade X from two senior high schools in Bantul Regency and 392 students of grade XI from three senior high schools in capital city of Yogyakarta to acquire construct validity. Exploratory factor analysis (EFA) was conducted to explore construct validity, whereas content validity was substantiated by Aiken's formula. The results showed that all items in integrated assessment instrument are asserted valid according to content validity and construct validity. Therefore, the integrated assessment instrument is suitable for measuring critical thinking abilities and science process skills of senior high school students on electrolyte solution and acid base matter.*

**Keywords:** acid base, critical thinking skills, electrolyte solution, item validity, science process skills

### **1. Introduction**

Nowadays, we are in the area of science and technology where scientific knowledge has grown exponentially and technology has progressed at a rapid pace. Science education plays a major role for society in the future. Countries around the world particularly the developing ones have continuously sought to improve the quality of science education particularly developing countries. Based on Indonesian Law number 20/2003, the government designs and constructs curriculum to produce human resources who are productive, creative, innovative, and affective throughout attitude, skills, and knowledge in integrated manner (Depdiknas, 2003). Assessment is an integral part of the learning process. It is an evaluation process of determining the quality of learning process. It is impacted by policy changes. Based on educational standard, it should be covered three aspects of learning process i.e. knowledge, attitude and skill (Kemendikbud, 2013). Therefore, assessment must include all aspect in learning process. When students engage in assessment, they should learn from those assessments (National Academy of Science, 1996, p.5).

Chemistry learning cannot be separated from the development of students' cognitive and students' skills. Process skills are essential for understanding and applying chemistry concept, so assessment in chemistry must cover both knowledge and process skills (Marasigan & Espinosa, 2014). Chemistry is a branch of science that apply higher order thinking skills in each matter and holding an

important role in developing students' critical thinking skills (Zhou, Huang & Tian, 2013). It is easily fastened to daily life phenomena. Furthermore, it can be applied in society. Some of chemistry matters in senior high schools are electrolyte solution and acid base matter. Electrolyte solution and acid base matter are contextual matters, so it can be easily fastened and applied in daily life. Many electrolyte solution phenomena can be found in everyday life such as storage battery application and body requirement for electrolyte solution. Moreover, electrolyte solution matter is an electrochemistry prerequisite matter. Meanwhile, the clear example of acid base phenomena is an acid base characteristic of things like environment pollution that is indicated by the presence of an acid base solution. Moreover, it is connected with other matters such as chemistry reaction, equilibrium, and stoichiometry (Demircioğlu, Ayas, & Demircioğlu, 2005). Furthermore, we can do experiments about electrolyte solution and acid base matter. Therefore this matter does not only develop students' critical thinking skills but also students' science process skills.

Critical thinking is one of thinking skills in higher order thinking skills which becomes one of learning goals in 21<sup>st</sup> century (Duran & Sendag, 2012). Critical thinking abilities can help students solve problem efficiently and it is a way to self-learning (Mahapoonyanont *et al.*, 2010). Critical thinking is an essential way on scientific investigation process especially for analyzing and evaluating scientific evidences (White *et al.*, 2011). Critical thinking abilities can emerge when students use higher order thinking abilities, so students' learning results is enable to increase. Students who have critical thinking abilities will have a high academic achievement. The result of Wenglinskys' research explained that students who are accustomed to think critically in learning process would be able to acquire high test score on learning evaluation (Moore & Stanley, 2010, pp. 16-21). Critical thinking abilities is not only necessary in school, but also in dialy life. In fact, drawing conclusion is very needed in all occupation section, so critical thinking abilities will increase a person's success in work pressure (Moore & Stanley, 2010). Moreover, increasing job requirement at worker who can analyzed and applying information critically and give efective solution about problem in work force (Stephenson & Sadler-McKnight, 2016). Therefore, critical thinking abilities are one of education goals in globalization era, so student have to be trained to use critical thinking abilities in chemistry learning.

