Jurnal Inovasi Pembelajaran Fisika
(INPAFI)
Available online http://jurnal.unimed.ac.id/2012/index.php/inpafi
e-issn 2549-8258, p-issn 2337-4624

THE DEVELOPMENT OF TWO-TIER MULTIPLE CHOICE TESTS TO ASSESS STUDENT'S CONCEPTUAL UNDERSTANDING IN PHYSICS LEARNING ASSISSTED BY ALGODOO

Nur Aini dan Wawan Bunawan
Jurusan Fisika FMIPA Universitas Negeri Medan
rhamidnuraini@gmail.com

Diterima: September 2020; Disetujui: Oktober 2020; Dipublikasikan November 2020

ABSTRACT

Physics learning has a relationship between one concept and another. There is an error in a concept, this can affect the understanding of other concepts, resulting in low student learning outcomes. Therefore, assistance is needed to measure student learning outcomes, for example by providing learning tests. This study aims to develop a two-tier multiple-choice test for the concept of "Newton's law of motion" in class XI SMA N 1 Percut Sei Tuan. This research is a development research using the 4D model to produce a product in the form of a two-tier multiple choice test on the material of Newton's law on modified motion without including the dissemination stage, namely: (1) define; (2) design; (3) developing. The final version of the multiple choice test consists of 12 items. The subjects of limited scale tests and broad-scale tests were class XI students of SMA N 1 Percut Sei Tuan in the 2019/2020 academic year. The student has completed a unit on Newton's Laws of motion. The reliability of the test was 0.76. Based on data analysis, ten alternative conceptions were identified. The results showed that a two-tier multiple choice test was effective in determining student misconceptions and could also be used as an alternative to traditional multiple choice tests. In conclusion, a two-tier multiple choice test can be used to assess students 'conceptual understanding as well as students' misconceptions about the concept of Newton's law material about motion.

Keywords: Two-Tier Multiple Choice Test, Conceptual Understanding, Misconceptions, Newton Law's About Motion Instruments

INTRODUCTION

Conceptual understanding in science learning has been the main concern of the researchers in the science education field. Students' conceptual understanding could not be easily observed or measured. Teachers have to investigate students' understanding before and after instruction. In order to measure students' conceptual understanding of several concepts in a science subject, various diagnostic tools have been developed and used such as open-ended tests, interviews, and multiple choice tests. Those are found to be the ones commonly employed in science education research (Gurel et al., 2015).

Multiple choice tests have been employed for measuring students' understanding of concepts since they allow a large number of students to be sampled in a provided amount of time as compared to time-consuming interviews. These tests are easy to administer and score; moreover, the results obtained are also easily processed and analyzed (Treagust et al., 1995). However, multiple-choice questions may not always indicate students' understanding or detect students' misunderstanding for a certain concept. The use of a two-tier diagnostic test has examined

a better way to evaluate students' conceptions. A two-tier diagnostic test was first developed with items precisely designed to identify auxiliary conceptions and misunderstandings in a defined content area of science. Since that time, a number of two-tier tests have been developed and reported in the literature (Treagust et al, 1995).

Two-tier tests have been considered as an effective assessment tool to establish students' conceptual understanding (Tuysuz. C. 2009). One of the factors affecting students' conceptual understanding is misconceptions. Misconceptions occur if a student's understanding of a concept differs from the scientific concept. Misconceptions are stable cognitive structures to change, affect students' conceptual understanding, and must be overcome so that students learn scientific concepts effectively (Hammer, 2000). Misconceptions have become a part in the science education area. Previous researchers have done lots of studies to investigate the students' misconceptions, particularly in Newton law's about motion instruments concept.

Physics is a branch of natural science, not only a collection of knowledge in the form of facts, concepts, or principles, but also a learning process that provides students with direct experience to understand the natural environment scientifically. Physics learning aims to improve students' mastery of knowledge, concepts, principles of physics, and develop student skills. Motion is an important concept for prospective students of Physics thus it is necessary to identify conceptions and causes of failure to understand and use Newton's laws of motion. Previous research has identified a level of understanding of the concept of Newton's Law. Handhika et al. (2016) tried to uncover students' conceptions of motion by reviewing the perceptions that students have. These studies have tried to uncover and map the level of understanding of Newton's law but have not revealed a more structured cause of failure to understand the concept of motion in accordance with Newton's law. Therefore, it is hoped that students will get a meaningful learning process when fulfilling all these aspects. Studies have also been conducted to determine the difficulties students have with Newton's laws. Newton's laws are important because they have easily visible applications in the daily lives of students.

