



Developments a three-tier multiple-choice diagnostic instrument to identify students' misconceptions about changes in the shape of object

Rifa Setiani Z^{*}, Akhmad Nugraha, Agnestasia Ramadhani Putri

Program Studi Pendidikan Guru Sekolah Dasar, Universitas Pendidikan Indonesia

**Corresponding author: rifasetianiz@gmail.com^{*}, akhmadnugraha@upi.edu, agnestasiarp@upi.edu*

Submitted/Received 02 April 2023; First Revised 28 April 2023; Accepted 24 May 2023

First Available Online 1 June 2023; Publication Date 1 June 2023

Abstract

The daily test scores obtained by 5th-grade elementary school students in the subject of changes in the states of matter are below the minimum passing grade criteria. Therefore, it is necessary to identify what makes these results so bad. Therefore, a diagnostic test instrument needs to be developed in an effort to detect the causes of low test scores below the minimum completeness criteria obtained by students in material about changes in the shape of objects. The development carried out was by developing a three-tier multiple-choice diagnostic test instrument. This three-tier multiple-choice instrument is very important to develop because it can detect students' understanding of concepts or misunderstandings that occur. This instrument consists of three tiers, including the first tier, which is related to answering questions. The second tier relates to the reasons for choosing answers at the first and third tier regarding the confidence index selected based on the answers and reasons on the first and second tiers. In this regard, the purpose of this study was to produce a three-level multiple-choice diagnostic test product that can be used to identify students' misconceptions about changes in the shape of objects in fifth grade elementary school students. The development method used is design-based research. The results of this development trial were able to identify student misconceptions, with an average percentage of students scientific knowledge in the shape of objects of 12.4%, lack of knowledge of 3.3%, misconceptions of 41%, misconceptions (false positive) of 25.4%, misconceptions (false negative) of 16%, and lucky guess of 1.9%. In addition, in relation to student responses in the development of this test tool, an average of 80.7% is in the good category. Therefore, it is clear that students can read diagnostic tests well.

Keywords: Change in object shape; Misconception; Three-tier multiple-choice.

INTRODUCTION

Learning is closely related to the occurrence of a teaching and learning process that results in the interaction of students with teachers and learning resources in a learning environment (Faizah, 2017). Students must understand a concept that is taught when teaching and learning activities take place because one indicator of success in learning is students' understanding of the concept. Concepts that students understand will affect their mastery of subsequent concepts. This is in accordance with the opinion of Suryani et al., (2016) that students' understanding of concepts is a form of learning outcome obtained by

students through teaching and learning activities. By understanding the concept, students will be able to solve problems. The aim of learning IPA is not only oriented to the concept of mastery, but students also focus on the concept of mastery of science and literature, where they are able to implement the concepts they have acquired in solving problems of everyday life. (Muyassaroh et al., 2022).

The term concept comes from the Latin word "conceptus" which means "to catch". In the Big Indonesian Dictionary, the concept is defined as a mental picture of an object, process, or anything outside of language that is understood through reason. Furthermore,

regarding understanding, understanding in the Big Indonesian Dictionary means understanding correctly. The purpose of understanding correctly is that someone knows correctly about a problem and can give certain examples about the problem; this is called understanding. Thus, if these two words are combined, namely understanding and concept, it can be interpreted as an ability possessed by a person when receiving and interpreting a mental picture of general knowledge about the same phenomenon (Dewi, 2019).

Students who understand the concept are not just memorizing it but are able to understand and interpret the concept that is being studied. According to Suryani et al (2016), students can be said to understand concepts by showing that they are able to understand a concept from various aspects and can provide more detailed explanations using their own language. Indicators of understanding concepts, according to the National Standards Agency (2006), are as follows: (a) repeating concepts in their own language; (b) classifying objects according to their nature; (c) giving examples and non-examples of concepts; (D) developing necessary and sufficient conditions for a concept; (e) applying concepts or being able to solve problems in everyday life.

Regarding the understanding of a concept, it is currently a very dangerous phenomenon. Understanding the concept is said to be a dangerous phenomenon, meaning that students often make mistakes in understanding the concept, commonly called misconceptions. Misconceptions can be considered someone's mistake in understanding a concept that is not in accordance with the thinking of scientific experts. (Wahyudi et al., 2021). This conception is built based on sound thinking or intuitively in an effort to interpret one's own experience in everyday life. Forms of misconceptions include conceptual errors, wrong relationships between concepts, and wrong views (Suparno, in Triumiana, 2017). Several ways that can be used to identify students' misconceptions are by using concept

maps, essay writing tests, diagnostic interviews, class discussions, practical questions and answers, and diagnostic tests. However, of all these methods, the one that is easier to use is a diagnostic test. Diagnostic tests can be interpreted as tests that can be used to identify the strengths and weaknesses of each student, determine student performance disabilities, and find out misconceptions that occur in students.