Science process skills are a necessary tool to produce and use scientific information, to perform scientific research, and to solve problems (Aktamis & Ergin, 2008). It is more important for the students to learn how to apply science than only learning reality, concepts, generalizations, theories and laws in science lessons (Karamustafaoğlu, 2011). Moreover, Science process skills have a correlation with critical thinking skills because it has a close relationship with the mastery concept (Chebii, Wachanga, & Kiboss, 2012). Science process skills are divided into two level skills i.e. basic science process skills and integrated science process skills (Rezba *et al.*, 2007, p.xx; Keil, Haney & Zoffel, 2009). Basic science process skills are fundamental skills that underlie scientific methods. It consists of six process skills i.e. observing, communicating, classifying, measuring, inferring and predicting. Integrated science process skills are the development of basic science process skills. It includes eleven skills i.e. identifying variable, constructing a table of data, constructing a graph, describing relationship between variable, acquiring and processing data, analyzing investigation, constructing hypotheses, defining variables operationally, designing experiment and experimentation. Based on the experts' statement, students should be prepared to address the challenges with higher level skills involving critical thinking and science process skills, so that they can solve problems in education learning and other fields.

Nowadays, learning process only emphasizes the cognitive aspect, while process skills get less attention (Zoller, 2001). The observation which was conducted at nine schools in Yogyakarta city and Bantul regency in the last 2015 indicated that many teachers didn't assess students exhaustively and they focused more on cognitive skills than science process skills. Teachers still set out assessment of students' knowledge, but assessment of science process skills was only based on teachers' subjective assumption. Assessing both critical thinking and science process skills needs more time allocation, so integrated assessment can be an alternative assessment. However, the integrated assessment is limited in number, so assessing students' skills cannot be exhaustively. Nowadays, cognitive and skills assessment are carried out separately (Keil, Haney, & Zoffel, 2009; Saribaz & Bayram, 2009; Tosun & Taskesenligil, 2013). If it is carried out separately, relationship between cognitive and skills can be acquired. Therefore, integrated assessment instrument is needed for measuring critical thinking abilities and science process skills in one instrument. The instrument development goals are to facilitate teacher for assessing the learning result effectively and efficiently and to develop valid instrument. By realizing the importance of integrated assessment instrument, this research study was aimed at developing and validating an integrated assessment instrument for measuring students' critical thinking and science process skills on electrolyte solution and acid base matter. This instrument contains 11 items on acid base matter and 8 items on electrolyte solution matter which each item covers two indicators at once.

## 2. Research Methods

This study adapted the development model of test proposed by McIntire through ten steps which are (a) defining the test universe, audience, and purpose; (b) developing a test plan; (c) composing the test items; (d) writing the administration instructions; (e) conducting piloting test; (f) conducting item analysis; (g) revising the test; (h) validation the test; (i) developing norms; and (j) completing the test manual. The development products are integrated assessment instruments for measuring critical thinking and science process skills of senior high school students on electrolyte solution and acid base matter.

The subject was 376 students of grade X from two senior high schools in Bantul Regency that have enrolled electrolyte solution matter, while the integrated assessment instrument of acid base matter was administered to a total of 392 high school students of XI grade in the academic year of 2015/2016 at five schools in capital city of Yogyakarta. Purposive sample technique was used for determining the sample based on schools that have implemented 2013 curriculum in learning process and school has got the rank of Acceptance Student Data both in Bantul Regency and capital city of Yogyakarta.

Data acquired in this study included development process data as qualitative data and item validity data (content validity data and construct validity data) as quantitative data. Content validity data were acquired by experts whereas construct validity were acquired by instrument field testing. To collect the data the researchers used questionnaire and integrated assessment instrument. Questionnaire was the item of content validity sheet used for verifying item content validity on integrated assessment instruments. Six peer reviewers from postgraduate students and 14 experts which consist of four teachers and four lecturers participated in this study. The initial product was served by peer reviewers and validated by experts.

