



## Effectiveness of Two-tier Multiple Choice Diagnostic Test for Analyzing Students' Misconceptions in High School Physics Learning

Nuning Rohmantika , Eko Setyadi Kurniawan, Sriyono

Universitas Muhammadiyah Purworejo

Jl. KH.A. Dahlan 3, Kabupaten Purworejo, 54111, Jawa Tengah, Indonesia

[|rohmantikanuning@gmail.com](mailto:rohmantikanuning@gmail.com)  | DOI: <https://doi.org/10.37729/radiasi.v15i2.2280> |

### Abstract

Physics is a subject related to the environment so that it requires students to study nature and apply it in life. However, in learning sometimes students have difficulty understanding physics concepts and cause students to experience misconceptions. Therefore, a two-tier multiple choice diagnostic test is needed to analyze students' misconceptions. This study aims to determine the effectiveness of the two-tier multiple choice diagnostic test to analyze students' misconceptions in terms of content validity, practicality, and effectiveness. The research model is instrument development. The research design used is true experiment research with pretest posttest control. The subject of this research is class X MIPA MA Negeri Purworejo. Data collection in this study used validation sheets, observation sheets, student response questionnaires, and two-tier multiple choice diagnostic tests. The results showed that: 1) content validation for the construct and language aspects obtained from 3 validators got an Aiken V index of 1 for all items included in the high validity category, 2) a practical two-tier multiple choice diagnostic test in terms of sheets, observations with reliable categories and good responses and responses from students, 3) student learning outcomes have increased with an N-gain of 0.5 which is included in the medium category and students' misconceptions are found on the subject of momentum, impulse, and collision by 41.20% which is at medium level. Thus the two-tier multiple choice diagnostic test is declared valid, practical, and effective so that it is feasible to analyze students' misconceptions in learning physics.

### Article Info:

Received:

28/08/2022

Revised:

12/09/2022

Accepted:

26/09/2022

**Keywords:** Effectiveness, Diagnostic test, Two-tier, Misconception



## 1. Introduction

Determinants of the quality of education can be seen from various aspects, one of which is learning [1]. Quality education will produce students who have the ability to think critically, innovatively, creatively, a sense of responsibility, and skills and have good character. Therefore, along with the times, educational innovations are needed. The 2013 curriculum requires students to master scientific skills [2]. Physics is one of the subjects related to nature, so to study phenomena related to physics students need direct interaction with nature. In addition, learning physics is considered a science that can explain natural phenomena with human images. Therefore, through physics learning, students can study the natural surroundings and apply it in everyday life factually [3]. In physics learning, students are directed to understand the concept of the material as well as possible. However, when learning takes place, students have difficulty in understanding the concepts in physics learning.

This makes it difficult for students to connect the material being studied with the knowledge used [4]. If the concepts possessed by students are not appropriate, it will lead to different interpretations of concepts, giving rise to misconceptions.

Misconception is an understanding that is different from scientific concepts that have been expressed or defined by experts [5]. The misconceptions that occur in physics learning are fatal, because the material in physics is related to each other. Students who experience misconceptions have a consistent wrong understanding [6]. If students experience misconceptions on basic concepts, it will allow students' misconceptions to get bigger. This is because students form their own understanding into the basic concepts that are believed to be true, where the concept comes from the environment or surrounding media. Sometimes the concepts that students already have are not the same as those agreed upon by the experts, giving rise to misconceptions. Misconceptions that occur in students need to be changed so as not to experience prolonged misconceptions, one method that can be used to identify misconceptions is to perform diagnostic tests. Through the diagnostic test, the teacher can see and find out where the students' misconceptions are based on the answers [7]. Students' misconceptions cannot be identified using a single-level multiple choice test because it allows students to answer questions by guessing so that the reason for the students' answers is unknown. The two-tier multiple choice diagnostic test is one of the tests that can be used to identify and analyze students' misconceptions in learning physics [8].