This concept is expanded and taught in the upper grade in senior high school. If the students' misconceptions about newton laws about motion instruments concept are not corrected, students will carry these misconceptions to the upper grades. Dealing with this issue, the development of two-tier multiple choice test for evaluating students' conceptual understanding of newton laws about motion instrument may lead to more meaningful learning. Therefore, the purpose of this study was to develop a two-tier multiple choice test to assess students' conceptual understanding, as well as to explore students' misconceptions of newton laws about motion instruments concept.

RESEARCH METHODOLOGY

This type of research uses research and development (Research and Development) which aims to develop two-level multiple choice in Physics for XI MIA 4 students at SMA N 1 Percut Sei Tuan based on curriculum 13 (K13). The product being developed is a 4D model on Newton's law of motion. In this research and development, researchers only arrive at the feasibility test of the resulting product. According to Putra (2018: 133) one of the development models can be in the form of a procedural model with the development objectives that the researcher wants to achieve, namely to produce a product and to test the feasibility of the product. This development research uses the 4D development model developed by Sivasailam Thiagarajan, Dorothy S. Semmel, and Melvyn I. Semmel (Trianto, 2010: 189) which contains 4 stages, namely: Define, Design, and Develop and stages without including the dissemination stage.

1. Define

The research and development procedure includes the define stage, The purpose at this stage is to determine and define learning needs by analyzing the objectives and limits of the material carried out a preliminary study to gather information that includes: literature study and literature study. The stages of definition consist of: Frontend Analysis, Learner Analysis, Task Analysis, Concept Analysis and Specifying Instructional Objectives. Based on the science textbooks, the content area of Newton law's about motion concept was identified. Concept analysis is carried out to analyze competency standards and basic competencies, and analyze learning resources for Newton's Law about Motion. The example of a content area of TTMCT item can be seen in Table 1.

Table 1. The Content Area of Two-Tier Multiple Choice Tests

Concept	Material	Indicator	Cognitive Aspect			
			C1	C2	C3	C4
Newton Law's	Newton Law's About Motion	Students explain Newton's First Law		1		11
		Students explain Newton's II law	24	2,3	25	12
		Students explain Newton's III law		9,16	20	
		Students explain Newton's law problems		5	22	13
	Penerapan Hukum Newton dalam kehidupan sehari-hari	Students apply the I Newton's law in daily life		8	6,19, 21	
		Students apply Newton's second law calculations in daily life	7	4,15, 18,23	14,17	10
		Students apply the calculation of Newton's third law in daily life				

2. Design

This stage begins with developing the characteristics that are developed include the design of questions, design of Algodoo-based test programs. Activities at this stage include: (1) Criterion-Test Construction, (2) Media Selection, (3) format selection and (4) initial design.

3. Develop

The third stage was the development of two-tier multiple choice test. This stage focused on developing the two-tier multiple choice test. Based on the specification grid, 25 items two-tier multiple choice tests were developed. Each item of the instrument consists of two sections. In a Two-tier multiple choice test (drafting I), the first tier asked students to choose about some specific concept related with newton law's about motion instruments concept; and the second tier questioned students about the reason or explanation for choice in the first tier. There were five choices for both tiers. The first version of the instrument was validated by two science lecturers, one science teacher, and one science student. The validation results will be tested on a small scale with 32 students. Then, after conducting a small-scale trial, it will then be tested on a large scale for the final TTMCT draft to be developed. An example of a TTMCT item can be seen in Figure 1.