Integrated assessment instrument was validated in order to represent what it is supposed to measure. Validation process involved theory and evidence collection to support interpretation against test score (Sumintono & Widhiarso, 2015, p.8). Content validity through Likert scale was processed from validation score and analyzed by using Aiken's V (Miller, 2003, p.75). Obtained V coefficient was matched with table of right-tail probabilities for selected values of the validity coefficient by Aiken (Aiken, 1985). The mathematical form of the Aiken's V is (Aiken, 1980):

$$V = \frac{\sum s}{[n(c - 1)]}$$

Explanation:

$s = r - l_o$

$l_o$  = rating scale in the lowest category (example: 1)

$c$  = rating scale in the highest category (example: 5)

$r$  = rating scale that be given by rater

$n$  = amount of rater

Collected data were initially coded in Excel spread-sheet, and then transferred to SPSS 16.0 data editor for data analysis. Exploratory factor analysis (EFA) is a powerful tool for explicating construct. Interpretation from this analysis was used to describe estimation upon dimension which can be measured from integrated assessment instrument. Thus, EFA was used as an assumption of unidimensionality. Factor analysis began by examining the adequacy of the sample, then using computer arranged variance-covariance to calculate eigen value. The eigen value was applied to calculate the percentage of explained variance, at once describe scree plot (Retnawati, 2014, p.19). Factor analysis output covered Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO), Bartlett's test, scree plot, and ant- image matrices. Table 1 shows the interpretation of the KMO (Beavers *et al.*, 2013).

KMO Value	Degree of common variance
0.90 to 1.00	Marvelous
0.80 to 0.89	Meritorious
0.70 to 0.79	Middling
0.60 to 0.69	Mediocre
0.50 to 0.59	Miserable
0.00 to 0.49	Don't factor

Table 1: Interpretation Guidelines for the Kaiser-Meyer-Olkin Test

### 3. Results and Discussion

The developing product of this study was integrated assessment instruments for measuring critical thinking and science process skills of senior high school students on electrolyte solution and acid base matter. Each item consists of critical thinking skills, science process skills, and learning indicators. The initial draft of integrated assessment instrument was essay tests including 8 items with 28 subitems on electrolyte solution, whereas on acid base including 11 items with 26 sub items. Furthermore, the composing scoring manual was appropriated with the difficulty item level and student thinking groove.

Scoring model of the assessment was Partial Credit Model 1-Parameter Logistic (PCM 1-PL). 1-parameter logistic is a difficulty level parameter. The choosing PCM 1-PL was based on its ability to analyze items which have different maximal score. The difference was caused by each item which has the different item quality to measure student ability and the different steps to solve problems. Both initial products were observed by three peer reviewers and the results suggested that almost all items should be revised. Most suggestions from the peer reviewers were taken to be considered and used. Besides, initial products were also reviewed by six experts for electrolyte solution matter and eight experts for acid base matter. The experts included two subject matter experts, two learning evaluation experts, and the rest were teachers.

Items reviewed were committed for validating initial product especially item of content validity by experts. Improvement was done based on its suggestion then product was validated by experts to make extent improvement. All items that have been validated by experts were arranged to be received, revised, and eliminated. Experts' suggestion was linguistic aspect suggestion especially word selection. It was used to make product perfectly. The involvement of lecturers as experts in validating product was aimed at gaining suggestion about depth of matter and suitability of scoring manual and test instrument. The chemistry teachers involvement in validating product was aimed to gain suggestion about visibility and effectiveness of product implementation in learning process such as time allocation, amount of ideal items in essay test, and product readability. Moreover, teachers were considered having experience in composing tests. Therefore, teacher can give some suggestions about the use of language including words and good sentences.

Based on the result of validating product of electrolyte solution matter, it is acquired amount of proper item questions that used in learning process; as many 8 items with 21 sub-items, whereas 7 sub-items were dropped. It was committed because of teachers'

considerations that the amount of item questions which were more than 25 sub-items were too many for examining within 90 minutes. In other words, the dropped sub-items have similar meanings to other sub-items. The dropped sub-items were chosen based on teachers' suggestion. Meanwhile, the result of acid base matter instrument shows that there is no eliminated item, nevertheless only three received items. Almost all of suggestions from the experts associated with the sentence structures. Because the time allocation was just 90 minutes, all items should be understood by students and the elusive terms were not used. The good time for students to finish the test range from 1,5 to 2,5 hours. If it is more than 2,5 hours, it will decrease students' thinking endurance (Utomo & Ruijter, 1994, p.60). If it happens, it will decline tests' reliability. In addition, to consider the use of language, revisions of the instrument were also done on the acid base concept. Composing good sentences are the early step to comprehend questions, so it was very influential for students' comprehension toward questions in the test (Beek & Louters, 1991).