Two-tier Multiple Choice is a multiple choice question with two levels which was first developed by David F. Treagus [9]. The first level of this test consists of multiple choices to answer the concept of the material, while the second level contains multiple choices about the choice of reasons for the answers to the first level. Two-tier multiple choice is very effective in identifying students' misconceptions, besides that it can be used as a tool for evaluating student learning outcomes [10]. In the field of education and especially in physics subjects, the identification of students' misconceptions has not been done much. The teacher assumes that if the student's score has reached the KKM, the child has understood the material presented. However, this is not necessarily true because the teacher has not identified the students' misconceptions because the tests used in the assessment of physics learning outcomes so far have only been one-level multiple choice, resulting in the teacher not knowing exactly where the students' misconceptions [11]. Besides coming from the students themselves, misconceptions are also caused by other things, one of which is the delivery of physics subject matter by the teacher so that it affects student learning outcomes [12].

From various materials in physics lessons, students have difficulty in understanding the concept of material, one of which is the material of momentum, impulse, and collision. Based on the results of previous research conducted by Lusiana., et al, it is proven that the students' misconceptions on the material of momentum, impulse, and collision are highest in the submaterial of the law of conservation of momentum and collision [13]. In addition, the results of other studies conclude that students experience misconceptions on questions related to the concept of momentum in everyday life [12].

Based on the researcher's observations and the results of interviews with physics subject teachers conducted at the MA Negeri Purworejo, information was obtained that teachers had not identified students' misconceptions in learning physics. The scores of students who are below and above the KKM are almost balanced so that Huru only assumes that students who have scores above the KKM have a good understanding of the concept. On the other hand, the questions used for daily tests, mid-semester tests, and final tests have not been able to identify students' misconceptions, so that students' misconceptions in learning physics have not been identified.

In assessing learning outcomes, teachers only use ordinary one-tier multiple choice questions. This results in most teachers not being able to identify students' misconceptions because there is no test instrument used to analyze misconceptions [14].

Referring to the description that has been presented, as an effort to analyze the misconceptions of students in learning physics, multiple choice questions are needed. Based on this background, this study intends to determine the effectiveness of a two-tier multiple choice diagnostic test to analyze students' misconceptions in high school physics learning on the subject of momentum, impulse, and collision.

## 2. Method

This research is part of development research of two-tier physics test instrument that refers to the modification of the Orindo & Antonio (1998) and Mardapi (2008). Instrument development model with the steps taken to develop this test, including: 1) test design, 2) trial, 3) measurement [15]. The subjects in this study were students of class X MIPA MA Negeri Purworejo in the 2021/2022 academic year. The research design to find effectiveness using true experiment with pretest posttest control [16].

The technique used to collect data in this study consisted of validation sheets, observations, questionnaires, and tests. The instruments used in this study were validation sheets, test implementation observation sheets, student response questionnaire sheets, and two-tier multiple choice diagnostic tests. Data analysis carried out is content validity, practicality, and effectiveness.

The content validation of the two-tier multiple choice diagnostic test for the construction and language aspects was compiled to obtain an assessment from the expert validator. Calculation of content validation using the Aiken V index presented in equation(1) [17].

$$V = \frac{\sum s}{n(c-1)} \quad (1)$$

To determining the validity category to determine the validity of the developed test, it is presented in Aiken V index interpretation. Score with interval  $V > 0.80$  is high category,  $0.40 \leq V \leq 0.80$  is moderate category, and  $V < 0.40$  in low category [15]. The test can be used if it has a moderate validity result.

The practicality of the two-tier multiple choice diagnostic test is seen from the observation of the implementation of the test carried out by two observers and the responses of students. The steps used to obtain observational data on the implementation of the test are to tabulate the data obtained and then calculated using the percentage agreement calculation in the equation (2)[18].

$$PA = \left(1 - \frac{A - B}{A + B}\right) \times 100\% \quad (2)$$

A and B are the scores given by the first and second observers, where  $A > B$ . The research instrument is declared reliable if the PA is more than one or equal to 75%. Furthermore, the PA criteria are interpreted using percantegge agreement criteria. Reliable on 76-100, enough reliable: 51-75, less reliable: 26-50, and unreliable: 0-25 [19].

Then to find out the student's response to the two-tier multiple choice diagnostic test, it was carried out using steps including recapitulation of student response questionnaire data. Next, tabulate the data obtained and calculate the percentage using the equation (3)[18].

$$NP = \frac{R}{SM} \times 100\% \quad (3)$$

Next step is to convert the percentage results into qualitative criteria based on the reference [18]. Criteria of students' response with percentage 86-100 is very good, 76-85: good, 65-75: enough, 55-59: less, and  $\leq 54$ : very less. The two-tier multiple choice diagnostic test is declared practical if the percentage of students' responses in interpretation is sufficient.