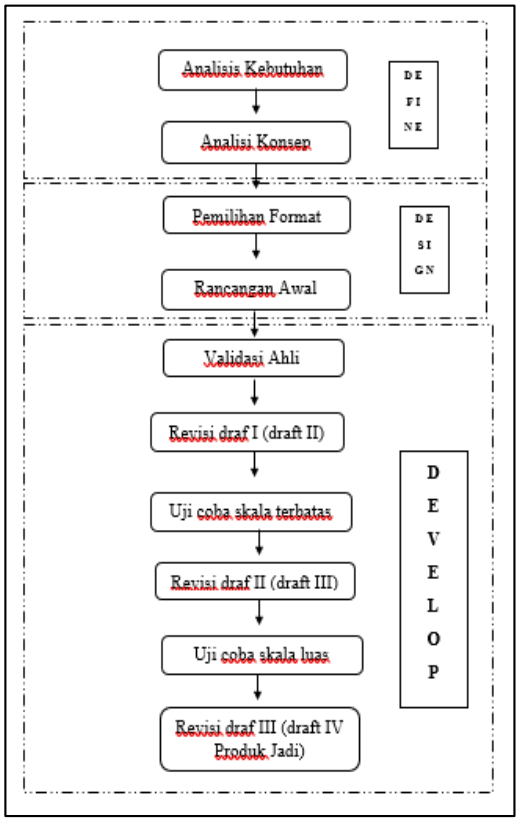


Figure 1. Learning System Development Model

RESULT AND DISCUSSION

The product being developed is a two-stage multiple choice test instrument with Newton's Law of motion as material. This development research uses the 4D model,

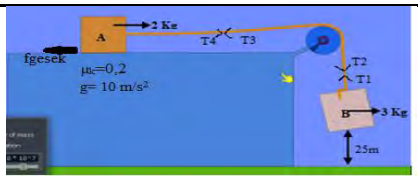
namely: define, design, development and disseminate. In research only at the development stage. The stages can be described as follows:

1. Define

This stage the objective is to analyze needs through field studies, literature studies and literature studies. Based on the field study conducted, the following information was obtained: (a) the teacher has not developed the test because it is not practical. They only use daily test tests as an indicator of the achievement of student competencies to map the level of understanding of students' concepts on a certain material. Furthermore, literature studies and literature studies obtained information that (a) Two-tier multiple choice has an advantage because in this test in addition to students working on test items that express certain concepts students must also reveal the reasons why choosing the answer (Suparno, 2005 : 8). By expressing their reasons in answering each question, the location of the misconception will occur. In addition, two-tier multiple-choice tests are easy to administer and also easy for teachers to provide assessments . (b) Algodoo has been used in research in the field of education in several countries. A study in Brazil shows that based on test results, students show a good and fast understanding of physics content when they use Algodoo simulations (Samir L, 2014). In this case, algodoo media as a supporting animation maker for the question.

2. Design

This stage begins with developing the characteristics that are developed include the design of questions, design of Algodoo-based test programs. The format used in the development of the test instrument was a closed multiple choice consisting of two tier. The first stage contains multiple choice questions with four distractors and one answer key. The second tier contains five choices of students' reasons for answering questions with one key of the correct reasons.



Balok *A* bermassa 2 kg dan balok *B* bermassa 3 kg dihubungkan dengan tali dan melalui katrol seperti gambar. Tentukan tegangan talinya . . . .

a. 7 N
b. 55 N
c. 24 N
d. 45 N
e. 10 N

Alasan jawaban:

a. tinjau benda B : $T = 30\text{-}3a$;
tinjau benda A : $T = 2a + 5$;
dengan metode subtitusi diperoleh nilai dari $a = 5\ ms^{-2}$;
Tegangan tali, $T = 45\ N$

b. tinjau benda B : $T = 30 + 3a$;
tinjau benda A : $T = 2a + 5$;
dengan metode subtitusi diperoleh nilai dari a = 7 ms−2;
Tegangan tali, $T = 7\ N$

c. tinjau benda B : $T = 30 - 3a$;
tinjau benda A : $T = 2a - 20$;
dengan metode subtitusi diperoleh nilai dari a = -5 ms−2;
Tegangan tali, $T = -45\ N$

d. tinjau benda B : $T = 30 + 3a$;
tinjau benda A : $T = 2a - 20$;
dengan metode subtitusi diperoleh nilai dari a = 2 ms−2;
Tegangan tali, $T = 24\ N$

e. tinjau benda B : $T = 30 + 3a$;
tinjau benda A : $T = -2a - 20$;
dengan metode subtitusi diperoleh nilai dari a = -10 ms−2;
Tegangan tali, $T = 10\ N$

Figure 2. The Example of TTMCT Item

3. Develop

The development stage, researchers create a two-level multiple-choice test. Then product was validated by 2 physics lecturers, 1 physics teacher, and 1 physics student. The validator provided a description and conceptual framework to complement the validity of the instrument. Validators or experts to provide an assessment of all test instruments developed in draft 1 which contain construction content, language, and time allocation provided in the instrument. The validation analysis method used is the Content Validate Ratio (CVR) method and the Content Validate Index (CVI) method. Of the 25 questions given, it shows that there are 5 items that are not used because the resulting CVR value is <0.99. Problems that were not good were revised according to the input of the expert team and the problems were rearranged. The results of the validation of multiple choice test questions are as follows: there are 9 questions that can be used without revision, 6 questions that can be used with a little revision, and 5 questions that cannot be used. After the

product is revised, this is obtained from the product with suggestions and improvements to the stage II product which will be tested on a trial-scale.