Besides the qualitative data, the validation process was applied to obtain quantitative data based on validation sheets. The estimating content validity was considered by suitability between learning indicators and critical thinking and science process skills indicators. To analyze quantitative data Aiken formula was used to calculate content-validity coefficient based on the expert's validation. The result of validating data by experts were analyzed by using Aikens' formula. It was used to determine whether the items were valid or invalid based on content validity coefficients (Aikens' V). Data analysis results of content validity are shown in Table 2.

Item	Electrolyte Solution	Acid Base
	Aiken's V	Aiken's V
Item 1	0.83	0.83
Item 2	0.94	0.75
Item 3	0.83	0.88
Item 4	1.00	0.83
Item 5	0.83	0.75
Item 6	0.94	0.90
Item 7	0.89	0.80
Item 8	0.89	0.88
Item 9		0.96
Item 10		0.92
Item 11		0.88

Table 2: Content Validity Coefficient of Integrated Assessment Instruments of Electrolyte Solution and Acid Base Matter

Content validity of coefficient significance statistic can be determined by correlating a rating category with amount of raters (Aiken, 1985). Electrolyte solution instrument involved six raters and four rating categories, so based on 0.05 significance level, allowed minimum content validity coefficient (Aikens' V) was 0.78 (Aiken, 1985). Meanwhile acid base matter involved eight raters and used the same rating categories, so based on 0.1 significance level, allowed minimum Aikens' V was 0.75 (Aiken, 1985). As indicated on Table 2, all of items have Aikens' V more than allowed minimum Aikens' V. It can be explained that all items in integrated assessment instrument are valid. Therefore, all items in integrated assessment instrument are appropriate to measure critical thinking and science process skills of senior high school students on electrolyte solution and acid base matter.

Unidimensionality means that each item only measures one dimension: students' ability in chemistry. Thus, each item is expected to only measure students' ability in chemistry on acid base matter. Factor analysis began by examining the adequacy of the sample and then used computer arranged variance-covariance to calculate eigen value. The eigen value is used to calculate the percentage of explained variance, at once describe scree plot. The output covers Kaiser-Mayer-Olkin Measure of Sampling Adequacy testing (KMO-MSA), Bartlett's test, scree plot, and anti-image matrices. Before interpreting the result of construct validity, the researchers committed sample adequacy analysis using KMO-MSA and Bartlett testing. To commit factor analysis, the value of KMO-MSA must be more than 0.5 and the significance of Bartlett testing must be less than 0.05 (Beavers *et al.*, 2013). In summary, the result of KMO and Bartlett's test of electrolyte solution and acid base instruments are respectively shown in Table 3.

		Electrolyte Solution	Acid Base
<b>Kaiser-Mayer-Olkin Measure of Sampling Adequacy</b>		0,605	0,714
<b>Bartlett's Test of Sphericity</b>	Approx. Chi-Square	134,964	446,201
	Df	28	0,55
	Sig.	0,000	0,000

Table 3: The KMO and Bartlett's Test of Electrolyte Solution and Acid Base Matter

As indicated on Table 3, KMO measure of sampling adequacy was 0.714, and Bartlett's test of sphericity was significant ( $\chi^2 = 446.201$ ,  $p < 0.05$ ). Both the KMO and Bartlett's test suggested that factor analysis was feasible because obtained KMO was more than 0.5 and Bartlett's test has significance less than 0.05 (Beavers *et al.*, 2013). In addition, based on Table 1, the KMO is in middling category. In other words, the sample size of 392 students is enough. Bartlett's test provides evidence that the observed correlation matrix is statistically different from a singular matrix and it is conforming that linear combinations exist. As well as in Table 3, the value of KMO-MSA was more than 0.5 and the significance of Bartlett's test must be less than 0.05. Thus, it can be explained that sample is adequate for factor analysis. Therefore, the interpretation of the result can be continued with construct validity analysis or unidimension assumption.

