11(1), 2022, 121-137 DOI: 10.23960/jppk.v11.i1.2022.12
Jurnal Pendidikan dan Pembelajaran Kimia
e-ISSN: 2714-9595 | p-ISSN: 2302-1772
http://jurnal.fkip.unila.ac.id/index.php/JPK/index

Development of Three-Tier Multiple Choice Diagnostic Test Instruments for Measuring Students' Misconceptions in Chemical Bonding

Allika Haya Fahrunisa, Ajat Sudrajat
Department of Chemical Education, Universitas Negeri Medan, Indonesia

Corresponding e-mail: allikahayafahrunisa91@gmail.com

Received: March 18th, 2022 *Accepted:* April 18th, 2022 *Online Published:* April 29th, 2022

Abstract: Development of Three-Tier Multiple Choice Diagnostc Test Instruments for Measuring Students' Misconceptions in Chemical Bonding. The aim of this study is to determine the need for the development of a three-tier multiple choice diagnostic test instrument in measuring students' misconceptions about chemical bonding material and to know the results of measuring misconceptions using a three- tier multiple choice diagnostic test instrument. This research was conducted at SMAN 1 Rantau Selatan in July 2021 - February 2022. The sample used in this study were students of class X MIA SMAN 1 Rantau Selatan who failed to achieve the KKM value on chemical bonding material. The development in this study used the ADDIE model with two trials. The conclusion of this research are showed that the three tier multiple choice diagnostic test instrument that has been developed meets the characteristics of the substance of the instrument to measure misconceptions in chemical bonding material, the validity of the test was 20 valid question and 20 invalid question. The reliability test showed a realiability lift 0.84. There are 40.3% of students who have misconceptions. For the category of not understanding the concept there are 57.3% of students and the category of understanding the concept there are 2.4% students.

Keywords: three-tier multiple choice, misconception, chemical bond.

Abstrak: Pengembangan Instrumen Tes Diagnostik Three Tier Multiple Choice Untuk Mengukur Miskonsepsi Siswa Pada Materi Ikatan Kimia. *Penelitian ini bertujuan untuk mengetahui kebutuhan pengembangan instrumen tes diagnostik pilihan ganda tiga tingkat dalam mengukur miskonsepsi siswa pada materi ikatan kimia dan untuk mengetahui hasil pengukuran miskonsepsi menggunakan instrumen tes diagnostik pilihan ganda tiga tingkat. Penelitian ini dilaksanakan di SMAN 1 Rantau Selatan pada bulan Juli 2021 – Februari 2022. Sampel yang digunakan dalam penelitian ini adalah siswa kelas X MIA SMAN 1 Rantau Selatan yang tidak mencapai nilai KKM pada materi ikatan kimia. Pengembangan dalam penelitian ini menggunakan model ADDIE dengan dua kali uji coba. Kesimpulan dari penelitian adalah instrumen tes diagnostik pilihan ganda tiga tingkat yang telah dikembangkan memenuhi karakteristik substansi instrumen untuk mengukur miskonsepsi pada materi ikatan kimia, validitas tes menunjukkan hasil bahwa terdapat 20 pertanyaan valid dan 20 pertanyaan tidak valid. . Uji reliabilitas menunjukkan hasil reliabilitas sebesar 0,84. Hasil pengukuran adalah terdapat 40,3% siswa yang mengalami miskonsepsi. Untuk kategori tidak paham konsep terdapat 57,3% siswa dan kategori paham konsep ada 2,4% siswa.*

Kata kunci: *pilihan ganda tiga tingkat, miskonsepsi, ikatan kimia.*

▪ INTRODUCTION

Chemical bonding is one part of class X chemistry which discusses the rules for an atom or element to be able to bond to form a compound called a chemical compound. Without chemical bonds, chemical elements cannot form compounds. There are so many students who consider this chemical bonding material is a difficult thing, this can be proven in research by Rahman et al. (2014) who wrote in their article "Based on the results of interviews with 27 students of class XI IPA 1, information was obtained that the interest of class XI students IPA 1 for chemistry lessons is quite high. However, their high interest actually stems from their interesting the teacher's ways and methods of teaching. Most students admitted that when they were in grade X, they did not like or lack interest in chemistry, including chemical bonding material, because the teacher's teaching method did not interest them. As a result, they pay less attention to what the teacher explains about the material presented, especially chemical bonds" (Rahman et al, 2014)

Misconception is basically an intersection between the theoretical concept of a material with the ability of the human brain to understand the concept. If human understanding is different from the theoretical concept, it can be stated that there is a misconception. Misconceptions can occur due to various factors, both internal and external factors. Internal factors include the lack of student interest in understanding the material so that when receiving an understanding of a material, the brain will automatically reject a little or a lot of the material and misconceptions occur. External factors include the teacher's lack of understanding of the material being taught and the preparation of language in convey the material can also be stated as a factor in the occurrence of misconceptions. Misconceptions are an obstacle for students to achieve learning objectives. One of the learning difficulties of students is when students experience misconceptions. Misconception is an understanding of the concepts contained in the minds of students as opposed to scientific concepts, which are influenced by the experiences of students (Hammer, 1996).