The Ministry of National Education (2007) describes the characteristics of a diagnostic test, namely: (1) it was designed with the aim of identifying student learning weaknesses; (2) it was developed based on an analysis of the source of errors that caused the problems that occurred to students; (3) it used questions in the form of descriptions or short answers, or using multiple choice forms accompanied by students' reasons for choosing a particular answer; and (4) it was accompanied by a follow-up plan in accordance with the difficulties identified. Thus, a diagnostic test can be said to be good if it can describe students' mastery of concepts, distinguish between concepts that students have understood and concepts that

still need to be learned, and provide information based on student answers so that weaknesses and inconsistent student thinking are identified.

Diagnostic test forms have been developed with various tiers, including: two- tier, three-tier, and four-tier multiple-choice diagnostic tests. The various types of diagnostic tests have their respective advantages and disadvantages. At this time, the choice for developing a diagnostic test is a three-tier multiple choice because it is easier to use when compared to other diagnostic tools (Kustiarini et al., 2019). The three-tier multiple choice diagnostic test is defined as a test that is in the form of multiple choices and consists of three tiers, including the first level regarding the choice of answers; the second tier is the reason for an answer contained in the first tier; and the third tier is related to the confidence index for the choice at the first tier and second tier, with closed choices namely "sure" and "not sure". This

three-tier multiple-choice diagnostic test is the development of a two-tier test, which is added with a confidence index. The confidence index for the choice of answers and reasons can measure differences in the level of student knowledge so that it can help identify misconceptions that occur in students. The advantages of the three-tier multiple choice diagnostic test include: (1) with the level of confidence in the answers and reasons chosen, it can dig deeper into the strength of students' understanding; (2) it can diagnose misconceptions that occur in students more deeply; and (3) it can determine parts of the material that require more learning (Nazura, 2021). This is in line with Wahyudi's opinion (2021) that a three-tier multiple-choice diagnostic test will be more valid for determining students' conceptions and misconceptions compared to a one-tier test or a two-tier test.

Based on the results of the initial interviews that have been conducted in one of the schools in Tasikmalaya, it shows that the material regarding changes in the shape of objects and the value of the daily test results for the last one year obtained by dominant students is below the minimum completeness

criteria set at school, which is 75. In addition Based on the results of interviews also conducted with four grade IV elementary school teachers from two different schools, it is known that, in general, teachers have never used the three-tier multiple-choice diagnostic test to measure students' understanding of concepts. In teaching and learning activities, the teacher's tendency to obtain information related to students' understanding of a material or concept is to use a measuring tool in the form of multiple-choice tests or essays. This, of course, makes it difficult for teachers to find out more information about concepts that students do not understand or misunderstand, which can lead to low student learning outcomes at the elementary school. Given the importance of understanding students' concepts in learning, it is therefore necessary to develop

diagnostic tests in schools, especially in science subjects, in the matter of changes in the shape of objects.

This three-level diagnostic test is still rarely used for the SD level. However, this test based on the results of a literature review is often used at the junior high/high school level with several subjects in science learning, including simple plane material (Nazura, 2021), elasticity (Wahyudi et al., 2021), and reaction rate (Natalia, 2023) For the elementary school level, based on the results of a literature review including the subject matter of style and motion (Nasution, 2021) as well as efforts to preserve and balance nature (Permadani, et al., 2022). The current research has differences and similarities with several previous studies. The similarities are that both use a three-tier multiple-choice diagnostic test instrument to uncover misconceptions about a concept. But the difference lies in the concept of the material. In this study, the concept of material was the main focus, namely regarding material changes in the form of objects in fifth-grade elementary school students. Besides that, the difference also lies in the method used, where some of the results of the literature review use the 4-D and ADDIE methods. Meanwhile, this research uses a design-based research method, which is often referred to as BDR. In this regard, the purpose of this research is to produce a three-tier multiple-choice diagnostic test product that can be used to identify students' misconceptions about material changes in the shape of objects in grade V at the elementary school level.