The effectiveness analysis was viewed from the results of the pretest and posttest of students and the percentage of students' misconceptions. The increase in the results of the students' pretest and posttest tests is calculated by the *N-Gain* shown in the equation (4) [20].

$$g = \frac{sf - si}{100 - si} \quad (4)$$

There are several criteria for increasing *N-Gain*, which high criteria  $g \geq 0.7$ ; medium  $0.3 \geq g > 0.7$ ; and low category  $g < 0.3$  [21]. The data from the pretest and posttest were then tested for normality with the aim of knowing that the data were normally distributed. Furthermore, to prove that the sample used came from a homogeneous population, a homogeneity test was carried out. Then to find out the difference between the students' pretest and posttest scores, a *t-test* was carried out.

The calculation of the percentage of students' misconceptions on the material of momentum, impulse, and collision is done by analyzing the data from the students' test results through a two-tier multiple choice diagnostic test. To find out the percentage of students' misconceptions, it is calculated using the equation (5) [22].

$$P = \frac{J}{N} \times 100\% \quad (5)$$

Furthermore, the percentage results are interpreted into several categories of student misconceptions which are presented on misconception category. High misconception between 61-100, medium: 31-60; and low category: 0-30 [4].

### 3. Result and Discussion

This research is part of the research on the development of two-tier test questions. This study to find the effectiveness of a two-tier multiple choice diagnostic test to analyze students' misconceptions in learning physics. Content validity for construction and language aspects aims to determine the feasibility of a two-tier multiple choice diagnostic test used to analyze students' misconceptions. This validation was carried out by 3 experts. Based on the results of the validity calculation using the Aiken V index, of the 30 questions tested for validity, a score of 1 was obtained for all items. The score indicates that the questions used in this study are in high validity criteria. These results are relevant to the results of other studies, namely if the Aiken score is above 0.75 then the diagnostic test is declared valid and feasible to use [23]. So based on content validation for construction and language aspects carried out by expert validators, it can be concluded that the two-tier multiple choice diagnostic test is valid or feasible to be used to analyze students' misconceptions on momentum, impulse, and collision materials.

The practicality of the two-tier multiple choice diagnostic test was reviewed through the test implementation observation sheet and student response sheets. The test implementation observation sheet was carried out by 2 observers who supervised the test. Observation of the implementation of the test is calculated using the percentage agreement and obtained a percentage of 95% so that it is in the reliable category. This proves that the implementation of the learning evaluation can be carried out properly starting from the beginning to the end of learning [24]. In this case, the assessment of student learning outcomes uses a two-tier multiple choice diagnostic test, so that based on the observations of the observer, it shows that the two-tier multiple choice diagnostic test is practical to use in the evaluation of physics learning.

A part of the observation sheet on the implementation of the test carried out by the observer, the practicality of the two-tier multiple choice diagnostic test was also reviewed from the results of the response questionnaire to determine student responses. Student response questionnaires were given after students took a two-tier multiple choice diagnostic test. The calculation of the student response questionnaire is presented in Table 1.

**Table 1.** Results of Student Response

No.	Aspect	Score	Percentage (%)	Category
1.	Eligibility of Content	716	82.87%	Good
2.	Language	352	83.80%	Good
	Average	539	83.33%	Good

Based on the student response questionnaire data in Table 1 shows that from the aspect of content feasibility, a score of 716 is obtained. If expressed in percentage, it gets a score of 82.87% and is included in the good category. The language aspect gets a score of 352, if expressed in percentage it gets a percentage of 83.80% and is in the good category. All aspects obtained a total score of 539 with a percentage of 83.33% so that they obtained a good category based on the reference in Table 4. These results are similar to the results of other studies which conclude that the location of the practicality of the assessment instrument is between 60% NRS 80%, then the criteria are strong [25]. From the results of student responses through student response questionnaires, information is obtained that the sentences used in the questions are easy to understand and understand making it easy for them to work on test questions. In addition, the language used does not cause multiple interpretations so that students have no difficulty in understanding the meaning of the questions. Through the diagnostic test questions given, students stated that they were more motivated to understand the concept of physics because in doing this test they were required to know the reasons besides only doing calculations using existing equations.