When viewed from the current state and situation of the teaching and learning process, where all students have to do online learning (on the network) due to the Covid-19 pandemic that has hit 215 countries in the world, one of which is Indonesia and for this condition, the convey of learning materials is not carried out directly or face-to-face but through online learning media, where this situation also contributes to the occurrence of student misconceptions about chemical bonding material.

There are many other efforts that can be used as an action in overcoming students' misconceptions about chemical bonding material. One of them is by developing a three tier multiple choice diagnostic test instrument to detect students' misconceptions about chemical bonding material. According to the Ministry of National Education (2007), a diagnostic test is a test used to find out the weaknesses of students so that these results can be used as a basis for providing follow-up in the form of appropriate treatment and in accordance with the weaknesses of the students. The three-tier multiple choice diagnostic test is one type of diagnostic test that can be used to identify and measure students' misconceptions. The three tier multiple choice diagnostic test is a development of the two tier multiple choice diagnostic test. This development is found in the addition of the confidence level of students in choosing the answers and reasons given (Mubarak, 2016).

Based on the results of interviews with chemistry teachers at SMAN 1 Rantau Selatan, Mrs. Saddiah Siregar as a chemistry teacher stated that many students thought

that chemistry subjects were difficult subjects to understand. This is because the chemistry subject contains many concepts that are interconnected with each other and difficult to understand. Mrs. Saddiah said that not all students could understand the concepts contained in chemistry subjects, especially in chemical bonding material. Understanding that cannot be mastered then causes misconceptions in students and this is one of the causes of not achieving maximum student learning outcomes and learning objectives are not achieved optimally. According to Mrs. Saddiah, about 30-40% of students fail to achieve the KKM value for chemical bonding material.

Based on the results of initial observations and interviews at SMAN 1 Rantau Selatan, it is known that the school has never conducted a test to analyze students' misconceptions, especially on chemical bonding material, generally in schools only use objective test instruments to measure students' cognitive abilities. Therefore, it is necessary to develop a three-tier multiple choice diagnostic test instrument to measure students' misconceptions about chemical bonding material

▪ METHOD

The type of research used in this research is development research with the ADDIE development model (Analysis, Design, Development, Implementation, Evaluation) which refers to the main processes of the learning system development process (Molenda, 2003). The research was conducted at SMAN 1 Rantau Selatan which is located at Jalan Ki Hajar Dewantara No.1, Labuhan Batu, North Sumatra in odd semesters, from July to February, academic year 2021/2022.

The population in the study was all students of class X MIA SMAN 1 Rantau Selatan who had studied chemical bonding material in the odd semester of 2020/2021 which consisted of 6 classes with a student load of 40 people in one class. Sampling in this study was taken using a purposive sampling technique, which is a sampling technique for certain reasons, is class X students who are confirmed to have studied chemical bonding material and are indicated to have misconceptions on chemical bonding material seen from the results of daily or formative exams that do not reach the KKM value based on provisions of the school curriculum.

Research Procedure

Analysis

At this step an analysis of the literature study is carried out as the basis for this research, namely the development of a three-tier multiple choice diagnostic test instrument, an analysis of basic competencies and content competencies based on learning devices according to the 2013 curriculum on chemical bonding material.

Design

At this step, product design is carried out in the form of a level one multiple choice test, a reasoned multiple choice test and a confidence test.

a. Tier one multiple choice test

The tier one multiple choice test was developed based on the results of basic competency analysis on chemical bonding material and produced indicators for each competency. These indicators are then used to compose level one multiple choice test items. Items that have been validated are then tested on research subjects.

b. Reasonable multiple choice test

The multiple choice reasoned test was developed from the data of the level one multiple choice test results. This test is in the form of a multiple choice test that contains a choice of reasons for the answer options chosen by students at the first tier. The reasons given by students are then analyzed and interpreted according to the predetermined combination of answers.

c. Confidence test

The third tier contains questions about how sure students are of the answers they have chosen. With this we can find out whether students only guess or really know the answer to the question.

Development

At this step, the product development of this research is carried out, namely the Three-Tier diagnostic test instrument. The test instrument consists of three-level multiple choice questions. The first tier consists of five answer choices obtained from the literature study. The second tier consists of five choices of reasons for students to answer questions at the first level. The third tier is the level of confidence in the form of the Certainty of Response Index (CRI) developed by Hasan et al (1999).