RESEARCH METHODS

The research method used in this research is design-based research (DBR). The DBR method is a method applied in the design, development, and improvement of an educational product according to basic theory and previous research (Plomp, 2013). This research method is used together with the development model adapted by Reeves, quoted

in Roghdah et al., (2021), and the stages of the development model, which consist of four steps as follows: (1) analysis and identification of problems; (2) designing solutions; (3) repeated cycles of testing and improving designs; and (4) reflection on products and strengthening solution implementation.

The first step is the analysis and identification of the problem. This step aims to obtain information related to learning needs by using an analysis of the objectives and limitations of the material used in the research. The second step of solution design aims to design a three-tier test evaluation to determine the level of students' conceptual thinking abilities and their misconceptions. The third step is a repeated cycle. This step aims to conduct repeated trials of products that have been developed and to find out student responses to the three-tier multiple-choice test instrument used. The fourth step is reflection on the product and strengthening the implementation of the solution, which aims to provide input based on what has been tested before so that the final product is more perfect.

In this study, the product developed was an evaluation instrument, namely a three-tier test to determine the level of students' understanding and misconceptions. The developed instrument consists of twenty items with tiered questions in each item and a validation sheet as an instrument scoring sheet. Researchers analyzed the data based on validity, reliability, and identification of the

level of understanding of concepts and students' misconceptions about material changes in the form of objects. The first trial stage in the development step is carried out by testing the content validity of the test instrument. The expert validation test was carried out by three science lecturers and one

science subject teacher. The first testing phase was carried out at SDN Kersanegara with 45 students as respondents.

The second trial was conducted at SDN 2 Picungremuk to find out about the ability to think about the concept of changing the shape of objects and their misconceptions. The interpretation of test results is carried out based on the analysis of test scores, where the results are interpreted into four categories: scientific knowledge, misconception (false positive), misconception (false negative), lucky guess and lack of knowledge. The test categories can be seen in table 1.

Table 1
Three-Tier Multiple-Choice Test Category

First Tier	Second Tier	Third Tier	Information
Correct	Correct	Certain	Scientific knowledge
Correct	Incorrect	Certain	Misconception (false positive)
Incorrect	Correct	Certain	Misconception (false negative)
Incorrect	Incorrect	Uncertain	Misconception
Correct	Correct	Uncertain	Lucky guess
Correct	Incorrect	Uncertain	Lack of knowledge
Incorrect	Correct	Uncertain	Lack of knowledge
Incorrect	Incorrect	Uncertain	Lack of knowledge

Source: (Source: Arslan et al., 2012).

Next, analyze how much the percentage of students who understand the concept, do not understand the concept, guess, and what happens are false positive and false negative misconceptions, using the following formula:

$$P = \frac{f}{N} \times 100$$

(Source: Sudijono, in
 Alawiyah, et al., 2017).

Information:

P = Percentage of the number of students with scientific knowledge, misconception (false positive), misconception (false negative), misconception, lucky guesses, and a lack of knowledge.

f = The number of students with scientific knowledge, misconception (false positive), misconception (false negative),

misconceptions, lucky guesses, and a lack of knowledge.

N = The total number of students taking the test

For the response of students and teachers to the result of the filling of the lift that has been done based on the analysis of the cuts, the results are interpreted into five criteria, namely: excellent, good, sufficient, not good, and bad. The range of teacher and student response criteria can be seen in Table 2.

Table 2
Percentage Range Student and Teacher
Response Criteria

Intervals	Criteria
85%-100%	Excellent
69%-84%	Good
53%-68%	Sufficient
37%-52%	Not Good
20%-36%	Bad

RESULTS AND DISCUSSION

Development of a Three Level Multiple Choice Test Instrument

This study uses the Design-Based Development Method (DBR), which consists of several stages including: (1) problem analysis and identification; (2) designing solutions; (3) repeated cycles of design testing

and improvement; and (4) reflection on the product and strengthening solution implementation.

The first step is to identify and analyze the problem. In identifying and analyzing problems, this research reviews the literature by collecting several sources in the form of journals about diagnostic tests that can be used to measure the level of understanding of concepts and misconceptions that occur in students. Meanwhile, observations and interviews were carried out in this research to obtain information and facts about the subject of the natural sciences, especially regarding changes in the shape of objects. According to the results of observations and interviews with teachers at SDN 2 Picungremuk and SDN Kersanegara, it is known that: (1) the predominance of students' daily test scores regarding changes in the shape of objects is below the minimum perfection criteria; (2) the teacher only uses essay questions or ordinary multiple choice questions as a tool to find out the level of understanding of concepts and misconceptions that occur in students; and (3) the teacher never uses a three-level multiple choice diagnostic test to identify student misconceptions.

