Content Validity and Scoring of Two Tier as Measuring Instrument of Science Process Skills for Knowledge Aspects in Chemistry Learning

Sri Yamtinah[a], Sulistyo Saputro[a], Haryono[a], Budi Utami[a]

Sebelas Maret University, Jl. Ir. Sutami, Surakarta, Indonesia

Corresponding e-mail: jengtina_sp@yahoo.com

Abstract: Problems in chemistry learning related to the three levels, namely the macroscopic, sub-microscopic, and symbolic (representational) (Barke, 2009: 27). During chemistry learning, student's abilities to understand the roles of each representation level and transfer to another level is an important aspect for produce an explanation that can be understood. The curriculum of 2013 mandates that the essence of a scientific approach to learns through the Science Process Skills. Two-tier used as an instrument for measuring knowledge aspect because its abilities to analyze the cognitive skills of the students. Through two-tier, students just not pick the answer on items but also provide a reason for the answers are chosen so the cognitive processes can be observed and measured. As an instrument to measure the Science Process Skills, the fulfillment validity of the content becomes important. This two-tier instrument has good content validity between 0.78 to 1.0. Therefore, the scoring system that used in the examination students' answer sheets is important part of this instrument to analyze science process skills student.

Keywords: science process skills, two-tier instrument, scoring system, content validity

1. INTRODUCTION

Chemistry is a science that is acquired and developed based on experiments that seek answers to the questions of what, why, and how natural phenomena; especially with regard to the composition, structure and properties, transformation, dynamics, and energetics substances. Therefore, subjects in the high school chemistry aims to learn everything about the composition, properties and structure, transformation, dynamics, and energetics of substances and their application to solve everyday problems. Chemistry is a product (chemical knowledge in the form of facts, concepts, principles, laws and theories) the findings of scientists and process (scientific work).

Chemical knowledge learned at three levels, namely the macroscopic, sub-microscopic, and symbolic *(representational)* (Barke, 2009: 27). Relations between the levels must be taught explicitly. Relations between the three can be described on the figure 1:
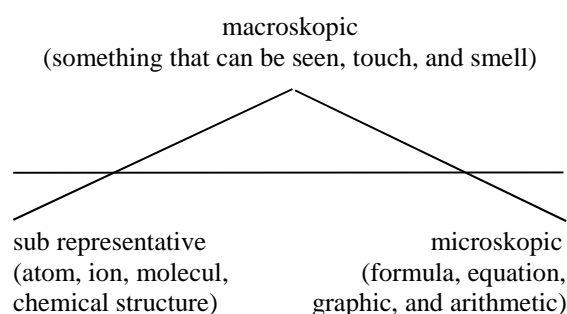
Figure 1. Three Level in Chemistry Sciences (Barke, 2009: 27)

Interaction and differences between the three levels is an important characteristic in chemistry learning and it is necessary to understand the concepts of chemistry. Observed phenomenon "corrosion of nails", is an example of the concept of chemistry at the macroscopic level. To better explain the phenomenon, the chemists to develop "models of atoms" and "molecules concept". While the level submicroskopic, explained that "the corrosion of nails" is a chemical process in which the surface of the iron reacts with oxygen in the air and produce iron oxide molecules. Another way to explain the concept of iron corrosion is through the equation with symbols, formulas and figures, namely: $4Fe(s) + 3O_2(g) \quad 2Fe_2O_3(s)$. In studying chemistry, the ability of students to

PROSIDING ICTTE FKIP UNS 2015
Vol 1, Nomor 1, Januari 2016
Halaman:

ISSN: 2502-4124

ICTTE
FKIP UNS 2015

understand the role each level of representation and transfer from one level to another level is an important aspect to produce an explanation that can be understood.

Basic materials in a very adrift chemistry between the materials with other material. Although not as strict as in mathematics and physics, in chemistry learning, the order of the concept is of considerable concern. For example before the learners master the concepts of stoichiometry students must first master the concept of the mole as well as learners must master the concept of reaction rate in advance in order to master the concept of chemical equilibrium. Similarly, before the students learn the concepts of chemical bonding in advance learners must master the concepts of atomic structure. Therefore, the steps learners in working on the problems of chemical must be ordered according to a hierarchy of existing concepts.

Curriculum 2013 has mandated the essence of a scientific approach to learning. The scientific approach is believed to be the golden bridge and the development of attitudes, skills, and knowledge of learners. Scientific method refers to the techniques of the investigation of phenomena or symptoms, acquire new knowledge, or correcting and integrating previous knowledge. To be called scientific, the search method *(method of inquiry)* should be based on evidence of the object observable, empirical and measurable principles specific reasoning.

