An Integrated Assessment Instrument: Developing and Validating Instrument for Facilitating Critical Thinking Abilities and Science Process Skills on Electrolyte Solution Matter

Sri Rejeki Dwi Astuti[1,a)], Suyanta[2,b)], Endang Widjajanti LFX[2], Eli Rohaeti[2]

[1]*Department of Chemistry Education , Graduate School Program, Yogyakarta State University, Indonesia*
[2]*Department of Chemistry Education, Faculty of Mathematics and Natural Science, Yogyakarta State University, Indonesia*

[a)]Corresponding author: srirejeki.dwiastuti@yahoo.com
[b)]suyanta@uny.ac.id

Abstract. The demanding of asessment in learning process was impact by policy changes. Nowdays, assessment not only emphasize knowledge, but also skills and attitudes. However, in reality there are many obstacles in measuring them. This paper aimed to describe how to develop integrated assessment instrument and to verify instruments' validity such as content validity and construct validity. This instrument development used test development model by McIntire. Development process data was acquired based on development test step and was analyzed by qualitative analysis. Initial product was observed by three peer reviewer and six expert judgment (two subject matter experts, two evaluation experts and two chemistry teachers) to acquire content validity. This research involved 376 students of X grade from two Senior High Schools in Bantul Regency to acquire construct validity. Content validity was analyzed used Aiken's formula. The verifying of construct validity was analyzed by exploratory factor analysis using SPSS 16. The result show that all item 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 matter.

INTRODUCTION

Assessment is evaluation process for determining quality of learning process. It is impact by policy changes. Assessment in education standard should covered three aspects i.e. knowledge, attitude and skill [1]. The three aspect are very important in chemistry education because the nature of chemstry are chemistry as a process and as a product. Therefore, assessment must covered all aspect in learning process. It called authentic assessment.

Education product are consist of declarative knowledge and procedural knowledge [2]. Declarative knowledge is cognitive knowledge including fact, principle, concept, theory and law. Whereas procedural knowledge is process skills oriented knowledge or called science experience [3]. Chemistry learning can't separate from the developing students'cognitive and students' skills. Process skills is essential for understanding and applying chemistry concept. So, assessment in chemistry must covered knowledge and process skills too [4].

Chemistry is branch of science that applying higher order thinking skills in each matter and holding important role in the development students' critical thinking skills [5]. It is easy fastened to dailylife phenomena. Furthermore, it can be applicated in society. One of chemisty matter in senior high school is electrolyte solution matter. Electrolyte solution matter is contextual matter, so it can be easy fastened and applied in dailylife. Many electrolyte solution phenomena can be find in daylilife such as storage battery application and body requirement for electrolyte solution. It can develope students' critical thinking skills. Moreover, electrolyte solution matter is electrochemistry prerequisite matter.

Critical thinking is one of thinking skills in higher order thinking skills which become one of learning goal in 21th century [6]. Critical thinking abilities can help students for solve problem efficiently and way to self learning [7]. Critical thinking as thinking process is thinking ability to make personal evaluation about problem based on authenticity, accuracy, process, theory, method, background and then make reasonable decision [8]. Critical thinking is essentiall way on scientific investigation process, especially fo analyed and evaluated scientific evidances [9]. Critical thinking covered concrete and abstract thinking process to make conclusion about fact and problem which appropriate with scientific evidences [10]. Critical thinking abilities can emerged when students used higher order thinking abilities, so students'learning result enable to increased.

Students who has critical thinking abilities, wiill has high academic achievement. The result of Wenglinskys' research explained that students who was familiarized to think critically in learning process, be able to aquire high test score on learning evaluation [11]. Critical thinking abilities not only necessary in school, but also in dialylife. In fact, drawing conclusion very needed in all occupation section, so critical thinking

abilities will raise person's success in work force [11]. Moreover, increasing job requirement at worker who can analyzed and applying information critically and give efective solution about problem in work force [12]. Therefore, critical thinking abilities is one of educations' goal in globalization era, so student have to familiarize used critical thinking abilities in chemistry learning.

Critical thinking covered three aspect i.e. to identify problem, to reconstruct argument and to evaluate argument [13]. Clasification of critical thinking by Watson-Glaser consist of defining a problem, determining possible solution and strong argument, drawing valid conclusion based on regarding solution and evaluating conclusion [6]. Students' critical thinking abilities seen from students arguments toward problem and conclusion toward arguments. Therefore, critical thinking indicators that used on developing integrated assessment instrument include (1) identifying problem, (2) reconstructing arguments, (3) determining solution, (4) evaluating arguments, dan (5) drawing conclusion.