The second step is designing a solution. An understanding of students' concepts is a sign of a successful learning process. However, for now, if we look at the results of students' daily test scores, they are below the minimum passing grade criteria. This is because there is an erroneous understanding of student concepts, commonly called misconceptions. To identify misconceptions, the appropriate thing to do is to use a diagnostic test. One valid diagnostic test to use is a three tier multiple-choice test. This test is used as a solution because, not only can you find out students' misconceptions, but you

can also find out the level of students' abilities and their weaknesses as a reference in providing appropriate treatment. Through this three-tier multiple-choice diagnostic test, it will be known how the level of students'

understanding of concepts in material changes in the form of objects. Students do not understand the concepts and misconceptions that occur.

In this step, a three-tier multiple-choice diagnostic test grid is also made. Making a grid includes two components, namely identity and matrix. The identity component includes educational units, classes, subjects, materials, number of questions, and learning objectives. The matrix component includes a description of the material, indicators, and formulation of the items. The grid used as a reference for the preparation of diagnostic test items is adjusted to the indicators and cognitive aspects used. The indicators in the grid are based on the reference in the Implementation Design of Elementary School Science Learning Curriculum 2013 Revision.

Table 3

Test of Changes in Existence

No	Material Changes in The Form of Object	Item Number
1	Freezes (liquid to solid)	1,7,8
2	Melts (solid to liquid)	2,14
3	Evaporates (liquid to gas)	3,6,9,10,17
4	Condenses (gas to liquid)	4, 15
5	Sublimes (solid to gas)	5, 11, 12, 16, 18
6	Crystallizes (gas to solid)	13, 19

The next step is to create a question sheet. Items include letter titles, student ID, work instructions, and question descriptions. Instructions for working on the questions include procedures for filling out the answer sheets and working on the questions. Each item used has three stages. The first stage is in the form of multiple-choice questions with three distractors and one correct answer. The second stage is the choice of reasons from the

responses in distractors and one correct answer. The third stage is the choice of the level of confidence in the response in the first and second stages. The number of questions in the initial product was 20.

The third step is an iterative cycle of design testing and improvement. At this step, two product trials were carried out. Product testing is carried out as an instrument improvement stage. The results of this stage are used as a reference to improve the instrument developed as an initial product. The initial product developed was assessed by the lecturer as an expert validator. At this stage of the assessment by the validator, the researcher obtained advice from experts regarding the feasibility of the instrument used. The validation report submitted to the validator contains three aspects of the assessment: content, test instructions, and language and spelling. Then the item is assessed constructively by experts, namely by giving a score of 1–5 for each indicator with the following conditions and criteria: (a) Score 5: all four indicators are met; (b) Score 4: only three indicators are met; (c) Score 3: only two indicators are met; (d) Score 2: only one indicator is met; and (e) Score 1: no indicators are met. The results of the assessment of all these indicators are then concluded and used as a reference to determine whether the item is valid or not. The following are the evaluation results obtained:

Table 4
Validity of Expert Decisions

Validator	Total Score	Information
Drs. H. Akhmad Nugraha, M.Si	36	Fit for use in the field with revisions
Agnestasia Ramadhani Putri, S.Pd., M.Pd	24	Fit for use in the field with revisions
Ghullam Hamdu, M.Pd	21	Fit for use in the field with revisions

Apart from going through a validation questionnaire, the validator also provides suggestions directly as input for revising the

initial product. The suggestions given are: (1) It is recommended that the images be clarified again; (2) It is recommended to improve the language and sentences used according to the EYD in several questions; (3) It is recommended to fix the distractor in the first level choice in one of the questions.

The first trial was conducted at Kersanegara elementary school with a total of 45 students as respondents. Implementation at this stage aims to determine the validity of the items, their reliability, their level of difficulty, and the differentiating power of the questions. The results of this step were then recapitulated and calculated using Microsoft Excel and SPSS 26 software. The data obtained in this trial was analysed to determine the validity of the items. Determining the validity of the objective items in this instrument uses the biserial point correlation formula. The r table value uses a significance level of 5%, so the r table used as a reference in interpretation is 0.2429.