The results of the revision of draft I after the validation test were carried out into draft II which was tested on students to see student assessments and tested on the teacher to determine the teacher's response to the developed multiple choice test. The results of the limited scale trial given to 32 students at SMA N 1 Percut Sei Tuan in the form of an assessment questionnaire obtained an average of 67% with a fairly good category. and Teacher response interviews were conducted to find out the teacher's opinion on the developed two-tier multiple choice test. 2 physics teachers at SMA N 1 Percut Sei Tuan commented that the content of the instruments included almost 95% of the syllabus and suitable to be used. The language used in the TTMCT was easily understood by the students. The reliability test is essential to examine the consistency of the items measured using the instruments.

Each multiple choice test question developed has a different level of difficulty. The test items can be said to be good if the test items have a difficulty level at the interval 0.31-0.70, this shows that the items are not too difficult and not too easy. A recapitulation of the difficulty level of multiple choice test items shows that of the 7 items, most of them are in the medium category because it can be seen that as many as 7 items or 35% of the questions are in this category. The questions in the difficult category are 1 item or 5%. There were 12 questions in the easy category or 60% and there are no items that are categorized as very difficult or very easy. Based on the explanation above, the difficulty level can be seen from the number of test takers who answered correctly the questions compared to the number of students who took the test. The more people who answer a question, the more easy the question is, conversely the fewer questions are answered, the more difficult the question.

The effectiveness of options is one of the characteristics that determine the quality of an item. The effectiveness of the options referred to here is in finding out whether a problem is working or not. Based on the results of the analysis, it is known that of the 20 questions as many as 5 (25%) have a very good category of interference index, 2 questions (10%) have a good category of disturbance index, as many as 8 questions (40%) have an intrusive index with a fairly good category, as many as 3 questions (15%) had a disturbance index in the bad category, and as many as 2 questions (10%) had an intrusive index. in very bad category. It can be said that the bully has sufficient function to trick the test taker's answer, because the test taker in the low-ability group of students chooses the wrong answer so that he is deceived by the correct answer.

Differential power needs to be analyzed to differentiate the level of ability of each student in working on two-level multiple choice test questions. The number of questions that can be done correctly by high-ability students is a good number. The number of questions that all students could or could not do correctly was a problem with a bad difference. Question numbers that can be done well by students with low abilities are not good questions so they must be discarded (Arikunto, 2012). The recapitulation of the difference in test items shows that the distinguishing power of two-layer multiple choice test questions, namely 2 questions or 10% of the questions have very poor distinguishing power, 2 questions or 10% of the questions have poor distinguishing power, 9 questions or 45% of the questions have distinguishing power enough and 9 questions or 35% of the questions have good distinguishing power. In other words, a two-tier multiple choice test item is sufficient to differentiate high-ability students from low-ability students.

Broad-scale test result data that includes reliability, level of difficulty and power differences serve as guidelines for revising draft III to draft IV (finished product). The requirements to become draft IV of questions must be reliable, the level of difficulty in the category is moderate, sufficient, good and the distinction is very good. There are 16 questions that passed into the final product. Tests (draft IV / final product) were used to identify

student misconceptions on Newton's Law material. A total of 16 test questions were tested on 35 students with the following results.

The reliability of Wide-Scale in 35 students at SMAN 1 Percut Sei Tuan, After the calculation is done, the r11 is 0.53 and then it is compared with the r table. The price of rtable for the number of samples 35 with a significance of 5% is 0.33. This value is included in $0.70 < r \leq 1.00$ So it can be concluded that the test questions developed have good reliability. In a traditional multiple choice test with four possible choices, the chance of predicting the correct answer is 25 percent. Nevertheless, in a TTMCT, the chance of predicting is 6.25 percent. By lessening the predicting chance from 25 percent to 6.25 percent, the arithmetic means of the students might decrease. Treagust (1995) stated that the development of the two-tier diagnostic test to reveal students' conceptions. The first tier of each item test is a multiple choice question dealing with proportional statements, and the second tier of each item is a composed of multiple choice set of reasons for the first tier's answer. The set of reasons includes students' scientific answer and possible misconceptions. A student's answer to an item was declared correct if the student selected both the correct answer and reason. Items of the TTMCT were evaluated for both correct and incorrect response combinations chosen by the students.