The effectiveness of the two-tier multiple choice diagnostic test was reviewed based on the results of the pretest and posttest of students in the control and experimental classes. Pretest is conducted in each class, it aims to determine the ability of students before being given treatment in the form of learning on the subject of momentum, impulse, and collision. In working on the pretest questions, the control and experimental classes were given the same questions. After being given the pretest, then each class carried out the same learning activities, namely the topics of momentum, impulse, and collision. At the end of the lesson, the control class was given a posttest in the form of a single-level multiple choice, while the experimental class was given a posttest of a two-tier multiple-choice diagnostic test. The recapitulation of the results of the pretest and posttest for the control and experimental classes is presented in Table 2.

**Table 2.** Students' Pretest and Posttest Scores

No	Aspects	<i>Control class</i>		<i>Experiment Class</i>	
		<i>Pretest</i>	<i>Posttest</i>	<i>Pretest</i>	<i>Posttest</i>
1	Subject	54	54	54	54
2	Mean	48.14	70.92	49.19	74.62
3	Median	46.67	70.00	51.67	74.58
4	Maximum	73.33	86.67	80.00	93.33
5	Minimum	20.00	53.33	20.00	53.33
6	Std Deviasi	11.33	7.53	13.45	8.18



Based on Table 2, information is obtained that students in the experimental class have higher scores than the control class. This proves that the score of students who get a posttest in the form of a two-tier multiple choice diagnostic test is higher than those who get a one-level multiple choice question. It can be seen that the posttest mean of students in the control class is 70.92 while in the experimental class it is 74.62, these results show that the two-tier multiple choice diagnostic test is effectively used to assess student learning outcomes. In addition, the maximum value obtained in the control class posttest was 86.67 while the experimental class obtained 93.33 results, this shows that the experimental class better understands the concept of physics after being given a two-tier multiple choice diagnostic test. Furthermore, to prove that the two-tier multiple choice diagnostic test is effective for analyzing students' misconceptions, it can be seen from the results of the pretest and posttest of each class. The answers to the pretest and posttest of one student from the control class are presented in Figure 2.

12. Dimas memukul bola kasti menggunakan pemukul dengan memberikan gaya sebesar 50 N untuk setiap pukulan. Waktu kontak antara pemukul dan bola kasti berbeda-beda seperti ditunjukkan pada tabel berikut. Pernyataan yang benar berdasarkan tabel tersebut adalah ....

Pukulan ke-	Waktu kontak
1	0,02 sekon
2	0,04 sekon
3	0,05 sekon
4	0,10 sekon

A. Impuls terbesar terjadi pada pukulan ke-2  
 B. Impuls terkecil terjadi pada pukulan ke-4  
 C. Impuls yang terjadi pada pukulan ke-3 lebih kecil dibandingkan pada pukulan ke-4  
 D. Impuls yang terjadi pada pukulan ke-4 lebih kecil dibandingkan pada pukulan ke-2  
 E. Impuls yang terjadi pada pukulan ke-1 lebih besar dibandingkan pada pukulan ke-4

12. Dimas memukul bola kasti menggunakan pemukul dengan memberikan gaya sebesar 50 N untuk setiap pukulan. Waktu kontak antara pemukul dan bola kasti berbeda-beda seperti ditunjukkan pada tabel berikut. Pernyataan yang benar berdasarkan tabel tersebut adalah ....

Pukulan ke-	Waktu kontak
1	0,02 sekon
2	0,04 sekon
3	0,05 sekon
4	0,10 sekon

A. Impuls terbesar terjadi pada pukulan ke-2  
 B. Impuls terkecil terjadi pada pukulan ke-4  
 C. Impuls yang terjadi pada pukulan ke-3 lebih kecil dibandingkan pada pukulan ke-4  
 D. Impuls yang terjadi pada pukulan ke-4 lebih kecil dibandingkan pada pukulan ke-2  
 E. Impuls yang terjadi pada pukulan ke-1 lebih besar dibandingkan pada pukulan ke-4

Figure 1. Control Class Pretest and Posttest Student Respons

Figure 1 show the answers to the pretest and posttest of students from the control class, namely the class that received the posttest in the form of one-level multiple choice. From the answers given by the students, both the pretest and posttest answers were not correct. Based on the concept of impulse material, it has been explained that the impulse is directly proportional to the force and contact time. This explains that if the contact time between two objects is longer, the impulse is also getting bigger. In this problem the ball is hit with the same force on each stroke but has a different contact time, so based on the data in the problem it can be seen that the order of the largest impulses is in the 4th, 3rd, 2nd, and 1st strokes. From Figure 1 it can be seen that students in the control class still gave incorrect answers before being given treatment. This shows that students still do not understand the concept of impulse well. While the answers to the pretest and posttest of students in the experimental class are shown in Figure 2.