Implementation

At the implementation step, after the instrument has been designed, then the questions tested on class X MIA students at SMA Negeri 1 Rantau Selatan. The results of the student's answers then tested for each item to be validated through a validation test and by an expert lecturer (chemistry lecturer). This validation is carried out with the aim of knowing the suitability of the items that have been developed with the aim of detecting misconceptions that occur in chemical bonding materials. After the validation test, the questions also tested for reliability, difficulty level, distractor test and different power test.

Evaluation

The evaluation step is carried out to see whether the developed instrument is perfect or still needs revision. This step is carried out based on the results of a series of validation tests, reliability tests, distractor tests and tests of different power of questions. After the series of tests are carried out, the product is further improved based on suggestions and input from the validator team. The valid items were then tested on the tenth grade students of MIA SMAN 1 Rantau Selatan. The results of this test then calculated in the form of percent and obtain descriptive data in the form of profiles of students' misconceptions on chemical bonding material.

Data Analysis Techniques

Validation Test

Validity test is used to determine whether the test items are valid or not. Invalid questions will be discarded and not used, while valid questions mean that they can be used. The formula used to test the validity is the biserial point correlation coefficient.

$$r_{pbs} = \frac{M_p - M_t}{S_t}\sqrt{\frac{p}{q}}$$

r_{pbs} = point biserial correlation coefficient
M_p = Mean score of the subjects who answered the item correctly
M_t = Mean total score (average score of all test takers)
S_t = Standard deviation of total score

(Roseti, 2013)

In this test, the validity coefficient obtained (rxy) is compared with the r values of the moment product table with the criteria ; if rxy > rtable, then the item is said to be valid (Silitonga, 2014)

Reliability Test

Reliability is the degree to which an instrument measures what it is supposed to measure. There are three ways of implementation to test the reliability of a test, namely: (1) single test, (2) retest, and (3) alternative test (BAPM, 2008). For the calculation of the reliability coefficient can be done using the Kuder-Richardson formula (KR-20) :

$$r_{11} = \frac{n}{n-1}\left(\frac{s_t^2 - \sum_{i=1}^{n} p_i q_i}{s_t^2}\right)$$

n = number of questions
P_i = the proportion of many subjects who answered correctly on the item i
q_i = the proportion of many subjects who answered incorrectly on the item i
S^2_t = total score variance.

Difficulty Level

The difficulty level of the question is the opportunity to correctly answer a question at a level of ability or it can be said to know that a question is classified as easy or difficult. To measure the level of difficulty, the researcher used the formula (Laila & Alfath, 2019) :

$$P = \frac{Np}{N}$$

P = Proportion or item difficulty index number
N_p = Number of samples who can answer the question correctly
N = Number of samples who answered the question

Different Power

According to Purwanto (2010) distinguishing power (DB) is the ability of the THB (Learning Outcome Test) items to distinguish students who have high and low abilities. This discriminatory analysis aims to determine the ability of the questions in distinguishing students who are classified as capable (high in achievement) with students who are classified as weak in achievement (Sudjana, 2012). To analyze the discriminatory power of the questions, the researcher used the formula:

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

JA = number of test takers in the upper group
JB = number of lower group test takers
BA = the number of the upper group who answered correctly
BB= the number of the lower group who answered correctly

Distractors

A good distractor is one that can be avoided by smart children and chosen by children who are less intelligent, lest the opposite happens (Wahyuni & Ibrahim, 2012). To analyze the effectiveness of the distractor, the researcher uses the formula :

$$\text{Distractor X} = \frac{\text{JPA} + \text{JPB}}{\text{JA} + \text{JB}} \times 100\%$$

JPA = Top group voter
JPB = Lower group voter
JA = Number of students in the upper group
JB = Number of students in the lower group

(Silitonga, 2014)

Test Scoring

Correct answers and reasons are given a score of 1 and if the answers to suggestions and reasons are wrong or do not provide answers, they are not given a score.

Interpretation of Three Tier Multiple Choice Diagnostik Test

This step is the peak of the research, the results of measuring the misconceptions of class X students on chemical bonding material using a three-tier multiple choice diagnostic test instrument and at this step a profile of students' misconceptions of class X on chemical bonding material will reveal where the students' misconceptions are located and the number of misconceptions load in percent.

Table 1. Criteria for Grouping Student Conceptions

No	**Cobination**	**Explanation**			**Keterangan**
		T1	**T2**	**T3**	
1	T-T-H	True	True	High	Understanding
2	T-T-L	True	True	Low	Not Understanding
3	T-F-L	True	False	Low	
4	F-T-L	False	True	Low	
5	F-F-L	False	False	Low	

6	F-T-H	False	True	High	Misconception
7	T-F-H	True	False	High	
8	F-F-H	False	False	High	

To calculate the percentage of students who experience misconceptions based on measurements using a diagnostic test instrument, the formula is used:

$$\%U = \frac{U}{N} \times 100\%$$

$$\%NU = \frac{NU}{N} \times 100\%$$

$$\%M = \frac{M}{N} \times 100\%$$

U = group of students who understand the concept
NU = group of students who do not know the concept
M = group of students who show misconceptions
N = number of students

▪ RESULT AND DISCUSSION

Data Collection and Literature Study

After collecting related references, then an analysis was carried out so that it could be seen that the cause of student failure in achieving learning objectives on chemical bonding material stems from the occurrence of misconceptions in students and in finding out the point of student misconceptions is very minimal done by the teacher and the students themselves.