The scientific approach that is intended in the curriculum of 2013 learning to approach scientific, namely learning that consists of the activities observed (to identify the things they want to know), formulate questions (and formulate hypotheses), try / collect data (information) with various techniques, associate / analyze / process the data (information) and to draw conclusions and communicate the results of which consists of a conclusion to acquire the knowledge, skills and attitudes.

Skills said as the ability to use the mind, reason, and act efficiently and effectively to achieve a particular result. Process is defined as a device that scientists use complex skill in

conducting scientific research. The process is a great concept that can be broken down into components that must be mastered if someone will do the research.

Skills process according Rustaman (2003) is a skill that involves cognitive skills or intellectual, manual and social. Cognitive skills involved because by doing process skills of students using thoughts. Manual dexterity skills clearly involved in the process because they involve the use of tools and materials, measurement, preparation or assembly tools. Social skills are also involved in the process skills as they interact with each other in carrying out the teaching and learning activities, for example, discuss the observations. Process skills need to be developed through direct experiences as a learning experience. Through direct experience, can be more feel the live process or activity is being conducted.

According Dahar (2012), science process skills is the student's ability to apply scientific methods to understand, develop and discover science. Science process skills is very important for every student in preparation to use scientific methods in developing science and are expected to acquire new knowledge or develop the knowledge that has been owned. Science process skills is the complex skills used by scientists to conduct scientific investigations into the series of the learning process.

The success of the learning process can be seen from the assessment conducted on learners. The assessment used in the learning santifik approach must be able to measure the skills in question, it is important to develop assessment instruments Science Process Skills.

One type of test that can be used to measure the skills Process Science (PPP), particularly in the aspect of knowledge (cognitive) is a *two-tier test*. Through the test instrument *two tier* where learners have to give a reason for choosing an answer, the teacher will be able to detect the ability of learners real. These instruments form an objective test two levels, the first question in the form of questions about both the main and the reasons for selecting the answer. W ith so teachers will be able to detect the actual ability of learners. These instruments form an objective

test two levels, the first question in the form of questions about both the main and the reasons for selecting the answer. Some researchers have developed an instrument types of problems grounded the model of different reasons. Some models are: (1) provide an opportunity to test participants to justify such a description answer, (2) provide an opportunity to test participants to choose an alternative reasons that have been prepared, (3) the reasons in two stages, the first reason in the form of multiple choice, continued The next reason for such a description. One example of *a two tier* item :

NaOH and HI will react and produce NaI and water.
Right theory for the reaction was….
A. Salt that formed can be hydrolysis in water
B. Salt that formed can not be hydrolysis in water
C. Salt that formed has acid criteria
D. Salt that formed has base criteria
E. Salt was forming from base and acid

Alasan dari jawaban saya adalah :
A. NaI from strong acid HI and strong base so will be ionized to Na^{+2} and I^-, that two ion can not react in water, so that the solution still netral.
B. NaI from weak acid HI and strong base so will be ionized to Na^{+2} and I^- that two ion will react in water that is hydrolyisis solution.
C. Na^{+2} and I^- , two ion can not react in water, that is hydrolyisis solution

---

As an instrument used to measure the Skills Process of Science, the fulfillment of the validity of the content becomes imperative characteristic of grain developed. The validity of the contents are intended to determine how attributes measured by an instrument in accordance with the performance developed. The validity of the content based on how far the test contents to describe the domain behavior or attribute to be measured in such tests. To obtain a high content validity, a test should be designed based on specific attributes are described with indicators which are then described in those items that really describe these attributes.

To obtain the content validity, the necessary rational analysis of experts in the field who developed the instruments or *professional judgment.* In this study, content validity was determined by using a formula Aiken, namely:

$V = S / [n * (c-1)]$ where $S = \sum n_i (r - r_o)$ (Aiken, 1985: 955)
Where:

V : Validity index of Aiken	n_i : Number of assessors *(raters)* which chose criteria i
c : Number of categories /criteria	r : Criteria to i
r_o : The lowest category	n : The total value appraiser

V value ranges at 0-1 and the criteria used to declare an item is said to be valid in the contents on the amount of rater (assessor) of 9 was 0.78 (Aiken, 1985). The instrument has met the content validity means rationally, instrument it deserves used to measure the performance to be measured.

Graded Response Model (GRM) is a model developed to handle the scoring on those items politomus (De Ayala, 1993). The use of GRM appropriate when the response examinees against the grain can be classified as a response category sequence and level of completion is likely to increase. That is by using a sequential response and completion rates are rising or in other words, the second step requires a prerequisite step one, and so on until the final settlement.

With the scoring procedure will be used to analyze the ability of learners. Additionally, it will be in the detection lies the weakness of students so that in addition can be used to measure the Science Process Skills in the aspect of knowledge, it can also be used as a diagnostic function.