12. Dimas memukul bola kasti menggunakan pemukul dengan memberikan gaya sebesar 50 N untuk setiap pukulan. Waktu kontak antara pemukul dan bola kasti berbeda-beda seperti ditunjukkan pada tabel berikut. Pernyataan yang benar berdasarkan tabel tersebut adalah ....

Pukulan ke-	Waktu kontak
1	0,02 sekon
2	0,04 sekon
3	0,05 sekon
4	0,10 sekon

A. Impuls terbesar terjadi pada pukulan ke-2  
 B. Impuls terkecil terjadi pada pukulan ke-4  
 C. Impuls yang terjadi pada pukulan ke-3 lebih kecil dibandingkan pada pukulan ke-4  
 D. Impuls yang terjadi pada pukulan ke-4 lebih kecil dibandingkan pada pukulan ke-2  
 E. Impuls yang terjadi pada pukulan ke-1 lebih besar dibandingkan pada pukulan ke-4

12. Dimas memukul bola kasti menggunakan pemukul dengan memberikan gaya sebesar 50 N untuk setiap pukulan. Waktu kontak antara pemukul dan bola kasti berbeda-beda seperti ditunjukkan pada tabel berikut. Pernyataan yang benar berdasarkan tabel tersebut adalah ....

Pukulan ke-	Waktu kontak
1	0,02 sekon
2	0,04 sekon
3	0,05 sekon
4	0,10 sekon

A. Impuls terbesar terjadi pada pukulan ke-2  
 B. Impuls terkecil terjadi pada pukulan ke-4  
 C. Impuls yang terjadi pada pukulan ke-3 lebih kecil dibandingkan pada pukulan ke-4  
 D. Impuls yang terjadi pada pukulan ke-4 lebih kecil dibandingkan pada pukulan ke-2  
 E. Impuls yang terjadi pada pukulan ke-1 lebih besar dibandingkan pada pukulan ke-4

ALASAN \*

A. Semakin lama waktu kontak antara dua benda, maka impulsnya semakin kecil  
 B. Semakin lama waktu kontak antara dua benda, maka impulsnya semakin besar  
 C. Semakin singkat waktu kontak antara dua benda, maka impulsnya semakin besar  
 D. Waktu kontak antara dua benda tidak mempengaruhi besar impuls yang terjadi pada benda  
 E. Besar kecilnya impuls ditentukan oleh waktu kontak antara dua benda

Figure 2. Control Class Pretest and Posttest Student Respons

Students' answers to the same questions for the experimental class using a two-tier multiple choice diagnostic test can be seen in Figure 2. The students' answers before being given treatment answered incorrectly as shown in Figure 2. However, the answers changed after the experimental class was given. The treatment was then given a posttest in the form of a two-tier multiple choice diagnostic test. From the posttest answer in Figure 2, it can be seen that the answer in the first tier is correct, namely the impulse that occurs in the 3rd stroke is smaller than the 4th stroke. This is because the longer the contact time between two objects, the greater the impulse. Based on the answers given by students from the experimental class, the answers from the first and second tiers are correct, so it can be concluded that they understand the concepts used in the question.

The results of the pretest-posttest control and experimental classes show that the two-tier multiple choice diagnostic test can make students understand the material being taught. This is evidenced by the answers of students in the control class who only use one-level multiple choice questions that have not been able to answer correctly according to the concept of physics, namely the concept of impulse. While the answers of students from the experimental class seemed to be correct in answering the first tier and the reasons for choosing the answer. So that these students can be concluded to understand the concept because the answers to the first and second tiers are correct. These results prove that the two-tier multiple choice diagnostic test is effective in increasing students' understanding of the material of momentum, impulse, and collision as well as being able to analyze whether these students do not understand concepts, misconceptions, or understand concepts.