At this step, curriculum analysis, student analysis and material analysis are also carried out. The selected material is adjusted to the material needs of class X MIA contained in the 2013 curriculum, namely chemical bonding material. Chemical bonding material was chosen because in that material there are concepts that are interconnected with each other and complex for students to understand so that students often experience failure in the material.

Product Description and Design

This step is the product design step, namely the test questions on the three-tier multiple choice diagnostic test instrument aimed at measuring students' misconceptions about chemical bonding material. The questions are designed according to the material that has passed the analysis stage and is based on indicators of achievement of the competence of chemical bonding material. Researchers design questions and design items by representing each sub-material which is then distributed into 17 indicators of competency achievement in chemical bonding material. The researcher developed 40 questions which were then validated by the expert validator which was then revised according to suggestions and corrective comments from the expert validator.

Product Validation by Expert

In this validation step by the expert validator, the expert validator assesses 22 assessment criteria related to the three-tier multiple choice diagnostic test that has been designed. The final stage of the assessment of the expert validator is categorized into

four categories, namely very good with item descriptions that can be used without revision, good category with item descriptions that can be used with little revision, fairly good category with item descriptions that can be used with many revisions and not good category with information on the item questions can not be used. There are three validator experts, consisting of one chemistry lecturer at the Faculty of Mathematics and Natural Sciences Unimed namely Mrs. Siti Rahmah, M.Sc., as validator 1 and two chemistry teachers at SMAN 1 Rantau Selatan namely Mrs. Saddiah Siregar, S.Pd as validator 2, and Mr. Nanang Husen , S.Pd., as validator 3. Based on the results, there are seven questions that do not have to go through revision and there are 33 questions that must be revised before being handed over to students.

Small-Scale Trial

The purpose of conducting a small-scale trial is to see a comparison of the processing time with the number of test questions given, whether it is appropriate or not. The small-scale trial phase also aims to obtain data in the analysis of the validity, reliability, level of difficulty, different power and distractor tests per item on the developed three-tier multiple choice diagnostic test instrument.

In this small-scale trial, the sample is class X MIA students who are included in the category of students who fail to achieve the KKM value or the value of mastery learning according to the school's provisions, namely scores below 75 on formative test results for chemical bonding material. For this small-scale trial, sample of 36 students was taken.

Validity Test

Validity test can be done after obtaining data from small-scale trials. Validity is calculated using the Microsoft Excel application. The validity test was carried out on 40 items based on the results of small-scale trials that had been carried out on a predetermined sample of 36 people. Based on the results of the analysis of the validity test of the small-scale test group, there were 20 valid three-tier multiple choice diagnostic questions and 20 invalid items.

Reliability Test

Based on the results of the reliability test in Microsoft Excel, the reliability value obtained is 0.84. This reliability value can be categorized as high reliability. The three tier multiple choice diagnostic test instrument in this development research has a high reliability value, so this instrument can be trusted in its use to measure students' misconceptions about chemical bonding material.

Difficulty Test

The calculation of the level of difficulty was carried out using the Microsoft Excel application. If the index shows a number between 0.20 - 0.80, then this indicates that the items are not too easy and not too difficult. The following is a table of the results of the calculation of the level of difficulty in the three-tier multiple choice diagnostic test questions:

Table 2. Results of the Analysis of the Difficulty Level of the Three Tier Multiple Choice Diagnostic Test

Item	Category	Item Number	Total
1	Hard	1,3,5,6,7,11,12,14,20,26,37	11
2	Moderate	2,4,8,9,10,13,15,16,17,18,19,21,22,23,24,25,27,28,29,30,31,32,33,34,35,36,38,39,40	29
3	Easy	-	0
		Total	40

Based on the calculation of the difficulty level of the items, it can be seen that there are 40 questions with 11 items in the difficult category and 29 items in the medium category. In percentage form, there are 27.5% items that are categorized as difficult and there are 72.5% items that are categorized as moderate. In the results of the calculation of the difficulty level of this item, there are more questions that are categorized as moderate and this indicates that the test questions used are classified as good. According to Kuncoro (2012:62) the best test category is if the test is filled with questions that have a moderate level of difficulty and have a small difficulty distribution range.