Factor analysis was used to analyze relationship of intervariable using correlation testing. The unrotated factor model suggested that there were three factors components satisfying the eigen value criteria. In factor analysis, the table of total explained variance is an important source of information about the number of significant factors. The result of total explained variance is shown in Table 4 and 5 for electrolyte solution and acid base instrument. The first column in Table 5 lists all variables (the 11 items of the scale) as the initial number of factors. The second column shows the eigen values (the variances of these factors). In the next two sub-columns, the variances are shown as the percent variance and the cumulative percentage variance are accounted for by the factors. Table 6 indicates that integrated assessment instrument contains three factor. If eigen values are less than one, then that factor could not be described further due to lack of information.

As indicated on Table 4, students' response data towards the instrument lade three eigen value (eigen value  $> 1$ ). According to Kaiser criteria, the instrument lade three factors, but there was one dominant factor (Beavers *et al.*, 2013). The dominant factor should be chemistry knowledge because the instrument was developed based on the definition of the grating item contained in basic competence in chemistry especially electrolyte solution matter. Chemistry knowledge tests consist of mathematics abilities and nonmathematics or languag abilities (Lord, 1980, p.20). Whereas, two factors were measured by the instrument including personality factor and administrative factor during the test such as anxiousness and students' motivation (Retnawati, 2014, p.1).

Component	Initial Eigen values		
	Total	% of Variance	Cumulative %
1	1,697	21,207	21,207
2	1,283	16,031	37,238
3	1,067	13,337	50,576
4	0,896	11,195	61,771
5	0,859	10,739	72,510
6	0,798	9,970	82,480
7	0,756	9,452	91,932
8	0,645	8,068	100,00

Table 4: Total Explained Variance on Electrolyte Solution Matter

Component	Initial Eigen values		
	Total	% of Variance	Cumulative (%)
1	2.479	22.539	22.359
2	1.470	13.365	35.904
3	1.134	10.313	46.217
4	0.996	9.057	55.274
5	0.847	7.702	62.976
6	0.814	7.400	70.375
7	0.789	7.171	77.547
8	0.708	6.439	83.986
9	0.670	6.095	90.081
10	0.583	5.301	95.381
11	0.508	4.619	100.000

Table 5: Total Explained Variance on Acid Base Matter

In addition, many factors formed can also be seen through the scree plot. As indicated on Figure 1a, eigen value began slightly on third eigen value. There was a dominant factor which was measured by integrated assessment instrument and two other factors also gave contribution towards instrument responses. Meanwhile, Figure 1b demonstrates the scree plot of the eigen values and factors from the extraction seen in Table 6. The scree plot shows that there are three bends while component number 4 and so on indicates that sloping graphic has the same eigen values. In other words, the eigen values have lower value than the previous factors. Practically, it is hard to find an item that only measure one dimension.