2. RESEARCH METHODS

This study is part of research development using the mode l Akker with the aim of developing an instrument Skills Process of Science. In the early stages of research conducted *Focus Group Discussion* (FGD) as preliminary *(preliminary investigation)* which is intended for analysis of needs *(needs assessment)* on the field about the chemistry teacher's ability to manage the assessment of learning. A preliminary study conducted on teachers, chemistry Surakarta and Karanganyar incorporated in MGMP

Chemistry. Analysis of preliminary data from a study conducted with descriptive techniques.

The second phase of this study is a literature review about the procedure of preparation of models of *two-tier* test instruments and scoring procedures using the system politomus model *Graded Response Model* (GRM). *Graded Response Model* (GRM) is a model developed to handle the scoring on politomus item (De Ayala, 1993). The use of GRM right when examinees response against the grain can be classified as category sequential response and level of completion is likely to increase. That is by using the response sequence and completion rates are rising or in other words, the second step requires a prerequisite step one, and so on until the final settlement. Scoring models GRM chosen to reward the thinking processes of learners, so that each stage of the process of thinking learners gain certain score (WAINER, H & Kiely, GL (1987).

3. DISCUSSION

Analisis needs *(needs assessment)* in the field of chemistry teacher's ability to manage the assessment of learning. This needs analysis is done with descriptive method is to capture the experience of teachers through questionnaires and interviews of teachers in MGMP Chemistry in Solo and Karanganyar. Some results of the needs analysis as follows:

a. Most of the teachers (94.73%) use the test to measure the ability of learners.
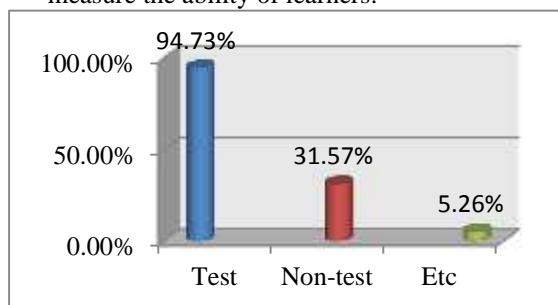
b. A total of 89.47% of teachers use essay tests and 68.42% using a multiple-choice test to measure the ability of students. Through this data it appears that teachers use a combination of essay and multiple-choice tests.
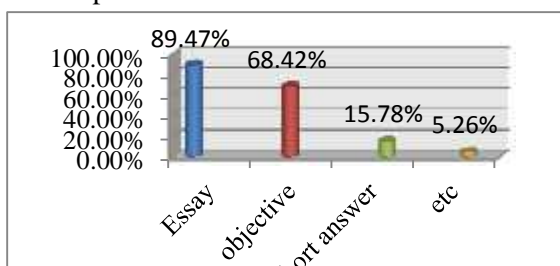
c. Opinions about the weakness of teachers form a multiple choice test, as in the following table:

Answer option	Percentage of voters (%)
a. Many students were speculative	36.84
b. The amount of the possibility of cheating	21.05
c. Students do not master the overall concept	10.52
d. Can not measure with a valid	42.10
e. Score is only true / false	5.26

Based on the opinion of the teacher, then the teacher actually understand about the weakness of multiple choice tests does.

d. Weakness form of essay tests in the opinion of the teachers are:

Answer option	Percentage of voters (%)
a. Takes a long time to answer	10.52
b. A limited number of questions	26.31
c. Only includes little material	36.84
d. Take a long time to correct	26.31
e. Subjectivity	10.52

Teachers also understand the weaknesses of the use of test types of essays, especially the weakness of the range of material items and long correction time.

A total of 94.73% of teachers expressed the need to develop the type of test in addition to multiple-choice and essay can measure the ability of learners to be more efficient without abandoning the purpose of measurement is to know closely the ability of learners.

By the time the teachers introduced the form of *two-tier* test for measuring aspects of knowledge (cognitive), there is the opinion of teachers as follows:

Answer option	Percentage of voters (%)
a. Good, could be developed	26.31
b. Can measure students' cognitive abilities	36.84
c. High difficulty level, because students must analyze	15.78

Based on the table it appears that teachers are interested to learn this type of test this further, then later researchers together with teachers to develop joint instrument Skills Process Science mainly on the type of test *two tier* for measuring aspects of knowledge (cognitive). In general, the majority of teachers revealed that *two tier* test instrument that refers to the science process skills will be good for students. Problem science process skills by teachers can train students to think of analysis, and can develop their skills in the field of chemistry.