Science process skills are important learnng ways for reaching knowledge [14]. One of science educations' goal is to prepare students for accumulating knowledge and to discover scientific knowledge by self [15]. Science process skills is a way to produce and to apply scientific information on scientifif research, and to solve problem [16-17]. Process skills was skills that scientist do to discover scientific knowledge and to investigate problem [18]. Science process skills is skills to discover knowledge, to define problem and to formulate problem [19]. Based on several different opinion about science process skills, it can be defined as a skills that used to discover knowledge by identiying problem, solving problem, and answering question about nature.. Science process skills are believed to ensure students' meaning knowledge because it help students to increase higher order thinkings. Therefore, developing science process skills for students is importent to be noted

Based on science process skills' level, it consist of two level skills i.e. basic science process skills and integrated science process skills [18,20]. Basic science process skills is fundamental skill that underlie scientific method. It consist of six process skills i.e. observing, comunicating, clasifying, measuring, inferring and predicting. Integrated science process skills is development of basic science process skills. It consist of 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 experimentating.

Based on science process skills' type, it consist of cognitive skills and sensorimotor skills [21]. Cognitive skills (soft skills) are process skills related to students' thinking process in learning process. It consist of clasifying, predicting, inferring, comunicating, constructing hypotheses, processing data, identifying variable, defining varaibles operationally, constructing a table of data, constructing a graph, analyzing investigation, describing relationship between variables, and designing experiment [22]. Sensoromotor skills (hard skills) are process skills related to body movement skills for discovering knowledge. It consist of observing, measuring, acquiring data and esperimentating [22].

Science process science indicators that used on developing integrated assessment instrument including (1) observing, (2) comunicating, (3) clasifying, (4) predicting, (5) inffering, (6) data analyzing, (7) constructing a table of data, and (8) designing experiment. Indicator selection was based on cognitive skills (science process skills that related to thinking process) in science process skills, so it can be measured using written test.

Now, learning process only emphasize to cognitive aspect, whereas process skills less get attention [23]. Based on interview result toward senior high school chemistry teachers in Bantul Regency indicate that comprehensive learning assessment haven't optimally done yet. Teachers still sets out assessment toward students' knowledge, whereas assessment toward science process skills only based teachers' subjective assumption. Furthermore, assessment toward students' critical thinking abilities on chemistry learining haven't got attention yet from them. However, assessment of learning result must be work simultaneously. Student who have a good science process skills in learning process, will have better cognitive abilities [24]. Nowdays, cognitive assessment and skills assessment are carried out separetely [20, 25-26] If it is caried out separately, so relationship between cognitive and skills can be acquired. Therefore, integrated assessment instrument is needed for measuring critocal thinking abilities and science process skills in one instrument. The instrument development goal are to facilitate teacher for asessing the learning result effectively and eficiently and to develope valid instrument.

RESEARCH METHODS

In this study, instrument test development model by McIntire was adopted to develop integrated assessment instrument. The subject in this study were 376 student at X grade from two Senior High School in Bantul Regency that have enrolled electrolyte solution matter. Sample determining used purpossive sampling technique based on school rank of Acceptance Student Data in Bantul Regency and school that has implemented 2013 curriculum in learning process.

To develop integrated assessment instrument, this study adopted test development model by McIntire that consist of 10 steps, i.e defining the test universe, audience and purpose; developing a test plan; composing the test items; writing the administration instructions; conducting piloting test; conducting item analysis; revising the

test; validation the test; developing norms; and complete test manual. In this study only used steps till conducting item analysis because only done till analysis of instruments' item.

Data were acquired in this research including 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 expert judgment whereas construct validity were acquired by instrument field testing. Collecting data instrument are questionnaire and integrated assessment instrument. Questionnaire is item content validity sheet for expert. It is used for verifying item content validity on integrated assessment instrument. Integrated assessment instrument is instrument for instrument field testing. It is used for verifying construct validity on integrated assessment instrument.