Table 5
Summary of Validity Analysis Results

Category	Total score	Item number
Valid	13	1, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 16, 17, 19
Invalid	7	2, 3, 4, 11, 16, 18, 20

In calculating the reliability of the questions, the researcher used the Cronbach's alpha formula. Based on the results of the students' answers, the calculated r count is 0.2429. According to these results, it can be seen that the diagnostic test questions developed are reliable. An analysis of the difficulty level of the items is done so that the criteria for easy or difficult work on the items can be identified. The results of the item difficulty level can be seen in Table 6.

Table 6
Difficulty level recapitulation

Level of Difficulty	Total Score	Item Number
Difficult	8	2, 3, 6, 9, 12, 15, 17, 18
Medium	7	1, 4, 5, 7, 8, 10, 11, 13, 14, 16, 19, 20

Giving a response questionnaire to the teacher was carried out so that the teacher's opinion regarding the diagnostic test that was developed was known. Following are the results of the teacher's response questionnaire in Table 7.

Table 7**Teacher Response Questionnaire Results**

No	Assessment Aspects	Percentage	Category
1	Questions on a three-tier multiple-choice diagnostic instrument that are developed according to learning objectives	80%	Good
2	The material for the three-tier multiple-choice instrument is in accordance with the material for changes in the shape of objects.	90%	Excellent
3	Items on a three-tier multiple-choice diagnostic instrument were developed to identify students' Misconceptions about material changes in the shape of objects	80%	Good

4	The choice of reasons corresponds to the answer choices.	90%	Excellent
5	The questions do not provide clues to the correct answer.	90%	Excellent
6	The developed three-tier multiple-choice diagnostic instrument can be used at the elementary level.	70%	Good
7	The format of the instructions for using the three-tier multiple-choice diagnostic instrument is easy to understand.	90%	Excellent
8	Completeness of three-tier multiple choice diagnostic instrument questions according to the grid and answer keys	80%	Good

9	Using good and correct Indonesian	70%	Good
10	Sing language according to the level of thinking of elementary students	60%	Sufficient

Of the 10 aspects of the assessment that are in the very good category, there are 4, namely aspects numbers 2, 4, 5, and 7. The aspects of the assessment that are in the good category are in numbers 1, 3, 6, 8 and 9. Meanwhile, there is only one aspect of the assessment in the sufficient category. aspect at number 10. Next is the student response questionnaire. Giving response questionnaires to students was carried out so that students' assessments for the test instruments that were developed could be known. The parts presented in the questionnaire are the title, identity, instructions for filling out the questionnaire, and descriptions of the 10 statements.

Table 8
Student Response Questionnaire Results

No	Assessment Aspects	Percentage	Category
1	Conformity with the material that has been studied	86%	Excellent
2	The readability of the test questions	82%	Good
3	Ease of the sentence to understand	79%	Good

4	Students' understanding of the purpose of the questions	78%	Good
5	Image clarity	80%	Good
6	Appropriateness of the Number of questions give	79%	Good
7	Appropriateness of the time given	75%	Good
8	The need to understand the concepts to Answer test questions	85%	Excellent
9	The functioning of the test to find parts of the material that have not been understood	79%	Good
10	Can provide better motivation to understand the concept	84%	Good
Average		80,7%	Good

The final product of a three-tier multiple-choice diagnostic test that has 13 questions was tested on 53 grade 5 students at Picungremuk 2 Elementary School as a second trial with the following details: 2 questions about frozen submatter, 1 item about melted submatter, 4 items about evaporated submatter, 1 item about condensed submatter, 2 submatter problems about sublimed submatter, and 2 items about crystallised submatter. In this experiment, it was found that the average percentage of people who had scientific knowledge about the nature of matter changing form was 12.4%,

lack of knowledge was 3.3%, misconception (Ms) was 41%, misconception (false positives) was 25.4%, misconception (false negative) was 16% and lucky guess was 1.9%.

The last step in this development is product reflection and the strengthening of solution implementation. Assessment and suggestions from the validator on the diagnostic test instrument are used as a reference in improving the initial product developed. Based on the results of the validation questionnaire and some of the suggestions obtained, the researcher made several revisions, as follows: (1) image deletion and clarification are performed. In this case, there are 2 images that must be deleted in 2 questions, namely item numbers 5 and 6. Then clarify the picture on five questions covering question numbers 4, 7, 8, 9, and 10; (2) improved the question sentence in item number 2 to make it shorter; and (3) change the distractor at the first level, namely in question number 2.