Student score increase in the pretest and posttest can be calculated using the *N-gain*. The results of the *N-gain* calculation show on Table 3.

Table 3. Calculation of *N-Gain*

No.	Class	Pretest	Posttest	<i>N-gain</i>
1.	Control	48.14	70.92	0.43
2.	Ekspériment	49.19	74.62	0.50

The results of the calculation of *N-gain* in the control class obtained an increase of 0.43 and in the experimental class an increase of 0.50 was obtained so that each class experienced an increase in *N-gain* in the medium category. Even though it is at a moderate level, based on these results, it can be seen that the increase in the experimental class is higher than the control class. The class that received the posttest in the form of a two-tier multiple choice diagnostic test experienced a higher *N-gain* increase because through the diagnostic test students were required to understand concepts in addition to only doing calculations. Based on acquisition of *N-gain* that two-tier multiple-choice diagnostic test is more effective for analyzing students' misconceptions. To find out that the pretest and posttest data were normally distributed, a normality test was carried out.

The results of the normality test for control class pretest obtained 0.243 and posttest 0.460; experiment class obtained pretest 0.775 and posttest 0.091. Based on the normality test table, if the results obtained are  $\text{Asym. Sig. (2-tailed)} > 0.05$ , it can be concluded that the distribution of pretest and posttest data in the control and experimental classes is normal. The results of the above calculations are obtained that  $\text{Asym. Sig. (2-tailed)}$  pre-test control class is 0.243, while for  $\text{Asym. Sig. (2-tailed)}$  in the posttest obtained 0.460, because the significance value of the pretest and posttest control class is greater than 0.05, meaning that the data is normally distributed.

The pretest and posttest data in the experimental class, Asym results were obtained. Sig. (2-tailed) of 0.775 and 0.091, which is also greater than 0.05, so the pretest and posttest data in the experimental class are normally distributed. From the calculation of the normality test, it can be concluded that the control and experimental classes have pretest and posttest data that are normally distributed. Furthermore, to find out the pretest and posttest data came from the same population, a homogeneous test was carried out. Table 4 shows the calculation of the homogeneity test of the pretest and posttest data.

**Table 4.** Homogeneity Test Calculation

No.	Tes	F <sub>count</sub>	F <sub>table</sub>	Sig.
1.	Pretest	0.192	3.93	0.662
2.	Posttest	2.086		0.152

The homogeneity test from Table 4 shows that the pretest  $F_{count}$  is 0.192 and the posttest is 2.086 with a significance level of 0.05, from the data it shows that  $F_{count} < F_{table}$ , it means that the pretest and posttest data come from a homogeneous population.

After carrying out the normality test, a paired *t*-test was carried out to see the difference in the students' pretest and posttest scores. The results of the calculation of the paired *t*-test are shown in Table 5.

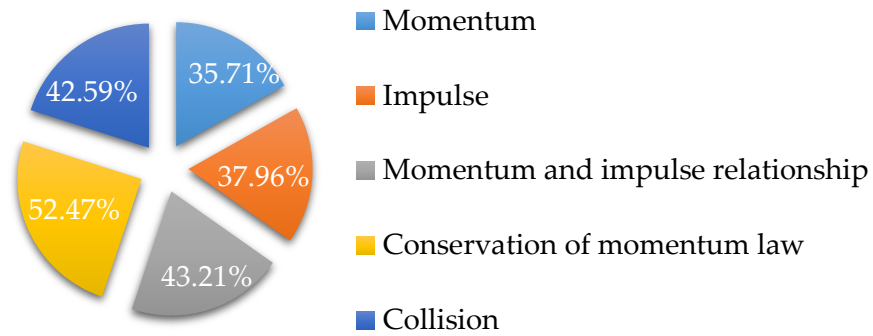
**Table 5.** Calculation of Paired t-test

No.	Class	Tes	df	t <sub>count</sub>	t <sub>table</sub>	Sig. (2-tailed)
1.	Control	Pretest & Posttest	53	-16.227	0.00575	0.000
2.	Eksperiment	Pretest & Posttest	53	-13.450	0.00575	0.000

The test used to determine the difference between the pretest and posttest is the result of significance. The provisions are if the significance is greater than 0.05 then  $H_0$  is rejected while  $H_a$  is accepted, but if the significance value is less than 0.05 then  $H_0$  is accepted and  $H_a$  is rejected.  $H_0$  is a hypothesis which states that there is a difference in the results of the pretest and posttest, while  $H_a$  is a hypothesis which states that there is no difference in the results of the pretest and posttest. The calculation results show that the significance value in the control and experimental classes is the same, namely 0.000, from these results it can be concluded that  $H_0$  is accepted and  $H_a$  is rejected. These results prove that there are significant differences in the results of the pretest and posttest of students.