1. Apabila resultan benda sama dengan nol, dapat dipastikan bahwa benda tersebut dalam keadaan. . . . a. bergerak lurus beraturan b. bergerak dengan kelajuan tetap **c. diam atau bergerak dengan kelajuan tetap** d. diam atau bergerak lurus beraturan e. bergerak dengan kecepatan berubah Alasan jawaban: a. Peristiwa tersebut menggambarkan Hukum Newton **b. peristiwa tersebut menggambarkan berlakunya hukum I Newton.** c. peristiwa tersebut menggambarkan berlakunya hukum II Newton d. peristiwa tersebut menggambarkan berlakunya hukum III Newton e. Peristiwa tersebut menggambarkan berlakunya hukum I dan II Newton	2. Muslimah berada di dalam lift yang bergerak ke atas dengan percepatan 0,4 ms-2. Jika massa Muslimah 40 kg dan percepatan gravitasinya 10 ms-2, maka gaya normal (*N*) yang bekerja pada Muslimah adalah . . . . a. 384 N b. 400 N **c. 416 N** d. 1000 e. 1600 N Alasan jawaban: a. $N = \frac{m.g}{a} = 1000N$ b. $N = w - m.\ a = 384$ N c. $N = w = 400$ N **d. $N = w + m.\ a = 416$ N** **e.** $N = m.g.a = 1600$ N

Item	Answer Option	Reason Option						Total
		a	b	c	d	e	Blank	
Q1	A	1,00	2,20	1,00	0,00	0,00	0,00	4,20
	B	1,05	3,16	1,05	0,00	0,00	0,00	5,26
	C	20,16	**56,84***	3,16	0,00	3,00	0,00	83,16
	D	1,11	1,00	0,00	0,00	0,00	0,00	2,11
	E	1,00	2,00	1,11	0,00	1,16	0,00	5,27
	blank	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Q2	A	3,16	1,05	2,11	0,00	0,00	0,00	6,32
	B	0,00	2,11	4,21	3,16	0,00	0,00	9,48
	C	12,24	8,42	12,63	**35,79***	0,00	0,00	68,96
	D	0,00	0,00	2,00	1,00	3,00	0,00	6,00
	E	4,21	0,00	3,16	0,00	2,11	0,00	9,48
	blank	0,00	0,00	0,00	0,00	0,00	0,00	0,00

Note: Figure in bold and with an asterik indicates the correct answer. texts in italics indicate a major alternative conception (>15%)

Figure 3. The Example of TTMCT Item and Percentage of Students Selecting Each Response Combination for Item Number 1 and Number 2 Dealing with the Newton Laws about Motion

The Analysis of Alternative Conceptions Using Two-tier Multiple Choice Test

Alternative conceptions are considered significant and conventional if they were found in more than 10% of the students' sample. Table 3 shows the summarize of significant common alternative conceptions of students in newton laws about motion instruments concept using two-tier multiple choice tests. The findings from Table 3 illustrate that students hold misconceptions about Newton's law of motion. Based on data analysis, thirteen alternative conceptions were identified. Alternative conceptions of Newton's law material about motion arise because of the difficulty and complexity of concepts, daily life experiences, textbooks, the language used, and teacher misunderstandings. Students come to school with different knowledge of these concepts based on their daily life

Students' understanding find out of the concept of Newton's legal material instrument about motion using the two tier multiple choice test instrument, the first level asks a

student to choose about certain concepts, and the second level asks students about reasons or explanations for choices at the first level. The score of the TTMCT method is considered for each item to be answered correctly if the student chooses either the first level (content knowledge) or the second level (first level reason) both correct. Table 3 shows the analysis of the percentage of correct answers and correct reasons of the two-level multiple-choice test.