The revisions made to the initial test instrument product caused several changes to the test item grid and test questions. Based on the results of the analysis of the validity of the items in Table 5, the items used as the final product are 13 questions. This is because the 20 questions tested resulted in 7 invalid questions and 13 valid questions. Based on the reliability analysis, it is known that the developed instrument obtained an r value of 0.685. This value is classified as $0.60 < r \leq 1.00$ which indicates that the test instrument developed has good reliability. An analysis of the degree of difficulty or level of difficulty is carried out in order to find out whether the items used in the test are good or not. In the results of the analysis of the level of difficulty in Table 6, it was found that 60% of the questions were in the moderate category. Good test items are items

that are not too difficult or too easy to work on (Sudijono, 2015).

Very good response results were obtained by giving a response questionnaire to teachers as an assessment of the three-tier multiple-choice diagnostic test instrument that was developed. The student assessment questionnaire was carried out in order to determine student assessments regarding the diagnostic test questions that were developed. The results of the student assessment questionnaire obtained an average of 80.7% in the good category. So that it can be seen that the diagnostic tests developed are well read by students.

Discussion of Three-Tier Multiple-Choice Diagnostic Test

After conducting field trials, it can be seen that the three-tier multiple-choice test has aspects that can differentiate it from other tests. Three-tier multiple choice (TMC) is a diagnostic test with a multiple-choice question model consisting of three levels. Aspects that differentiate TMC from other multiple-choice diagnostic test models are as follows:

- **Three tiers**

TMC questions consist of three tiers: the first tier, the second tier, and the third tier. The first tier deals with answers to questions usually consisting of factual or definitional questions; the second tier concerns the reasons for what was chosen in the first tier; the second tier consists of applied or analytical questions; and the third tier is slightly different from the two types of questions in the previous tier. The third tier relates to the student's confidence index in choosing answers and reasons that are considered correct in the previous tier. With this level of confidence, it is possible to find out which students answered the questions by guessing and even distinguishing the

wrong student answers due to misconceptions or a lack of knowledge (Andriani, 2019; Basyiroh, 2022).

- Hierarchical Structure

TMC is designed with a hierarchical structure, where the answers to one level of questions must be known before answering questions at a higher level. That is, the answers from the first tier are the basis for answering questions in the second tier, and the answers from the first and second tiers are the basis for selecting the confidence index for the third tier.

- Improve understanding and skills

By using TMC, students will be able to deepen their understanding and improve their skills in answering questions at these three levels. In addition, TMC also encourages students to think critically and creatively because they have to integrate information from lower levels to answer questions at higher levels.

In addition to these aspects, in development, there are advantages and disadvantages to three-tier multiple choice. The development of a three-tier multiple-choice (TMC) diagnostic test has several advantages, including:

- Can measure students' understanding and skills at various levels

TMC enables the measurement of students' understanding and skills at three different levels, namely basic knowledge, understanding of concepts, and the ability to apply concepts in new situations. In this way, the TMC test can provide more comprehensive information about students' abilities than the traditional multiple-choice test model, which only measures factual knowledge. Furthermore, the teacher will also gain deeper insight into the misconceptions experienced by

students by using the three-tier multiple-choice test (Savira et al., 2019).

- Facilitate the identification of student learning difficulties and provide more detailed feedback

The TMC test can assist teachers and educators in identifying student learning difficulties at each level of questions. In addition to having three levels of questions, the TMC test can provide more detailed feedback about the strengths and weaknesses of students at each level. This can assist teachers in planning more appropriate and effective teaching and learning.

- Maintain consistency and validity

The TMC test can help maintain consistency and validity in assessments. Each level of questions must be made according to the applicable curriculum and be measurable in evaluating student understanding. Thus, TMC test results are reliable and useful in making educational decisions.

- Provide more detailed teaching guidance

TMC test results can provide more detailed information about student understanding at each level of questioning. In this way teachers can identify areas that need more emphasis and select more appropriate teaching methods for each level of inquiry.

- Effective time management

Students can complete the TMC test in a relatively short amount of time. This provides efficient use of time and allows teachers and educators to focus time and energy on more in-depth and productive learning.

- Can be used in a variety of subjects:

The TMC test can be adapted for a variety of subjects. This makes it possible to develop TMC tests for different subjects and obtain useful information about student understanding in each subject.

In addition, there are various reasons why developing a three-tier multiple choice test instrument is advantageous, apart from objectivity in assessment, broad content domain sampling, mechanical assessment, and generalisation. The three-tier multiple-choice test has the advantage of allowing researchers to check the validity of the instrument and estimate misconception scores. The correlation between the two-level score, the level of confidence, and the percentage of misconceptions (false-negative) and misconceptions (false-positive) provides more accurate proof of the validity of the test (Laksono, 2020). and even with this three-tier system, it is possible to determine sub-materials that require special emphasis and design learning that is more suitable to be applied so that the misconceptions experienced by students are reduced. (Lubab, 2023).