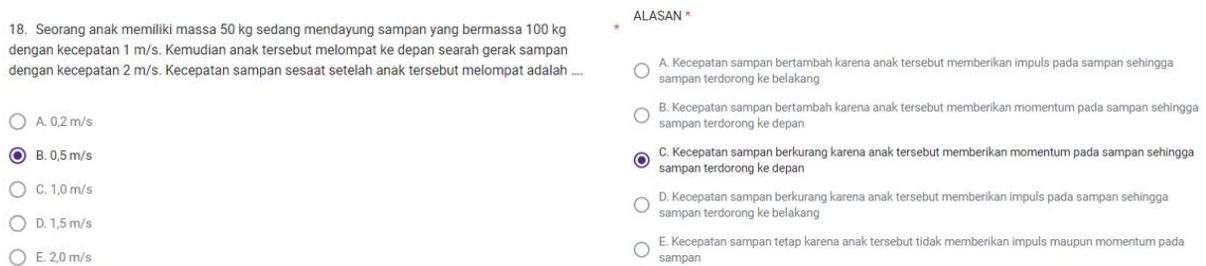
The purpose of the two-tier multiple choice diagnostic test is to be able to analyze students' misconceptions so that based on the results of the analysis the teacher can correct them so that similar misconceptions do not occur. Based on the posttest results of the experimental class using a two-tier multiple choice diagnostic test, it shows that as many as 41.20% of students have misconceptions and 45% of students understand the concept while the other 13.80% do not understand the concept. Referring to the category of misconceptions in Table 6, the misconceptions are at a moderate level. The percentage gain proves that the two-tier multiple choice diagnostic test can classify students' conceptual understanding abilities [26]. The percentage of students' misconceptions in each sub-material of momentum, impulse, and collision is shown in Figure 3.





**Figure 3.** Students' Misconceptions

Based on Figure 3, information can be obtained that students' misconceptions in different sub-materials, including momentum of 35.71%, impulse of 37.96%, momentum and impulse relationship of 42.21%, conservation of momentum law of 52.47%, and collision of 42.59%. From the percentage of students' misconceptions that exist, all are in the medium level of misconception. Students experience the highest misconception in the sub-material of the law of conservation of momentum, they have difficulty answering questions on this matter. Most students make mistakes in choosing the reason or second tier. The highest misconception experienced by students when they are asked to solve problems regarding the application of the law of conservation of momentum in everyday life, one of which is the case of a child jumping from a canoe. Students are able to calculate the speed of the canoe after the child jumps, but cannot explain the cause and direction of the canoe shortly after a child jumps. This shows that students have difficulty in understanding the concepts applied to the law of conservation of momentum, they only calculate using existing physical equations. One sample of the students' answers of questions about the momentum law of conservation sub-material is presented in Figure 4.



**Figure 4.** Student Answers of The Law Conservation Momentum

Based on the data of students who experienced misconceptions, students who answered correctly in the first tier were more than students who answered correctly in the second tier. In this question, students are asked to calculate the speed of the canoe after a child jumps from the canoe. The problem can be solved by the equation of the law of conservation of momentum to determine the final speed of the boat. However, students did not understand the reason for the reduced speed of the canoe and which direction the canoe finally moved due to the impulse given by a jumping child. This can be seen from the answers of students in the second tier which are mostly wrong.

Meanwhile, of the 5 sub-materials, the lowest misconception is found in the momentum sub-material. In this question, students are able to complete and answer questions correctly with the right reasons. Most of the students already understand that momentum is determined by mass and velocity and can apply it in everyday life.

If the mass and velocity are greater, the momentum will be greater because they are directly proportional to each other, so that students do not experience high misconceptions in this sub-chapter. This is evidenced by the correct answers of students in performing calculations and giving the right reasons for the selected answers. The presentation inline with [27] proves that the two-tier multiple choice diagnostic test is effective for analyzing students' misconceptions in learning physics so that it can help teachers to determine students' understanding of the material being taught.

