THE DEVELOPMENT OF THREE TIER DIAGNOSTIC TEST TO IDENTIFY STUDENT MISCONCEPTION IN CHEMICAL BONDING ON 10TH GRADER

Annisaa Cahya S. and I Gusti Made Sanjaya

Jurusan Kimia, Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Negeri Surabaya

E-mail: Annisaa.Harmanto@gmail.com

Abstrak

Telah dilakukan penelitian pengembangan tes diagnostik *three tier* untuk mengidentifikasi miskonsepsi siswa pada materi Ikatan Kimia dengan menggunakan program PHP-MySQL. Penelitian ini dilakukan dengan tujuan untuk mengetahui kelayakan tes diagnostik *three tier* berdasarkan kriteria isi, konstruk dan bahasa. Selain itu untuk mengetahui reliabilitas dan efektivitas tes diagnostik *three tier* berdasarkan hasil uji coba terbatas. Jenis penelitian ini adalah pengembangan dengan menggunakan model 4-D terdiri atas empat tahap yaitu pendefinisian *(define),* perancangan *(design),* pengembangan *(develop)* dan penyebaran *disseminate),* namun penelitian ini hanya dibatasi hingga tahap pengembangan. Uji coba dalam penelitian ini dilakukan dua kali di SMAN 18 Surabaya, uji coba I dilakukan di kelas X MIA 5 dengan jumlah siswa 20 orang untuk mendapatkan hasil berupa tingkat kesukaran, daya beda, dan reliabilitas butir tes. Uji coba II dilakukan di kelas X MIA 3 dengan jumlah siswa 10 untuk mengetahui letak miskonsepsi siswa dan efektivitas tes diagnostik *three tier* dalam mengidentifikasi miskonsepsi siswa. Hasil dari validasi butir tes berdasarkan kriteria isi, konstruk dan bahasa berturut-turut adalah 82%, 83% dan 80,75%, sehingga dikategorikan sangat layak. Validasi media 456omputer tes diagnostik *three tier* pada kriteria 1 tentang tampilan media sebesar 83% dan pada kriteria 2 tentang kemudahan penggunaan media sebesar 85,83%, sehingga dikategorikan sangat layak. Sedangkan tingkat reliabilitas yang diperoleh adalah 0,968 yang berarti sangat tinggi dan nilai efektivitas media sebesar 67,12% yang berarti efektif.

Kata Kunci:Tes diagnostik *three tier*, ikatan kimia, miskonsepsi, kelayakan

Abstract

There had been done a research about the development of three tier diagnostic test to identify student's misconception in chemical bonding supported by program called PHP-MySQL. The aims of the reaseach is knowing the feasibility of three tier diagnostic test based on construct, content, and linguistic validity, and also to know the reliability of the test and the effectiveness of three tier diagnostic test in identifying the student's misconception based on the result of limited trial. The type of this research is development research using 4-D model that consist of four steps, that are define, design, develop and disseminate, but this research is limited and ended up at develop phase. The limited trial in this research is done two times at SMAN 18 Surabaya. First limited trial was done at X MIA 5 with 20 students to get data regarding item difficulty and item discrimination and reliability of the three tier diagnostic tests. The second limited trial was done at X MIA 3 with 10 students to get data of students' misconception and effectiveness of three tier diagnostic tests in identifying misconception. The result of validity scoring based on contain, construct, and language respectively 82%, 83% dan 80,75%, thus categorized as very valid. The validity of three tier diagnostic test media obtained at indicator 1 regarding display of media is 83%, at indicator 2 regarding the ease in using media obtained 85,83%, thus categorized as very valid . The reliability of the test obtained is 0,968 which is categorized as very reliable and the effectiveness of three tier diagnostic test media obtained is 67,12% which is categorized as effective.

Key word : *Three tier diagnostic tests, chemical bonding, misconseption, feasibility*

INTRODUCTION

Chemistry is a science that studies the material that covers the structure, composition, properties and changes of matter and energy that accompany it. The materials are studied in chemistry are so thorough and very broad coverage, this is an obstacle in the learning process chemistry in school when students are required to understand the chemical material with such a broad scope of learning at school while time is limited [1]. Sometimesthe real learning process does not go smoothly and often creates confusion instead of learning on the students many things that students do not understand while learning continues in accordance with lesson plans that have been made. These things will lead to misconceptions students to the material that is being studied in other words learning in class will be ineffective. This is supported by the results of interview chemistry teacher at SMAN 18 Surabaya on March 16th 2015 regarding common misconception happened in learning process especially chemistry.