Table 3. The percentage of correct answer-correct reason from two tier multiple choice test of 35 student

Topic	Indicator of item Test	Item	Total	
			N	Percentage
Newton's Laws of Motion	Students explain Newton's First Law	1	28	80%
	Students explain Newton's First Law	2	27	77%
The application of Newton's Law in everyday life	Students apply Newton's second law calculations in everyday life	3	24	69%
Newton's Laws of Motion	Students explain Newton's legal issues	4	23	66%
The application of Newton's Law in everyday life	Students apply Newton's second law calculations in everyday life	5	26	74%
	Students apply Newton I law calculations in everyday life	6	20	57%
Newton's Laws of Motion	Students explain Newton's First Law	7	21	60%
Newton's Laws of Motion	Students explain Newton's legal issues	8	27	77%
The application of Newton's Law in everyday life	Students apply Newton's second law calculations in everyday life	9	21	60%
	Students apply Newton's second law calculations in everyday life	10	23	66%
	Students apply the calculation of Newton's law III in everyday life	11	21	60%
	Students apply Newton's second law calculations in everyday life	12	19	54%
Newton's Laws of Motion	Students explain Newton's legal issues	13	26	74%
	Students explain Newton's III law	14	18	51%
The application of Newton's Law in everyday life	Students apply Newton's second law calculations in everyday life	15	28	80%
Newton's Laws of Motion	Students explain Newton's II law	16	23	65%

The science curriculum in Indonesia is stated that the assessment is geared towards measuring conceptual understanding (Widiyatmoko & Shimizu, 2018). TTMCT in this study discusses conceptual understanding in two topics: (1) Newton's Laws of Motion; (2) The application of Newton's Law in everyday life.

The concept of Newton's legal material instrument concept of motion, Indonesian students have been provided with examples of the application of Newton's law in everyday life. For example, in Table 3 it is shown that 77% of students hold a misconception about "students explain Newton I law". Based on this conception, students have not been able to distinguish between "no force" with a resultant force equal to zero. The fact is $\sum F$ or the resultant force has meaning the total amount of force affecting an object so that what is zero is not the force affecting an object, but the total or resultant amount of force affecting the object that is zero.

This also means that Newton's First Law is still affected by force or $F \neq 0$ unless the object is in a vacuum. Another misconception is the eye accommodation process. Based on the analysis, 60% of students have a misconception about "Students apply Newton's third law calculations in everyday life." Based on this conception, students assume that the action force appears earlier than the reaction force and students think gravity affects the state of the normal force. The action force has the same value as the reaction force, The action force has the opposite direction to the reaction force, The action-reaction force occurs in different objects, The action-reaction force has a coincide catch point, The action force has a different meaning from the reaction force. FAB is not the same as the meaning of FBA, the action force and reaction force appear simultaneously and the normal force with gravity is not a pair of reaction action forces. with indicator.

Table 4. The percentage of Recap of the Concept Indicator Understanding

No	Indicator	Persentase	Category
1	Students explain Newton's First Law	67%	High

2	Students explain Newton's II law	63%	High
3	Students explain Newton's III law	51%	Middle
4	Students explain Newton's law problems	67%	High
5	Students apply the I Newton's law in daily life	51%	Middle
6	Students apply Newton's second law calculations in daily life	65%	High
7	Students apply the calculation of Newton's third law in daily life	54%	Middle

The results of the analysis explained that the students' conceptual understanding of Newton's law of low motion. The percentage of achievement of indicators in Newton's legal material regarding the motion of the concept of the instrument is above 60%. The highest percentage in Table 4 is 80% of students who are shown an understanding of students explaining Newton's first law and students applying Newton's second law calculations in everyday life. On the other hand, the lowest percentage was 351% of students showed that students understood the understanding of Newton's III law. These results indicate that a number of students do not know the correct answer and hold misconceptions about Newton's law of motion. The two-stage multiple-choice test in this study can assess students' conceptual understanding as well as analyze student misconceptions.

Two-tier tests have been used by previous researchers to identify students' misconceptions in science. A two-tier diagnostic test, as Treagust (1995) reported, was first developed with items specifically designed to identify alternative conceptions and misunderstandings in clearly defined content areas of science. Since that time, a number of two-tier tests have been developed and reported in the literature.

The use of two-tier multiple choice tests allows teachers to achieve students' conceptual understanding, and also to explore students' reasoning behind these ideas. Moreover, it facilitates the assessment of misconceptions of a larger sample of students in an effective way in science education research. Misconceptions are stable, unscientific conceptions that obstacle the real learning of individuals listed the properties of misconceptions as follows: (1) strong and stable cognitive structures; (2) Differ from scientific conception; (3) Affecting how students understand scientific explanations; and (4) Must be overcome, avoided and eliminated to achieve the scientific conceptions.