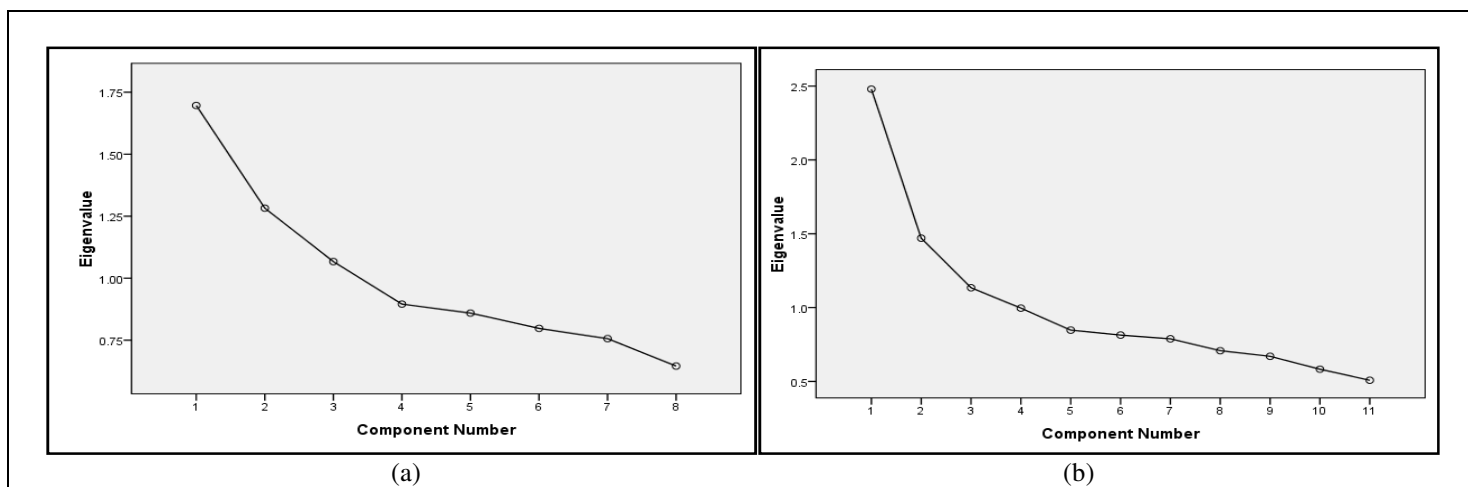


Figure 1: EFA Scree Plot of (a) electrolyte solution matter and (b) acid base matter

Unidimension assumption was very difficult to fulfill ideally (Hambleton & Swaminathan, 1985, p.16). However, unidimension assumption can be considered to fulfill if tests contained one dominant factor (Hambleton, Swaminathan & Rogers, 1991, p.56; Wiberg, 2004, p.5). Unidimensionality can be seen from the first eigen value which is more than 20% (Reckase, 1979). Both in Table 5 and 6 the first eigen values is more than 20%, so the unidimensionality requirement has been fulfilled. It means that the integrated assessment instrument found to comply the construct validity, because it only measured one knowledge dimension i.e chemistry knowledge.

#### 4. Conclusion and Suggestions

Based on the results of study, we can conclude that the integrated assessment instrument on electrolyte solution and acid base matter was asserted valid according content validity and construct validity. It was suitable for measuring critical thinking abilities and science process skills on electrolyte solution and acid base matter. Our interest when designing this research stemmed from our willingness to assess students' cognitive and skills aspects in chemistry. Due to the limited work in this area, our first step was to design integrated assessment instrument, so that the instrument validity could be measured. For further study, characteristics of the instrument can be examined. Besides, the integrated assessment instrument can be designed for other matters in chemistry. By applying the integrated assessment instrument, it is possible to assess not only students' cognitive but also students' skills in a certain period and an effective way.

#### 5. Acknowledgment

The authors would like to express profound thanks and appreciation to the DP2M DIKTI (Directorate of Higher Education), Ministry of Research, Technology and Higher Education Republic of Indonesia through "Tim Pascasarjana" Research Grant 2016 for the financial support.

#### 6. References

- i. Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. *Educational and Psychological Measurement*, 40, 955-959.
- ii. Aiken, L.R. (1985). Three coefficients for analyzing: The reliability and validity of ratings. *Educational and Psychological Measurement*, 45, 131-142.
- iii. Aktamis, H., & Ergin, O. (2008). The effect of scientific process skills education on student' scientific creativity, science attitudes and academic achievements. *Asia-Pasific Forum on Science Learning and Teaching*, 9(1), 1-21.
- iv. Beavers, A.S., Lounsbury, J.W., Richards, J.K., Huck, S.W., Skolits, G.J., & Esquivel, S.L. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research & Evaluation*, 18(6), 1-13.
- v. Beek, K. V., & Louters, L. (1991). Chemical language skills investigating the deficit. *Journal of Chemical Education*, 68 (5), 389-392
- vi. Chebii, R., Wachanga, S., & Kiboss, J. (2012). Effects of science process skills mastery learning approach on students' acquisition of selected chemistry practical skills in school. *Scientific Research*, 3(8), 1291-1296.
- vii. Demircioğlu, G., Ayas, A., & Demircioğlu, H. (2005). Conceptual change achieved through a new teaching program on acids and bases. *Chemistry Education and Practice*, 6(1), 36-51.
- viii. Depdiknas. (2003). Undang-Undang Republik Indonesia Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional. Jakarta: Depdiknas.
- ix. Duran, M., & Sendag, S. (2012). A preliminary investigation into critical thinking skills of urban high school students: Role of an IT/STEM program. *Creative Education*, 3 (2), 241-250.
- x. Hambleton, R. K., & Swaminathan, H. (1985). *Item Response Theory*. New York: Kluwer Inc.