The chemical bonding is one of topics in chemistry subjects at class X SMA and it is also coverage at level macroscopic, microscopic and symbolic. So on this subject the students demanded to be able to understand the concepts being taught, but the reality on the learning activity shows that many students do not understand the concepts in chemical bonding. Student difficulties in studying chemical bonding is supported by the results of the questionnaire at preresearch on March 16, 2015 which is about 75% of students in class X-MIA 3 have difficulties in understanding the concept of chemical bonding, these things will lead to be students' misconception.

There are two main reasons for the difficulties faced by the students, firstly topic in chemistry are very abstract and secondly the words commonly used in everyday life has a different meaning in chemistry terms. Differences meaning of the students' concepts with experts can be called misconceptions. Misconceptions can happen to anyone, especially students in the learning process where students have the initial concept (preconceptions) which are brought into the formal class, and difficult to be addressed, especially when these misconceptions can help students in solving problems [2]. Students sometimes make their own interpretation to a concept they learned but the results of interpretation in the form of ideas that exist in the cognitive structure of the students regarding the attributes and criterion of the concepts sometimes are not appropriate or even contrary to the concept that has been agreed by the experts.

Identify the level of understanding and misconceptions students become major problems in research in recent years. Problems misconceptions attract researchers to try to find ways how to locate the student misconceptions and misconceptions of the problems that occurred, so many misconceptions detection instrument developed in studying the chemical form of multiple choice questions tiered [3]

The instruments to detect the presence of misconception particularly about the subject matter of the chemical boning has not been found and developed even if it is available, largely as usual multiple choice questions, open ended, or interviews. Besides difficult to obtain such instrumentsfrom developers, the curriculum, the depth and the content capacious of the material in the chemical bonding of other countries is different, so it is relatively not appropriate when

applied in Indonesia. Therefore it is very important to the development of detection instrument in Chemistry to identify the student misconception, especially in the subject matter of the chemical bonding in accordance with the prevailing curriculum is curriculum material 2013. The deepening and widening adjusted to predetermined indicators that can be used easily by teachers in detecting the presence of misconceptions on high school students in Indonesia.

Several previous studies by Victoria Diana Chong in 2013 stated that the three tier test is effective used in detecting students' misconception [4]. In addition the study done by Akhmad Subekti about diagnostic tests with the help of PHP-MySQL program developed is feasible to use [5]

METHOD

This type of research is the development of research with reference model 4-D with four steps: define, design, develop and disseminate, but in the present study is limited to the development step only.

The purpose of the development of diagnostic tests three tier is to determine the feasibility and reliability of three tier diagnostic test, in addition to knowing the effectiveness of the three-tier diagnostic test that supported by PHP-MySQL in identifying students' misconception. So in this research will be developed the three tier diagnostic test which is supported by PHP-MySQL program as the diagnostic media.

This study was conducted at SMAN 18 Surabaya in class X-MIA 5 with the number of students are 20 as second trial and MIA-3 with the number of students are 10 as first trial. The first trial conducted to obtain data on the quality of

the items such as: item difficulty, item discrimination, and reliability of diagnostic test. The second trial conducted to obtain data of the location and number of students' misconceptions. And then students' responses is used to know as the basic value of the effectiveness of the three-tier diagnostic tests to identify student misconceptions. To determine the validity of diagnostic tests three tier based on criterion of content, and language constructs, validate the contents, constructs, and language items is done by expert they are lecturers and chemistry teachers'. While the of three tier diagnostic test validated by learning media experts those are chemistry lecturer.

The method used to obtain the data are questionnaires, pre-study, and test. The questionnaire consist of a questionnaire study of material and media experts to obtain experts' suggestions for better improvements in developing three tier diagnostic test included its computer based devic, validation questionnaire for experts to obtain a quality product developed assessment, student questionnaire responses to determine the level of media effectivness.

Both qualitative and quantitative analysis is done in this research. Qualitative analysis includes validation and quantitative analysis includes item discrimination, item difficulty and item relibiality of diagnostic tests.

Item and media of three tier diagnostic test both are validated by expert by using this following formula:

$$\% \text{ Percentage} = \frac{Score\ obtained}{Criteria\ score} \times 100\%$$

The alternative scoring in this study using Likert Scale as in the Table 1:

Table 1 Likert Scale

Valuation	Score
Invalid	1
Less valid	2
Enough Valid	3
Valid	4
Very valid	5

[6]

Based on interpretation table score, feasibility achieved if the score obtained $\geq$ 61%.

Item difficulty index analysis was conducted to obtain data on the level of difficulty of items to be tested. Difficulty index calculation is done by using the following formula:

$$P = \frac{B}{Js}$$

The value of item difficulty interpered into some catagories such as difficult, medium and easy. The best item to be tested to students is an item with item difficulty of $0,30 \leq P \leq 0,7$ catagorized as medium[7].