Although three-tier multiple choice (TMC) is a very useful and effective assessment method for measuring students' understanding and skills in various subjects, there are still some deficiencies in the development of TMC diagnostic tests, including:

- Requires more time and effort to develop TMC requires more time and effort to develop compared to other multiple-choice test models. An effective TMC diagnostic test requires different levels of difficulty and complexity at each question level, thus requiring careful thought and planning in test development.
- Difficulty in evaluating the correct answer In the TMC model, the correct answer often consists of information obtained from lower levels. Therefore, determining the correct answer can be more difficult and requires more careful interpretation than other multiple-choice test models.
- Lack of flexibility TMC lacks flexibility in terms of scoring, as each question must be answered in a

specified order and answers at one level must be known before answering questions at a higher level. Therefore, this test model may not be suitable for all types of subjects or learning topics.

CONCLUSION

This three-tier multiple-choice diagnostic test instrument can be used in an effort to identify misconceptions about material properties and changes in the shape of objects. The three-tier multiple-choice diagnostic test instrument consists of three tiers. The first tier is related to the choice of answers to questions, with 1 choice and 3 distractors. The second tier is related to the choice of reasons from the response in the first stage, with 3 distractors and 1 answer. The third tier is a choice of confidence levels for responses in the first and second tiers, with closed options of sure and not sure.

The three-tier multiple-choice diagnostic test instrument developed has 13 valid questions and a reliability value of 0.685, which states that the test instrument is reliable. From the results of this study, it is known that the average percentage of students' understanding of the material nature of changes in form is 12.4% scientific knowledge, 3.3% lack of knowledge, 41% misconception (Ms), 25.4% misconception (false positive), 16% misconception (false negative), and 1.9% lucky guess. The last stage in this development is reflection after two field trials, so that a perfect product can be produced.

Three-tier multiple-choice was developed because it has aspects that are different from other diagnostic test instruments. Three aspects that differentiate TMC from other multiple-choice models are the three levels of questions, the hierarchical structure, and the emphasis on understanding and skills. In addition, it has advantages and disadvantages in terms of development. The advantage is that the TMC test is a useful assessment method for

measuring student understanding and provides more detailed teaching guidelines for teachers and educators. While the drawbacks require greater time and effort in development and are less flexible. Nonetheless, TMC remains an effective assessment method for measuring students' understanding and skills in various subjects.

THAN-YOU NOTE

Mr. Drs. H. Akhmad Nugraha, M.Si as supervisor 1 and Mrs. Agnestasia Ramadhani Putri, S.Pd., M.Pd as supervisor 2 who has helped a lot in giving guidance, instructions and suggestions to the writer.