Research in overcoming students' misconceptions involves three main steps, namely developing diagnostic test instruments, analyzing the causes of misconceptions, and remediation of misconceptions. Misconceptions are difficult to replace with new, correct understandings; they consistently influence the effectiveness of further learning. This condition happens because of misconceptions was difficult to change. Students' misconceptions interfere with students' learning of scientific concepts. Overcoming students' misconceptions require teaching methods which provide chances for students to reveal their pre-concepts and dissatisfaction with their concepts, particularly in light and optical instruments concept. According to Indonesian's national curriculum, the beginning of eighth grade in junior high school is the stage prior to receiving formal instruction about light and optical instruments concept, and this concept is expanded and taught in the upper grade in senior high school. If the students' misconceptions about light and optical instruments concept are not corrected, students will carry these misconceptions to the upper grades.

The previous literature indicated that there are various advantages of using two-tier multiple choice tests. Two-tier multiple choice test provided a reliable and valid pencil-and-paper, easy to score instruments for the teacher to evaluate students' idea better. Furthermore, this test has been used to evaluate students' misconceptions (Treagust, 1995) and very useful as the instruments that provide the teachers with students' understanding of particular science concept. The test is more readily administered and scored than the other method. The TTMCT is relatively convenient for students to respond and more practical and valuable for teachers to use regarding reducing

guesswork, allowing for large-scale administration and offering insight students' reasoning.

The complexity and difficulty of the newton laws about motion instruments concept can cause the students' misconceptions. In the learning process, students will try to link the new knowledge to their cognitive structures. If the students have misconceptions, these will interfere with their learning and they will difficult to connect new knowledge with their existing knowledge. Because of this condition, students will difficult to achieve conceptual understanding in a learning process. Thus, the teacher should guide prerequisite concepts for the students as the bridging between students prior knowledge and the understanding of the concept being learning.

CONCLUSION & SUGGESTION

Based on this study, the development of two-tier multiple choice test is comfortably used to assess students' conceptual understanding as well as students' misconceptions on newton law's about motion instruments concept. The two-tier multiple choice test could help teachers to enhance students' knowledge level and prevent students' misconceptions.

Thus, two-tier multiple choice test help to improve teaching- learning process in the science classroom. This study exhibits several limitations. One of the limitations is that it lacks generalizability. Since the study involving a small number of participants, the findings from this study may not be generalized to the other contexts.

REFERENCES

Arikunto, S. 2012. *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara

Gurel, D. K. 2015. A Review and Comparison of Diagnostic Instrumens to Identify Students' Misconceptions in Science. *Eurasia journal of Mathemetics, Science & Technology Education*. 11 (5): 989-1008

Hammer, D. (2000). Student resources for learning introductory physics. American Journal of Physics, 68(S1), S52–S59. https://doi.org/10.1119/1.19520.

Putra, I.A. & Sujarwanto, E. (2018). Analisis pemahaman konseptual mahasiswa pada materi kinematika partikel melalui tes diagnostik. Jurnal Riset dan Kajian Pendidikan Fisika, 5(1), 10–16. http://dx.doi. org/ 10.12928/jrkpf.v5i1.8923.

Suparno, P. *2005. Miskonsepsi dan perubahan konsep dalam pendidikan fisika.* Jakarta: Grasindo

Trianto. 2010. Model Pembelajaran Terpadu, Konsep, Strategi dan Implementasinya dalam KTSP. Jakarta: Bumi Aksara

Treagust, D. F. 1995. Diagnostic Assessment of Students" Science Knowledge". In: Glynn, S.M, Duit, R. (Eds.), Learning Science in The Schools: Research Reforming Practice. Mahwah, New Jersey: Lawrence Erlbaum Associates. pp. 327-346.

Tüysüz,Cengiz. 2009. Development of two-tier diagnostic instrument and assess students' understanding in chemistry. Jurnal yang di publikasikan: Oleh *Scientific Research and Essay*

Widiyatmoko, A., & Shimizu, K. 2018. An overview of conceptual understanding in science education curriculum in Indonesia. *International Conference on Mathematics, Science and Education 2017 (*ICMSE2017*)*. 1