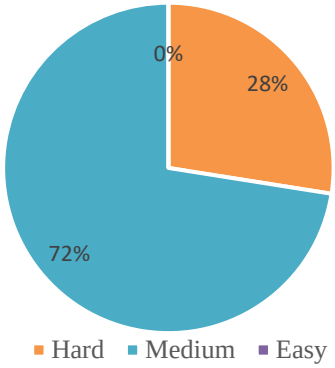

Figure 1. Percentage of Difficulty Level of Three Tier Multiple Choice Diagnostic Test

Different Power Test

The different power test of items was conducted to determine the number of differences between students with high abilities and students with low abilities. Three-tier multiple choice diagnostic test questions are categorized as very good if the items have a differentiating power value between 0.4-1, categorized as good enough if the discrepancy value is between 0.30 - 0.39, categorized as moderate if they range from 0.20-0 .29 and categorized as bad if it has a different power value 0.19. The following are the results of the calculation of the differentiating power for the three-tier multiple choice diagnostic test obtained from a small-scale trial :

Table 3. Results of Differential Power Calculation of Three Tier Multiple Choice Diagnostic Test

No	Category	Item Number	Total	Percentage
1	Very Good	10,14,28,35,38	5	12,5%
2	Good Enough	13,20,26,34,37,40	6	15%
3	Moderate	7,8,9,11,24,27,29,30,31,36	10	25%
4	Bad	1,2,3,4,5,6,12,15,16,17,18,19,21,22 ,23,25,32,33,39	19	47,5%

Based on the table, it can be seen that the three-tier multiple-choice diagnostic test questions that are categorized as bad are 19 items or 47.5%, the questions that have the power of difference in the medium category are 10 items or as much as 25%, the items that are included in the category of different power are sufficient. good as many as 6 items or by 15% and items that are included in the category of excellent discriminating power are 5 items or as many as 12.5% items.

Distractors

The item distractor test was carried out with the aim of seeing whether the distractors on the three-tier multiple choice diagnostic test items worked well or not in misleading students when working on test questions. The distractor is eligible if the student chooses the distractor more than 5% and the distractor is said to be functioning. If the distractor chosen by students is less than 5%, then the distractor is said to be not functioning. The three-tier multiple choice diagnostic test instrument consisted of 40 questions when tested on a small scale, with five answer options, one correct option and four options as distractors. Based on these circumstances, there are 160 distractor option items in the test instrument of this study. The following is a diagram of the results of the distractor function test calculation in the three-tier multiple choice diagnostic test :

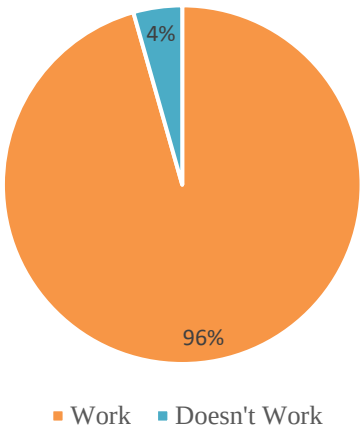

Figure 2. Percentage of Distractor Functionality Test Answer

Based on the Figure, it can be seen that there are 153 distractors or as many as 95.6% distractors that function well and there are 7 distractors or 4.4% distractors that do not function properly.

Product Analysis and Revision

After calculating the analysis of the test items obtained from small-scale trials, the next step is to revise the product based on the results of the test items and consider the results of the student response questionnaires to the three-tier multiple choice diagnostic test instrument that has been distributed. In the response questionnaire there are eight questions and there are five categories of assessment of student answers, answers that disagree are given a value of 1, answers that do not agree are given a value of 2, answers that are quite agree are given a value of 3, answers that agree are given a value of 4 and answers that strongly agree are given a value of 5. is the result of questionnaire analysis of student responses to the three-tier multiple choice diagnostic test instrument:

Table 4. Analysis of Student Response Questionnaire Results on Three Tier Multiple Choice Diagnostic Test Instruments

No	Question	Results (%)	Category
1	The material contained in the test questions has been studied previously	89%	Good
2	Sentences in the test questions can be read clearly and easily understood	83%	Good
3	The time for finishing the test is adequate for the number total of questions given.	71%	Enough
4	Finishing the test requires an understanding of the concept of chemical bonding.	87%	Good
5	The Figures, symbols and formulas in the questions are clearly legible.	80%	Good
6	The three- tier multiple-choice diagnostic test that you have taken helps you to find parts of the material that you do not understand.	82%	Good
7	You become more motivated to better understand the concepts in chemical bonding material after working on the test,	80%	Good
8	It is necessary to use a three-tier multiple choice diagnostic test on materials other than chemical bonds.	78%	Good
	Average	81.25%	Good

The table above shows the results of the questionnaire analysis of student responses to the three-tier multiple choice diagnostic test and obtained an average result of 81.25% and is classified as a good result. For the aspect of processing time on the three-tier multiple choice diagnostic test, the results were 71% and included in the fairly good category, meaning that students judged that the processing time was still

inadequate with the number of questions, for that it was necessary to add more time to the next trial after make revisions. Furthermore, for aspects of material loading, sentence clarity, image clarity, and the need for using concepts in working on test instruments, the results are included in the good category, this is because students have studied the material contained in the test and the test questions can be read and observed properly.