- xi. Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. California: Sage Publications, Inc.
- xii. Karamustafaoğlu, S. (2011). Improving the science process skills ability of science student teachers using I diagrams. *Eurasian Journal of Physics and Chemistry Education*, 3(1), 26-38.
- xiii. Keil, C., Haney, J., & Zoffel, J. (2009). Improvements in student achievement and science process skills using environmental health science problem-based learning curricula. *Electronic Journal of Science Education*, 13(1), 1-18.
- xiv. Kemendikbud. (2013). *Permendikbud Nomor 66 tahun 2013 tentang Standar Penilaian Pendidikan*. Jakarta: Kemendikbud.
- xv. Lord, F. (1980). *Application of item response theory to practical testing problems*. Mahwah: Lawrence Erlbaum Associates, Inc Publisher.
- xvi. Mahapoonyanont, N., Krahomwong, R., Kochakornjarupong, D., & Rachasong, W. (2010). Critical thinking abilities assessment tools: reliability generalization. *Procedia Social and Behavioral Science*, 3(38), 434-438.
- xvii. Marasigan, A. C., & Espinosa, A. A. (2014). Modified useful-learning approach: effects on students' critical thinking skills and attitude towards chemistry. *International Journal of Learning, Teaching and Educational Research*, 1(1), 35-72.
- xviii. Miller, J.M. (2003) *Test validation: A literature review*. Florida: University of Florida
- xix. Moore, B., & Stanley, T. (2010). *Critical thinking and formative assessment*. New York: Eye on Education, Inc.
- xx. National Academy of Sciences. (1996). *National science education standards*. Washington, DC: National Academy Press.
- xxi. Reckase, M. D. (1979). Unifactor laten trait models applied to multifactor tests: results and implications. *Journal of Educational Statistics*, 4(3), 207-230.
- xxii. Retnawati, H. (2014). *Teori respons butir dan penerapannya*. Yogyakarta: Nuha Medika
- xxiii. Rezba, R., Sprague, C., McDonnough, J.T. & Matkins, J.J. (2007). *Learning and assessing science process skills fifth edition*. USA: Kendall/Hunt Publishing Company.
- xxiv. Saribaz, D., & Bayram, H. (2009). Is it possible to improve science process skills and attitudes towards chemistry through the development of metacognitive skills embedded within a motivated chemistry lab?: A self-regulate learning approach. *Procedia Social and Behavioral Sciences*, 1, 61-72.
- xxv. Stephenson, N. S., & Sadler-McKnight, N. P. (2016). Developing critical thinking Skills using the science writing heuristic in the chemistry laboratory. *Chemistry Education Research and Practice*, 17, 72-79.
- xxvi. Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan rasch pada assessment pendidikan*. Jakarta: Trim Komunikata.
- xxvii. Tosun, C., & Taskesenligil, Y. (2013). The effect of problem-based learning on undergraduate students' learning about solution and their physical properties and scientific processing skills. *Chemistry Education Research and Practice*, 14, 36-50.
- xxviii. Utomo, T., & Ruijter, K. (1994). *Peningkatan dan pengembangan Pendidikan*. Jakarta: (Gramedia Pustaka Utama.
- xxix. White, B., Stains, M., Escriu-Sune, M., Medaglia, E., & Rostamjad, L. (2011). A novel instrument for assessing students' critical thinking abilities. *Journal of College Science Teaching*, 40(5), 102-107.
- xxx. Wiberg, M. (2004). *Classical test theory vs item response theory*. Umea: Umea University.
- xxxi. Zhou, Q., Huang, Q., & Tian, H. (2013). Developing students' critical thinking skills by task-based learning in chemistry experiment teaching. *Creative Education*, 4(12A), 40-45.
- xxxii. Zoller, U. (2001). Alternative assessment as (critical) means of facilitating HOCS-promoting teaching and learning in chemistry education. *Chemistry Education Research and Practice in Europe*, 2(1), 9-17.