Data analysis techniques are qualitaive data analysis and quantitative data analysis. Qualitative data analysis was used for analyzing development process data. It were expert sugesstion for finishing product. Quantitative data analysis was used for estimating content validity and construct validity. The content validity data was analyzed with Aikens' Formula for calculating content validity coefficient [27]. A coefficient content validity may be define as

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

Explanation:

s = r – lo

Io = 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

The construct validity data was analyzed with SPSS software (ver.16.0). Exploratory factor analysis were used to estimate construct validity of item.

RESULT AND DISSCUSSION

The developing product of research is integrated assessment instrument for mesuring critical thinking abilities and science process skills of senior high school students on electrolyte solution matter. The writing item was based on the definition of the grating electrolyte solution matter contained in basic competence in 2013 curriculum. Each item consist of learning indicator, critical thinking indicator, and scienc process skills indicator. The initial draft of integrated assessment instrument was essay test including 8 items with 28 subitems. Futhermore, the composing skoring manual was appropriated with difficulty item level and student thinking groove.

Scoring model of the assessment is Partial Credit Model 1-Parameter Logistic (PCM 1-PL). 1 parameter logistic is difficulty level parameter. The choosing PCM 1-PL was based on its ability to analyze item which have different maximal score. The difference was caused each item has the different item quality to measure student ability and has the different steps to solve item. After that, initial product was reviewed by three peer reviewer and six expert (two subject matter experts, two learning evaluation experts, and two chemistry teacher).

Item reviewed was committed for validating initial product especially item content validity by expert. The estimating content validity was considered by suitability between learning indicators with critical thinking indicators and science process skills indicators. The initial product was validated by two subject matter experts, two learning evaluation experts, and two chemistry teacher who are skilled in composing test. The validating initial product aimed to verify item content validity.

The result of validating data by expert were analyzed with Aikens' formula. It used to determine valid items or invalid items based on content validity coefficients (Aikens' V).

Table 1. Content Validity Coefficient In Integrated Assessment Instruments

No	Item	Aikens' V
1	1	0.83
2	2	0.94
3	3	0.83
4	4	1.00
5	5	0.83
6	6	0.94
7	7	0.89
8	8	0.89

To know content validity coefficient significance statistic, can be determine by correlating ratings category with amount of raters [28]. This research involved six raters and four ratings category. In 0.05 significance level, allowed minimum content validity coefficient (Aikens' V) was 0.78 [28]. As indicated on Table 1, all of items has Aikens' V more than 0.78. So, it can be explained that with 0.05 significance level, all

of items in integrated assessment instrument was asserted valid. Therefore, all of item in integrated assessment instrument be able to measure critical thinking abilities and science process skills of senior high school students on electrolyte solution matter.

The result of item review by experts are not only acquired validating data, but also qualitative data i.e experts' sugesstion. Experts' suggestion was linguistic aspect sugesstion especially word selection. It was used to make product perfectly. The involvement of lecturers as expert in validating product was aimed to gain suggestion about depth of matter and suitability between scoring manual with test instrument. The involvement of chemistry teachers in validating product was aimed to gain sugesstion about visibility and effectiveness implementation product in learning process such as time allocation, amount of ideal item in essay test, and product readability. Moreover, teacher was considered having experience in composing test. Therefore, teacher can give some suggestion about language using, words and composing a good sentences.

Based on the result of validating product, it was acqured amount of proper item question that used in learning process as many 8 item with 21 subitem, whereas 7 subitem was dropped. It was committed because of teachers' consideration that amount of item question more than 25 subitem is too many for examining during 2 hour lesson (90 minutes). The other words, this dropped subitem have similarity meaning with other subitem. The drooped subitem choose based on teachers' suggestion. The good time for students to finish the test ranging from 1,5 hour to 2,5 hour because if it more than 2,5 hour will cause declining students' thinking endurance [29]. If it happen, it will cause declining tests' reliability

The emendation of instrument was committed on properity if concept, writing technique, and selection words. Sentence is early step to comprehend question, so it was very influential for students' comprehension toward question in the test [30]. Though there were a dropped subitems, but 8 item questions has covered all learning indicators. Futhermore, all of item question in integrated assessment instrument was ordered and was arranged again become valid instrument according to content validity.

The valid instrument according to content validity was used in field testing. It aimed to verifying construct validity. Construct validity was called unidimention testing. Its' purpose was to know the instrument only measure one factor or dimension. The result of field testing data were analyzed with exploratory factor analysis using SPSS software (ver.16.0).