Analysis of Discrimination items were conducted to obtain items that can distinguish students' cognitive level. item discrimination calculation is done by dividing the students into groups of upper and lower, and calculation using the following formula:

$$D = \frac{\sum BA}{JA} - \frac{\sum BB}{JB}$$

The value of item discrimination can be interpreted into categories such as Very good, good, Enough, Low, and very low. The best item to be tested to students is an item that has item discrimination 0,6-1,0 can be categorized as very good or good [7].

Analysis of the reliability of diagnostic tests done to find out about the item relibiality done by using product

moment correlation formula calculation with rough numbers as follow:

$$r_{xy} = \frac{N\Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{\{N\Sigma X^2 - (\Sigma X)^2\}N\Sigma Y^2 - (\Sigma Y)^2}}$$

Furthermore, to calculate relibialitas entirely then use the formula

$$r_{11} = \frac{2_{r1/21/2}}{1 + r_{1/21/2}}$$

The amount correlation coefficient calculation results can be interpreted with the following requirements scale.

Table 2 Coefficient Correlation Of Product Moment

Coefficient Correlation	Criterion
0,8000-1,0000	Very High
0,6000-0,8000	High
0,4000-0,6000	Enough
0,2000-0,4000	Low
0,0000- 0,2000	Very Low

[7]

Those items are Catagorized to be reliable to use items with a value of 0.6000 up to 0.8000 with the reliability category high.

Afetr All, the items selection as a diagnostic test, the selection criterion used diagnostic test items. Under the rules Nitko (1983) and Nunnally (1970) [9] about the good selection criterion are as in Table 3:

Table 3 Criterion item selection

Criterion	Coefficient	Decission
Item Difficulty	0,30 - 0,70	Accepted
	0,10 - 0,29 or 0,70 - 0,90	Revised
	<0,10 or>0,90	Rejected
Item Discrimination	>0,30	Accepted
	0,10 - 0,29	Revised
	<0,10	Rejected

[9]

item with the decision "acceptable" can be directly used, while the matter of the decision "revised" may be used on condition that must be repaired in advance and for the matter of the decision "rejected" should not be used.

Media which developed need to be evaluated formatively focusing on effectiveness and efficiency in order to achieve the goals set [8]. So it is necessary to analyze the effectiveness of the media that have been developed, the value of the effectiveness are student response percentage data obtained and student misconceptions that can be determined by using the following formula.

$$P = \frac{S}{Js} \times 100\%$$

And then Student responses were then determined using the following formula:

$$P = \frac{F}{N} x 100\%$$

Media received a positive response when the response of students who gained ≥61% [6]

In purpose to analyze the students 'level of understanding and misconceptions can be done by looking at the results of the students' answers after taking the test is given. There are several categories that classify the students' answers which written in the Table 4:

Table 4 Answer Classification

First tier	Second tier	Third tier	Kategori
Correct	Correct	Sure	Understand concept (TK)
Correct	Incorrect	Sure	Misconseption 1 (MK1)
Incorrect	Correct	Sure	Misconseption 2 (MK2)

First tier	Second tier	Third tier	Kategori
Incorrect	Incorrect	Sure	Misconseption 3 (MK3)
Correct	Correct	Doubt	Not understand concept (TTK)
Correct	Incorrect	Doubt	Not understand concept (TTK)
Incorrect	Correct	Doubt	Not understand concept (TTK)
Incorrect	Incorrect	Doubt	Not understand concept (TTK)

[10]

RESULT AND DISCUSSION

Results of the validation analysis chemistry lecturer and teacher might be explained in detail as follows:

a) Validation construct, contain, and language

There were 100 items at beginning before passing study stage by expert, and then only 94 were left or selected to pass the next are the validation of the content aspect, constructs and language while the other 6 items were rejected due to incorrigible and not correlate to the topic. Here the result of validation 94 questions

Table 5 Validation Result

Category	Validation of		
	Content	Construct	Language
Very Valid	21	22	18
Valid	73	72	75
Enough	0	0	1
Less	0	0	
Invalid	0	0	

b) Validation of three tier diagnostic test media

The media of the three tier diagnostic test was judged by two criterion developed into the 10 indicators as in the Table 6:

Table 6 Validation Result of The Three Tier Diagnostic Test Media

No	Criterion	Indcator	Score by validator			∑	%	Criterion
			1	**2**	**3**			
1	Display of the three tier diagnostic test media	a.Systematic display	4	4	4	12	80.00	Valid
		b. background	4	4	4	12	80.00	Valid
		c. Font size	5	4	4	13	86.67	Very valid
		d. Colour of text	5	4	4	13	86.67	Very valid
		e. Animation	4	4	4	12	80.00	Valid
2	The ease in operating the the three tier diagnostic test media	a. the ease of indicator input into media	5	4	3	12	80.00	Valid
		b. the ease of questions input into media	5	4	4	13	86.67	Very valid
		c. the ease of answer input into media	5	4	4	13	86.67	Very valid
		d. the media is communicative	5	4	4	13	86.67	Very valid
		e. the text structure is correct grammatically	5	4	4	13	86.67	Very valid
		f. fast in indentifying the student misconception	5	4	4	13	86.67	Very valid
		g. .fast in showing the result (%) of student misconception individualy	5	4	4	13	86.67	Very valid
		h. fast in showing the result (%) of student misconception classicaly	5	4	4	13	86.67	Very valid

Based on the table above at indicator 1 the validity score is 82,67% and at indicatoe 2 the score obtained is 85,83%, thus the media of three tier diagnostic test is categorized as very valid [6].

c) Item Difficulties

Level of difficulty of items is calculated by comparing the number of students who answered correctly and the number of students who answered incorrectly. Good Item is which has medium difficulty level, it is intended to prevent too many student answer right or vice versa [7]. Here are items classification based on the difficulty level that conducted after second trial done presented in the Table 7:

Table 7 Calculation of Item Difficulties

Criterion	∑ items	Items number
Easy	25	2,5,9,10,11,27,33,35,43,44,45,47,51,53,59,66,67,68,70,71,72,78,81,84,85,88
Medium	65	1,3,4,6,7,8,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,28,29,30,31,32,34,36,37,38,39,40,41,42,46,48,49,50,52,54,55,56,58,60,61,62,64,65,69,73,74,75,76,77,79,80,82,86,87,90,92,93,94
Difficult	4	57,63,89,91.

Based on the Table 7, showed that 25 items can be categorized easily, 65 questions were categorized and four categorized difficult matter. Those items which ideal for use in a test are categorized items easily and medium or in the range of 0.3 up to 0.7 [7].

d) Item Discrimination

By splitting 20 students into upper groups and lower group, so the groups with lower score are student 3, student 6, student 19, student 18, student 2, student 7, student 16, student 15, student 11, student 20 and upper group are student 1, student 4, student 5, student 8, student 9, student 10, student 12,student 13, student 14, student 17, it will result the value of the item discrimination of 94 selected items which presented in Table 8:

Table 8 Calculation of item discrimination

Criterion	∑ item	Number of item
Very low	4	15,48,85,89
Low	16	16,30,38,43,45,46,53,56,63,78,79,83,84,93

Criterion	∑ item	Number of item
Enough	38	2,4,5,6,8,9,10,11,12,14,18,20,25,27,31,33,34,35,37,40,47,49,51,54,55,57,59,61,62,67,68,70,76,80,81,87,88,90,91
Good	35	1,3,7,13,17,19,21,22,23,24,26,28,29,32,36,39,41,42,44,50,54,58,64,65,69,71,72,73,74,75,77,82,86,94
Very Good	1	66

Based on Table 8, from 94 questions that have been tested at first trial only four questions categorized as very low, 16 items categorized as low, 38 items were categorized fairly, 35 items were categorized as good and only 1 item categorized as very good. Thus only 36 items of diagnostic test that allowed to be tested at next trial which categorized as good and very good only [7].

e) Reliability

in this study, the reliability is calculated using moments products equation and half-split method by means of an objective test is split into two equal parts[7], in condition the number of questions that analyzed must be even (that evenly). After calculation using equation of moments products, the reliability value obtained 0.993 and overall reliability calculated by means split-half method obtained 0.968 so that the reliability of the test can be categorized as very high [7].

d) Items Decision

Selection of items to put in the media is a item whichis very valid or valid according to the expert, an item which has a item discrimination good or very good,

and an item which has midle difficulty level with a value of 0.30 up to 0.7 in accordance with the rules of Nitko [9] and also must be reliable. Thus the number of the best three tier diagnostic test item that classified to input into media is 35 questions. Those questions that have input into media will be used to identify students' misconception.

c) Effectiveness of three tier diagnostic test

The media developed considered as effective when it is able to achieve the goal set or reason of the media is developed [8]. The three tier diagnostic media in this research is aimed to detect the location of the misconceptions. students so the media is considered effective if it is able to show the location of the misconceptions students The value of the effectiveness of this diagnostic test is the percentage of tier three misconceptions experienced by the students. After the second trial activities conducted was known that the misconceptions of students is 67,12%. The other way to see the effectiveness of media that can be seen from practicality or flexibility in operating the media [8]. Here is the result of students' responds with two criterion that expanded become eight indicators as in the table 9:

Table 9 Students' responds result

No	Criterion		Indicator	Yes	No	%
1	The display of three tier diagnostic test media	a.	Systematic display is clear	10	0	100
		b.	The background is interesting	6	4	60
		c.	The size of font is clear to read	10	0	100
		d.	Colour of text is clear	10	0	100
		e.	The animations help to understanding the question	8	2	80
			$\sum$			88
2	The ease of media operation	a.	Chosing answering is easy to do	9	1	90
		b.	The media is communicative	9	1	90
		c.	The media is easy operating	10	0	100
			$\sum$			93

The values obtained from criterion 1 regarding media display is 88% thus categorized as very valid. and the value obtained from criterion 2 regarding ease of operating media is 93% thus categorized as very valid based on the interpretation of the score.

The developed media get positive response as the score of student response obtained $\geq 61\%$ or it can be concluded that the three-tier diagnostic test media said to be effective in detecting misconceptions experienced by students [6].

CLOSING

Conclution

Based on the results of research conducted it was concluded that the three

tier diagnostic tests in chemical bonding is proper for using as a tool to detect students' misconceptions with following details:

1. The three tier diagnostic test in chemical bonding topic is very feasible based on the content validity is 82%, validity construct is 83% and language validity is 80,72%. While the three-tier diagnostic test media gained 84% on the criterion 1 regarding the display media and 85.83% on criterion 2 regarding the ease of media to use. So that the media developed is very feasible as a tool to detect the student misconceptions.

2. The three-tier diagnostic test developed has reliability value of 0.968 which means the value of reliability of three tier diagnostic test is very high.

3. The three tier diagnostic test media is categorized as effective by the ability to detect student misconceptions experienced by 67.12% and based on the student's response to the ease of media by 93%.

Suggestion

1. This Research was done limited until develop phase, therefore it is necessary to conduct further research so that later obtained information about the threetier diagnostic shortage developed for further improvement.

2. it would be better if the three-tier diagnostic test chemical bonding materials tested immediately after completing student teaching in chemical bonding topic at school to get more overview about misconceptions.

3. The use of the three tier diagnostic media in the long term will cause the database bigger and will make the PC access getting slow down, to anticipate this problem items that have been tested should be deleted.

The computer access slow down when the test is an obstacle that often occurs when using PHP-MySQL program, the diagnostic media should be hosted on a website.

BIBLIOGRAPHY

1. Wu,H.K.2001.*Promoting Understanding of Chemical Representations: Students' Use of a Visualization Tool in the Classroom.Journal of Research in Science Teaching*, 38(7): 821–842.

2. Suparno, Paul. 2013. *Miskonsepsi & perubahan konsep pendidikan fisika*. Jakarta: Grasindo.

3. Treagust, D. F. , Chandrasegaran A. L. ,and Mocerino,Mauro (2007). *The development of a two-tier multipl choice diagnostik instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation*.The Royal Society of Chemistry, 8 (3), 293-307.

4. Chong, Victoria Diana. 2013. *Using An Activity Worksheet To Remediate Students' Alternative Conceptions Of Metallic Bonding:American International Journal of Contemporary Research. Vol. 3 No. 11; November 2013 39-52.*

5. Subekti, Akhmad. 2011. *Pengembangan Tes Diagnostik*

Berbasis Komputer Menggunakan Program PHP-Mysql PadaMateri Ikatan kimia SMA kelas X. Skripsi tidak dipublikasikan. Surabaya: Universitas Negeri Surabaya.

6. Riduwan. 2012. *Skala Pengukuran Variabel-Variabel Penelitian*. Bandung: Alfabeta.

7. Arikunto, Suharsimi. 2013. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.

8. Sadiman. 2010.*Media Pendidikan: Pengertian Pemngembangan, Dan Pemanfaatan.* Jakart: PT. Raja Grafindo Persada.

9. Surapranata, Sumarna. 2004. *Analisis, Validitas, Reliabilitas dan Interpretasi Hasil Tes Implementasi Kurikulum 2004*. Bandung: PT. Remaja Rosdakarya.

10. Arslan, H. Ozge, Ceyhan Cigdemoglu and Christine Moseley. 2012. *A Three-Tier Diagnostic Test to Assess Pre-Service Teachers' Misconceptions about Global Warming, Greenhouse Effect, Ozone Layer Depletion, and Acid Rain.* Internation Journal of Science Education. Vol. 34 (11): hal. 1667-1686.