BIBLIOGRAPHY

- Adriani, W., Hijrah Selaras, G., Yogica, R. (2019.). Identifikasi tingkat pemahaman konsep menggunakan three-tier multiple choice diagnostic test. *Atrium Pendidikan Biologi*. 1-8
- Alawiyah, N. S., Ngadimin, & Hamid, A. (2017). Identifikasi miskonsepsi siswa dengan menggunakan metode Indeks Respon Kepastian (IRK) pada materi impuls dan momentum linear di SMA Negeri 2 Banda Aceh. *Jurnal Ilmiah Mahasiswa (JIM)*, 2(2), 272–276.
- Arslan, H. O., Cigdemoglu, C., & Moseley, C. (2012). A three-tier diagnostic test to assess pre-service teachers' misconceptions about global warming, greenhouse effect, ozone layer depletion, and acid rain. *International Journal of Science Education*, 34(11), 1667–1686. <https://doi.org/10.1080/09500693.2012.680618>
- Basyiroh, U., Mulyani, B., & Ariani, S. R. D. (2022). Analisis kesulitan belajar pada materi kesetimbangan kimia dengan tes diagnostik three-tier multiple choice pada siswa kelas XI SMA Negeri 1 Surakarta tahun pelajaran 2020/2021. *Jurnal Pendidikan Kimia*, 11(1), 51-59.
- Dewi, S, Z & Ibrahim, H, T. (2019). Pentingnya pemahaman konsep untuk mengatasi miskonsepsi dalam materi belajar IPA di sekolah dasar. *Jurnal Pendidikan Universitas Garut*, 13(1), 130-136, www.jurnal.uniga.ac.id
- Faizah, N, S. (2017). Hakikat belajar dan pembelajaran. *At-Thullab: Jurnal Pendidikan Guru Madrasah Ibtidaiyah*, 1(2), 175-185.
- Kustiarini, F. T., Susanti VH, E., & Saputro, A. N. C. (2019). Penggunaan tes diagnostik three-tier test alasan terbuka untuk mengidentifikasi miskonsepsi larutan. *Jurnal Pendidikan Kimia*, 8(2), 171. <https://doi.org/10.20961/jpkim.v8i2.25236>
- Laksono, P, J. (2020). Pengembangan three tier Multiple Choice Test Pada Materi kesetimbangan kimia mata kuliah kimia dasar lanjut. *Orbital: Jurnal Pendidikan Kimia* 4 (1), 44-63
- Lubab, H., Alvy Siswanto, N., & Galuh Octaviani, R. (2023). Identifikasi miskonsepsi tanaman rimpang pada siswa kelas XI di MA Sabilul Ulum Mayong. *Jurnal Cendekia Ilmiah*, 2(3).
- Muyassaroh, I., Mukhlis, S., & Ramadhani A. (2023). Model project based learning melalui pendekatan STEM untuk meningkatkan kemampuan masalah siswa SD. *Jurnal Educatio*, 8 (4), 1607- 1616
- Natalia, D, P., & Sudrajat, A. (2023). Development of three-tier diagnostic test instrument to measure misconceptions of class XI students on reaction rate materials. *Jurnal Teknologi Pendidikan: Jurnal Penelitian dan Pengembangan*

- Pembelajaran, 8(1), 1-11.
- Nasution, R. H., Wijaya, T. T., Putra, M. J. A., & Hermita, N. (2021). Analisis miskonsepsi siswa SD pada materi gaya dan gerak. *Journal of Natural Science and Integration*, 4 (1), 11-21.
- Nazura, Saputri, D. F., Anggraeni, L. (2021). Pengembangan tes diagnostik three tier test pada materi pesawat sederhana untuk peserta didik kelas VIII SMP. *Jurnal Pendidikan Sains dan Aplikasinya (JPSA)* 4 (2) 54-60, <http://dx.doi.org/10.31571/jpsa.v4i2.2425>
- Permadani, S. P., Aini, J. D. N., & Thohir, M.A. (2022). Identifikasi miskonsepsi materi upaya pelestarian dan keseimbangan alam kelas V SDN Bendogerit 1. *Jurnal Pembelajaran, Bimbingan, Dan Pengelolaan Pendidikan*, 2(10), 916–924. <https://doi.org/10.17977/um065v2i102022p916-924>
- Plomp. (2013). *Educational design research educational design research*, 1–206.
- Roghdah, S. J., Zammi, M., & Mardhiya, J. (2021). Development of four-tier multiple choice diagnostic test to determine students' concept understanding level on thermochemical material. *Jurnal Phenomenon*, 11(1), 57-74.
- Savira, I., Wardani, S., Harjito, & Noorhayati, A. (2019). Desain instrumen tes three tiers multiple choice untuk analisis miskonsepsi siswa terkait larutan penyangga. *Jurnal Inovasi Pendidikan Kimia*, 13 (1), 2277 – 2286.
- Suryani, E., Rusilowati, A., & Wardono. (2016). Analisis pemahaman konsep IPA siswa SD menggunakan two-tier test melalui pembelajaran konflik kognitif. *Journal of Primary Education*, 5(1), 56-65 <http://journal.unnes.ac.id/sju/index.php/jpe>
- The Ministry of National Education. (2007) *Tes diagnostik*. Jakarta: Direktorat Pembinaan SMP pada Direktorat Jenderal Manajemen Pendidikan Dasar dan Menengah.
- Triumiana, D. A., & Sudarsono, F. (2017). Pengembangan instrumen tes diagnostik mata pelajaran fisika kelas XI SMA Negeri 2 Bantul semester genap 2017/2018. *Jurnal Penelitian dan Evaluasi Pendidikan*, 5(2), 45-54
- Wahyudi, F., Didik, L., A., Bahtiar. (2021). Pengembangan instrumen three tier test diagnostik untuk menganalisis tingkat pemahaman dan miskonsepsi siswa materi elastisitas. *RELATIVITAS: Jurnal Riset Inovasi Pembelajaran Fisika*. 4(2),48-58, <http://ojs.unimal.ac.id/index.php/relativitas/>