Those items are arranged based on indicators agreed upon teachers in *the Focus Group Discussion* (FGD). Validation is done by lecturers who acted as an expert, in addition to the lecturer validation is also performed by high school teachers. Validation of the expert and the teacher is then used to validate the contents, where in addition to the inputs used to revise the test items, also used to calculate the validity of the content of every item by applying the formula Aiken. Aiken validation results that the entire instrument has index > 0.78, which means to qualify as an instrument that has good content validity.

Limited trial results obtained all student participants limited scale trials have not been working on the model of Two-Tier test. But seen from the given problem, most of the students stated that the questions were given easy to understand, only a few problems there are insufficient data is complete and there are some questions that have no answers on the options given, so it needed some improvements in matters The. Judging from the level of difficulty of questions, there are some questions that pertained difficult so that the process takes longer than the other questions. The time given for working on Two-Tier Test for 90 minutes according to the student's enough salt hydrolysis matter the number of items is done as many as 20 items two-tier test.

The scoring system uses a model GRM because scoring is done with a model for mem eriksa politomi cognitive processes that occur in learners. P enskoran using GRM can be seen in the following table:

No.	Aspects of assessment	Score
1	Answer wrong-wrong reasons (SS) or did not answer	0
2	Answer wrong-reasons correctly (SB)	1
3	The answer was wrong reason (BS)	2
4	The answer is correct reason (BB)	3

GRM using scoring models, in addition to the final values obtained as the ability of learners, can also be traced back to an understanding as science process skills possessed by learners.

4. CONCLUSION

Analysis needs of the high school chemistry teachers in Surakarta and Karanganyar found that teachers use forms of essay and multiple-choice tests on learning. The teacher looked at the need to develop forms of instrument that can be used efficiently and effectively in measuring the ability of learners. Through FGD that the material is appropriate KPS instruments developed for the material stoichiometry Hydrolysis of class X and XI. Results of the validation of the contents of the instrument using the KPS formula is obtained that the entire instrument Aiken has a validity index Aiken> 0.76 so that the instrument has good validity Aiken. Limited trial results obtained throughout learners trial participants limited scale never before working on *Two-Tier* model *test.* Some inputs associated with data in several questions that needed to complete work on the problems but has not written on the matter. Provided sufficient time to work on the problems. Methods of scoring with *Graded Response Model* (GRM) provides an opportunity to analyze the ability of teachers of Science Process Skills (KPS) owned learners.

5. ACKNOWLEDGEMENT

6. BIBLIOGRAPHY

Abin Syamsudin. (2002) *Psikologi Pendidikan: Perangkat Sistem Pengajaran Modul*. Bandung: PT Remaja Rosdakarya

Aiken, L.R. (1985). Three Coefficient for Analyzing the Reliability and Validity of Ratings. *Educational and Psychological Measurement*, 45

Barke,H.D, Al Hazari & Yitbarek,S (2009). *Misconceptions in Chemistry, Addressing Perceptions in Chemical Education.* German: Springer-Verlag Berlin Heidelberg

De Ayala, R.J. (1993). An Introduction to Polytomous Item Response Theory Models. *Measurement and Evaluation in Counceling and Development, 25, 127-189*

Embretson, SE & Reise, SP. (2000). *Item respon theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates

Oriondo, L.L., & Dallo-Antonio, E.M. (1984). *Evaluating educational outcomes (tesis, measurement and evaluation)*. Manila: REX Printing Company, INC.

Wainer, H & Kiely, G.L (1987). Item cluster and computerized adaptive testing: a case for testlet. *Journal of education measurement, 24, 185-201*

Wainer, H & Lewis, C. (1990). Toward a psychometrics for testlet. *Journal of Education, 27: 1-14*

Abin Syamsudin. (2002) *Psikologi Pendidikan: Perangkat Sistem Pengajaran Modul*. Bandung: PT Remaja Rosdakarya

Aiken, L.R. (1985). Three Coefficient for Analyzing the Reliability and Validity of Ratings. *Educational and Psychological Measurement*, 45

Barke,H.D, Al Hazari & Yitbarek,S (2009). *Misconceptions in Chemistry, Addressing Perceptions in Chemical Education.* German: Springer-Verlag Berlin Heidelberg

De Ayala, R.J. (1993). An Introduction to Polytomous Item Response Theory Models. *Measurement and Evaluation in Counceling and Development, 25, 127-189*

Embretson, SE & Reise, SP. (2000). *Item respon theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates

Oriondo, L.L., & Dallo-Antonio, E.M. (1984). *Evaluating educational outcomes (tesis, measurement and evaluation)*. Manila: REX Printing Company, INC.

Wainer, H & Kiely, G.L (1987). Item cluster and computerized adaptive testing: a case for testlet. *Journal of education measurement, 24, 185-201*

Wainer, H & Lewis, C. (1990). Toward a psychometrics for testlet. *Journal of Education, 27: 1-14*