Before interpretating the result of construct validity, it was committed sample adequancy analysis using Kaiser-Mayer-Olkin Measure of Sampling Adequancy testing (KMO-MSA) and Bartlett testing. KMO-MSA testing was used to know what sample have adequated for factor analysis. Bartlett testing was used to determine presence correlation between variables.

Table 2. The Result of KMO-MSA Testing dan Bartlett Testing

Kaiser-Mayer-Olkin Measure of Sampling Adequacy		0,605
Bartlett's Test of Sphericity	**Approx. Chi-Square**	134,964
	Df	28
	Sig.	0,000

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 [31]. As indicated on Table 2, the value of KMO-MSA was more than 0.5 and the significance of Bartlett testing must be less than 0.05. So, it can be explained that sample has adequated for factor analysis. Therefore, the interpretation result can be continued with construct validity analysis or unidimension assumption.

Futhermore, factor analysis was continued with interpreting eigen value from correlation variant-covariant matrix.

Table 3. The Result of Initial Eigenvalue from Correlation Variant-Covariant Matrix

Item	**Initial Eigenvalues**		
	Total	**% of variance**	**Commulative %**
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

Factor analysis was used to analyze relationship intervariable using correlation testing, so it was get a new variable named factor. As indicated on Table 3, students' respond data towards the instrument lade three eigen value (eigen value > 1). According to Kaiser criteria that the instrument lade three factors, but there was dominant one factor [31]. 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 test consist of mathematics abilities and nonmathematics

abilities or languagen abilities [32]. Whereas two factor was measured by the instrument including personality factor and administrative factor during the test such as anxiousness and students' motivation [33].

The result of factor analysis was presented in scree plot for eigen value visualization. As indicated on Figure 1, eigen value began slightly on third eigen valuee. So, there was a dominant factor which was measured by integrated assessment instrument and two another factor also gave contribution towards instrument respondses.

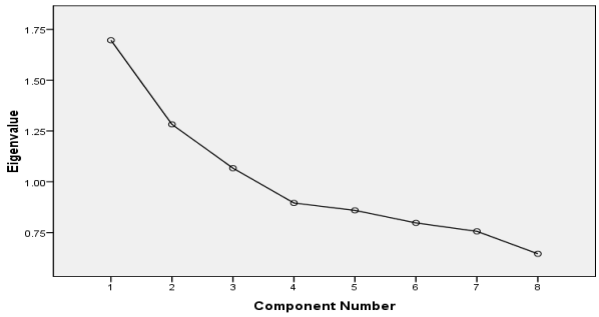

Figure 1 . Scree *Plot*

Unidimension assumption was very difficult to fulfilled ideally [34]. However, unidimension assumption can be considered to fulfilled if test contained one dominant factor [35-36]. If the result of factor analysis refered that first factor has comulative percentage more than 20%, so unidimension has fulfilled [37-39]. As indicated on Table 3, comulative percentage of first factor was 21.207%, so it can be stated that unidimension assumption has fulfilled. If unidimension assumption has fulfilled, construct validity has fulfilled too. Therefore, the integrated assessment instrument has proven to be valid according construct validity, because it only measure one knowledge dimension i.e chemistry knowledge.

CONCLUSION

Based on the result of study, we can conclude that the integrated assessment instrument on electrolyte solution matter was asserted valid according content validity and construct valdity. So, it was suitable for measuring critical thinking abilities and science process skills on electrolyte solution matter.

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.