Before the large-scale trial, the three-tier multiple-choice diagnostic test instrument was developed through a revision stage based on the results of the item test and the results of the student response questionnaire analysis that had been distributed, namely revisions to the number of questions and time to answer questions. The number of questions used in the wide-scale trial was 20 questions out of 40 questions. The selection of items is based on the results of the validation test, and according to the results of the validation test using the Microsoft Excel application, there are 20 valid items, namely items number 2, 8, 9, 10, 11, 13, 14, 20, 22, 26, 27 , 28, 29, 30, 31, 34, 35, 36, 37, 38. The working time of the instrument was revised from 60 minutes to 90 minutes with a close book system.

Wide-Scale Trial

In the wide-scale trial, the sample used was a sample whose criteria had been determined, namely students of class X MIA who had symptoms of misconceptions with the provision that they did not reach the KKM value of the education unit set by the school, namely the value of 75. The sample is filled with 32 students who are a mixture of several classes and are united in one class to take the three-tier multiple choice diagnostic test. Some time before starting the test, the researcher gave directions and instructions to the sample about the guidelines for working on the test instrument and the steps. On the first page of the test instrument, guidelines for working on questions have also been loaded which can then be used as a reference for students while working on diagnostic test questions.

In the large-scale trial, an assessment questionnaire was also distributed to students. The purpose of filling out the student assessment questionnaire is to evaluate the assessment of the three-tier multiple choice diagnostic test according to a large-scale trial sample. The analysis of this assessment questionnaire will later become a determinant for the next product revision, if the results of the analysis have shown a good value and do not require revision, then the interpretation of the diagnostic test results can be carried out. If the results of the analysis show an unfavorable score and see several aspects whose value is inadequate, the product revision will be carried out again based on the student assessment questionnaire. The following are the results of the analysis of student assessment questionnaires on the three-tier multiple choice diagnostic test instrument.

Table 5. Analysis of Student Assessment Questionnaire Results on Three Tier Multiple Choice Diagnostic Test Instruments

No	Aspect	Value	Category
1	Material Coverage	88%	Good
2	Grammar and sentence structure	85%	Good
3	Display of Figures, symbols, reactions and chemical structures	82%	Good
4	Finishing time	84%	Good

5	Instrument development on other materials	81%	Good
6	The benefits of the instrument in measuring students' misconceptions	80%	Good
7	The three tier multiple choice diagnostic test recommendation is used for ability tests in schools for better understanding of concepts	81%	Good

Based on the table above, it is known that all aspects asked in the student assessment questionnaire on the three tier multiple choice diagnostic test instrument have reached the good assessment criteria, so revisions are no longer carried out after a wide-scale trial and interpretation of the results of the three tier multiple choice diagnostic test can be carried out after a wide-scale trial.

Interpretation of Three Tier Multiple Choice Diagnostic Test Results

After carrying out a wide-scale trial, data were obtained for the interpretation of misconceptions on the three-tier multiple choice diagnostic test instrument for chemical bonds. The interpretation of the category of misconceptions, understanding concepts and not understanding concepts is calculated based on the scores obtained by students in correctly answering the three-tier multiple choice diagnostic test questions then divided by the number of questions and multiplied by 100%. If the diagnostic test questions at the first and second tier are answered correctly, they get a score of 1 each and if they answer incorrectly they get a score of zero, on the third-tier test questions if the CRI is high, they get a score of 1 and if the CRI is low, they get a score of 0. Students are categorized as having a high confidence level or CRI if they choose options 4, 5, 6 and students are categorized as having a low CRI if the options selected are 1, 2, 3 on the third level questions. After each student's score is entered into a table in microsoft excel, it is then interpreted into the category of misconceptions, not understanding concepts and understanding concepts.

Based on the research, it is known the percentage of students' understanding categories which include misconceptions, understand concepts, and do not understand concepts. The highest percentage of misconceptions was obtained in item number 9 of 59.4% and the lowest percentage of misconceptions was obtained by item number 11 of 21.9%. The following is the percentage data for each category of overall items in the form of a graph.

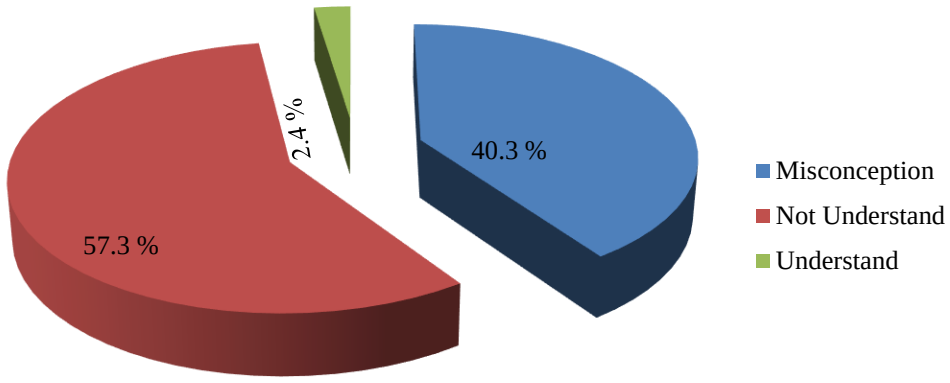

Figure 4. Percentage of Each Understanding Category

Figure 4 shows the results of the percentage of students in the form of graphs for the categories of understanding concepts, not understanding concepts, and misconceptions of all items on chemical bonding material. The graph shows that the category with the highest percentage is not understanding the concept of 57.3% and the lowest percentage is the category of understanding the concept of 2.4%. The results of the wide-scale trial were also presented for each category of misconception, concept understanding, and not understanding the concept of chemical bonding material concepts. The following table presents the percentage of each category of student understanding on the concept of chemical bonding material based on indicators of competency achievement.

Table 6. Percentage of Each Category of Students' Understanding of Chemical Bond Material Concepts Based on indicators of competency achievement.

Indicators of Competency Achievement	Item	Category Percentage (%)		
		M	NU	U
Can explain the tendency of an element to achieve stability based on its electron configuration.	1, 10, 17	43	55	2
Can explain the process of forming ionic bonds and the properties of ionic compounds	11, 12, 13, 15, 18	40.6	57.5	1.9
Can explain the process of formation covalent bonds	2, 3	29.7	65.6	4.7
Can analyze the octet rule in covalent compounds	4, 7	29.7	65.6	4.7
Can compare the polarity of a compound	5	37.5	62.5	0
Can classify the types of bonds by drawing the Lewis structure of a compound	6, 9, 16	46.9	51	2.1

Can classify the types of bonds based on chemical reactions	20	37.5	59.4	3.1
Can explain the concept of intermolecular forces	8	34.4	62.5	3.1
Can explain the concept and properties of metallic bonds	14	56.2	43.8	0
Can explain the properties and concepts of PEI and PEB of a molecule	19	50	50	0

Table 9 shows that the highest percentage of misconceptions category lies in the indicators of competency achievement explaining the concept and properties of metallic bonds which is 56.2% and the lowest percentage figure is located on the indicators of competency achievement explaining the process of forming covalent bonds and analyzing the octet rule in covalent compounds which is 29, 7%. For the category of not understanding the concept, the highest percentage figure lies in the indicators of competency achievement explaining the process of forming covalent bonds and analyzing the octet rule in covalent compounds, which is 65.6% and the lowest percentage figure in the category of not understanding the concept lies in the indicators of competency achievement explaining the concept and properties of metallic bonds, namely by 43.8%. The highest percentage for the concept understanding category lies in the indicators of competency achievement explaining the process of forming covalent bonds and analyzing the octet rule for covalent compounds, which is 4.7%. The lowest percentage figure for the concept understanding category lies in the indicators of competency achievement comparing the polarity of a compound and explaining the concepts and properties of metallic bonds and explaining the properties and concepts of PEI and PEB of a molecule, which is 0%.

▪ CONCLUSION

Based on the research that has been done regarding the development of a three-tier multiple choice diagnostic test instrument on chemical bonding materials, it can be concluded that the three tier multiple choice diagnostic test instrument has the characteristics of the substance of the instrument to measure misconceptions on chemical bonding material, the number of questions after the test is 20 questions in the form of three tier multiple choice. The three tier multiple choice diagnostic test instrument on the chemical bonding material developed has reached a good category and is suitable for use based on the assessment of 3 expert validators in the field of chemistry. Based on the validity test, there are 20 valid questions and 20 invalid questions. The reliability test reached 0.84 and was classified as very reliable. So the three tier multiple choice diagnostic test on chemical bonding material can be trusted to measure the level of understanding of the concept of chemical bonding material. Based on the results of the trials that have been carried out, there are 40.3% of students who have misconceptions. For the category of not understanding the concept there are 57.3% of students and the category of understanding the concept there are 2.4% of students.

▪ REFERENCES

Anintia, R., Sadhu, S., & Annisa, D. (2017, December). Identify students' concept understanding using three-tier multiple choice questions (TTMCs) on stoichiometry. In *International Journal of Science and Applied Science:*

Conference Series (Vol. 2, No. 1, pp. 308-317).

Bajpai, S., & Bajpai, R. (2014). Goodness of measurement: Reliability and validity. *International Journal of Medical Science and Public Health*, *3*(2), 112115.

Banerjee, A. C. (1991). Misconceptions of students and teachers in chemical equilibrium. *International journal of science education*, *13*(4), 487-494.