REFERENCES

1. Kemendikbud, Permendikbud Nomor 66 tahun 2013 tentang Standar Penilaian Pendidikan (4 June 2013).
2. R.J. Marzano, *The Art of Science Teaching* (ASCD, Alexandria, 2007), p.59.
3. Z. H. Ismail and I.Jusoh, Journal of Science and Mathematics Education in S.E Asia **xxiv**, 67-77 (2001).
4. A.C. Marasigan and A.A.Espinosa, International Journal of Learning, Teaching and Educational Research **1**, 35-72, (2014).
5. Q. Zhou, Q.Huang and H. Tian, Crative Education **4**, 40-45 (2013).
6. M. Duran and S. Sendag, Creative Education **3**, 241-250 (2012).
7. N. Mahapoonyanont, R. Krahomwong, D. Kochakornjarupong, and W. Rachasong, Procedia Social and Behavioral Science **3**, 434-438 (2010).
8. Z. Qing, S. Ni, and T. Hong, Procedia Social and Behavioral Science **2**, 4561-4570 (2010).
9. B. White, M. Stains, M. Escriu-Sune, E. Medaglia and L. Rostamnjad, Journal of College Science Teaching **40**, 102-107 (2011).
10. Arslan, The Malaysian Online Journal of Educational Science **3**, 1-10 (2015).
11. B. Moore and T. Stanley,.*Critical Thinking and Formative Assessment* (Eye on Education Inc, New York, 2010), pp.16-21.
12. N.S. Stephenson and N.P. Sadler-McKnight, Chemistry Education Research and Practice **17**, 72-79 (2016).
13. T. Bowel and G. Kemp, *Critical thinking: A concise guide* (Routledge, London, 2002), p.6.
14. R.A. Rauf, M.S. Rasul, A.N. Mansor, Z. Othman and N. Lyndon, Asian Social Science **9**, 47-57 (2013).
15. J. Li and D. Klahr, "The psychology of scientific thinking: Implications for science teaching and learning,"in *Teaching Science in the 21st Century*, edited by J. Rhoton, & P. Shane (NSTA Press, Arlington, 2006), p.2.
16. W. Harlen, Assessment in Education **6**, 129-140 (1999).
17. J. Huppert, S.M. Lomask and R. Lazarorcitz, Journal of Science Education **24**, 803-821 (2002).

18. R.J. Rezba, C.R. Sprague, J.T. McDonnough and J.J. Matkins, *Learning & Assessing Science Process Skills Fifth Edition.* (Kendal/Hunt Publishing Company, Dubuque, 2007).
19. A. Aydin, International Journal of Education and Practice **1**, 51-63 (2013).
20. C. Keil, J. Haneyand J. Zoffel, Electronic Journal of Science Education **13**, 1-18 (2009).
21. B. Subali, Cakrawala Pendidikan **xxx**, 130-144 (2011).
22. M.M. Sheeba, Educationia Confab **2**, 108-123 (2013).
23. U. Zoller, Chemistry Education Research and Practice in Europe **2**, 9-17 (2001).
24. A.C. Austin, H. Ben-Daat, M. Zhu, R. Atkinson, N. Barrows and I.R. Gould, Chemistry Education Research and Practice **1**, 168-178 (2015).
25. D. Saribaz and H. Bayram, Procedia Social and Behavioral Sciences **1**, 61-72 (2009).
26. C. Tosun and Y.Taskesenligil, Chemistry Education Research and Practice **14**, 36-50 (2013).
27. L.R. Aiken, Educational and Psychological Measurement **40**, 955-959 (1980).
28. L.R. Aiken, Educational and Psychological Measurement **45**, 131-142 (1985).
29. T. Utomo and K. Ruijter,. *Peningkatan dan pengembangan Pendidikan* (Gramedia Pustaka Utama, Jakarta, 1994), p.60.
30. K.V. Beek and L. Louters, Journal of Chemical Education **68**, 389-392 (1991).
31. A.S. Beavers, J.W. Lounsbury and J.K. Richards, Practical Assessment, Research & Evaluation **18**, 1-13 (2013).
32. F. Lord, *Application of Item Response Theory to Practical Testing Problems* (Lawrence Erlbaum Associates, Inc Publisher, Mahwah, 1980), p.20.
33. H. Retnawati, *Teori Respon Butir Dan Penerapannya* (Parama Publishing, Yogyakarta, 2014, p.1.
34. R,K, Hambleton and H. Swaminathan, *Item Response Theory* (Kluwer Inc, Boston, 1985), p.16.
35. R.K. Hambleton, H. Swaminathan and H.J. Rogers, *Fundamentals of Item Response Theory* (Sage Publications Inc, California, 1991), p.56.
36. M. Wiberg, *Classical Test Theory vs Item Response Theory* (Umea University Press, Umea, 2004), p.5.
37. M.D. Reckase, Journal of Educational Statistics **4**, 207-230 (1979).
38. N. Smits, P. Cuijpers and A. van Straten, Psychiatry Research **188**, 147-155 (2011).
39. Q. Wu, Z. Zhang, Y. Song, Y. Zang, Y. Zhang, and F. Zhang, International Journal of Advancements in Computing Technology **5**, 209-216 (2013).