Bayuni, T. C., Sopandi, W., & Sujana, A. (2018, May). Identification misconception of primary school teacher education students in changes of matters using a five-tier diagnostic test. In *Journal of Physics: conference series* (Vol. 1013, No. 1, p. 012086). IOP Publishing.

Burud, I., Nagandla, K., & Agarwal, P. (2019). Impact of distractors in item analysis of multiple choice questions. *International Journal of Research in Medical Sciences*, *7*(4), 1136-1139.

Departemen Pendidikan Nasional. (2007). *Tes Diagnostik*. Jakarta: Direktorat Pembinaan Sekolah Menengah Pertama.

Hammer,D. (1996). Miscoception or P-Prims: How May Alternative Perpective of Cognitive Structure Influence Intructional Perceptions and Intensions?. *The Journal of The Learning Science*, (5)2:97-127.

Hasan, S., Diola B., & Ella L. K. (1999). Misconceptions and the Certainty of Response Index (CRI). *Physics Education*, 34(5), 294 – 299.

Jacob, J. (2017). Reliability: How? When? What?. *International Journal of Advances in Nursing Management*, *5*(4), 372-374.

Kiray, S. A., & Simsek, S. (2021). Determination and evaluation of the science teacher candidates' misconceptions about density by using four-tier diagnostic test. *International Journal of Science and Mathematics Education*, *19*(5), 935-955.

Koto, I., & Gusma, S. E. (2021). Using certainty response index to differentiate lack of knowledge and misconception about basic electrical concepts. In *Journal of Physics: Conference Series* (Vol. 1731, No. 1, p. 012070). IOP Publishing.

Maison, M., Safitri, I. C., & Wardana, R. W. (2019). Identification of Misconception of High School Students on Temperature and Calor Topic Using Four-Tier Diagnostic Instrument. *Edusains*, 11(2), 195-202.

Masykuri, M., & Rahardjo, S. B. (2018, April). Student Certainty Answering Misconception Question: Study of Three-Tier Multiple-Choice Diagnostic Test in Acid-Base and Solubility Equilibrium. In Journal of Physics: *Conference Series* (Vol. 1006, No. 1, p. 012018). IOP Publishing.

Mubarak,S., Susilaningsih, E & Cahyono,E. (2016). Pengembangan Tes Diagnostik Three-Tier Multiple Choice Untuk Mengidentifikasi Miskonsepsi Peseta Didik Kelas XI. *Journal of Innovative Science Education*,5(2):101-110.

Molenda, M. (2003). In search of the ellusive ADDIE model. Pervormance improvement, 42 (5), 34-36. Submitted for publication in A. Kovalchick & K. Dawson, Ed's, Educational Technology: An Encyclopedia.

Pacala, F. A. A. (2018). Development and Validation of Three-Tier Multiple Choice Test for Conceptual Understanding in Momentum and Collision. Journal of *Multidisciplinary Approach and Studies*, 5, 2.

Prodjosantoso, A. K., & Hertina, A. M. (2019). The Misconception Diagnosis on Ionic and Covalent Bonds Concepts with Three Tier Diagnostic Test. *International Journal of Instruction*, 12(1), 1477-1488.

Purwanto,. (2010). *Evaluasi Hasil belajar*. Yogyakarta: Pustaka Pelajar.

Putz, M. V. (2010). The bondons: The quantum particles of the chemical bond. International Journal of Molecular Sciences, 11(11), 4227-4256.

Rahman, A., Enawati, E., & Erlina. (2014). *Miskonsepsi Siswa Kelas XI IPA SMA Negeri 9 Pontianak Pada Materi Ikatan Kimia*. Program Studi Pendidikan Kimia : FKIP UNTAN.

Rintayati, P., Lukitasari, H., & Syawaludin, A. (2021). Development of Two-Tier Multiple Choice Test to Assess Indonesian Elementary Students' Higher-Order Thinking Skills. *International Journal of Instruction*, *14*(1), 555-566.

Sadhu, S., Tima, M. T., Cahyani, V. P., Laka, A. F., Annisa, D., & Fahriyah, A. R. (2017, August). Analysis of acid-base misconceptions using modified certainty of response index (CRI) and diagnostic interview for different student levels cognitive. In *International Journal of Science and Applied Science: Conference Series* (Vol. 1, No. 2, pp. 91-100).

Silitonga,P.M., (2014). Statistik : Teori dan Aplikasi dalam Penelitian. Medan: FMIPA Unimed.

Taherdoost, H. (2016). Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research. *How to test the validation of a questionnaire/survey in a research*.

Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International journal of science education*, 10(2), 159-169.

Türkoguz, S. (2020). Investigation of Three-Tier Diagnostic and Multiple Choice Tests on Chemistry Concepts with Response Change Behaviour. *International Education Studies*, 13(9), 10-22.