## 4. Conclusion

Based on the analysis and presentation of the research results on the effectiveness of the two-tier multiple choice diagnostic test, the test is declared valid, practical, and effective so that it is appropriate to be used to analyze students' misconceptions in learning physics on the subject of momentum, impulse, and collision. The conclusions that can be drawn from this study are: 1) the two-tier multiple choice diagnostic test has high validity criteria based on the results of content validity calculations for the construct and language aspects carried out by expert validators so that it is said to be valid and feasible to use, 2) two diagnostic tests -tier multiple choice has met the practicality criteria based on data from the observation results of the implementation of the test which shows reliable criteria and the responses of students who meet the criteria are good, 3) the two-tier multiple choice diagnostic test is declared effective. It is suggested for teachers to apply a two-tier multiple choice diagnostic test to analyze students' misconceptions. For further researchers, the two-tier multiple choice diagnostic test can be developed in other subjects and other subjects with better content. In addition, the two-tier multiple choice diagnostic test is expected to be developed at different school levels to obtain various information and variations in student answers. This research should be used as a reference for further research by relating aspects of learning that are not discussed and have not been developed in this study.

## Acknowledgement

The researcher would like to thank the University of Muhammadiyah Purworejo and the MA Negeri Purworejo who have allowed and assisted in conducting the research as well as various parties who assisted in the completion of this research.

## References

- [1] S. Suryaningsih dan F. A. Nisa, "Kontribusi STEAM Project Based Learning dalam Mengukur Keterampilan Proses Sains dan Berpikir Kreatif Siswa," *J. Pendidik. Indones.*, vol. 2, no. 6, hlm. 1097–1111, 2021.
- [2] S. S. Lumbantobing dan S. F. Azzahra, "Meningkatkan Kemampuan Berpikir Kreatif Siswa Dalam Menghadapi Revolusi Industri 4.0 Melalui Penerapan Pendekatan STEAM (Science, Technology, Engineering, Arts And Mathematics)," hlm. 8, 2020.
- [3] C. Damayanti, A. Rusilowati, dan S. Linuwih, "Pengembangan model pembelajaran IPA terintegrasi etnosains untuk meningkatkan hasil belajar dan kemampuan berpikir kreatif," *J. Innov. Sci. Educ.*, vol. 6, no. 1, hlm. 116–128, 2017.
- [4] M. Ali, "Analisis Miskonsepsi Siswa Berdasarkan Gender dalam Pembelajaran Fisika dengan Menggunakan Tes Diagnostik Two-tier Di Kotabaru," vol. 7, no. 1, hlm. 8, 2019.
- [5] I. Savira dan S. Wardani, "Desain Instrumen Tes Three Tiers Multiple Choice Untuk Analisis Miskonsepsi Siswa Terkait Larutan Penyangga," vol. 13, no. 1, hlm. 10, 2019.

- [6] S. Winarni dan S. Syahrial, "Miskonsepsi Kimia yang Disebabkan Pernyataan Nonproposisi," *J. Pendidik. Sains*, vol. 4, no. 4, hlm. 122–129, 2016.
- [7] Q. Fariyani dan A. Rusilowati, "Pengembangan four-tier diagnostic test untuk mengungkap miskonsepsi fisika siswa sma kelas x," *J. Innov. Sci. Educ.*, vol. 4, no. 2, 2015.
- [8] D. S. Jubaedah, I. Kaniawati, I. Suyana, A. Samsudin, dan E. Suhendi, "Pengembangan tes diagnostik berformat four-tier untuk mengidentifikasi miskonsepsi siswa pada topik usaha dan energi," 2017, vol. 6, hlm. SNF2017-RND.
- [9] F. M. Dawati, S. Yamtinah, S. B. Rahardjo, A. Ashadi, dan N. Y. Indriyanti, "Uji Validitas Computerized Two-Tier Multiple Choice (CTTMC) Melalui Focus Group Discussion (Fgd) Untuk Mendiagnosis Kesulitan Belajar Siswa," 2017, hlm. 260–265.
- [10] A. A. I. Ardiansyah dan E. S. Bahriah, "Persepsi siswa terhadap penggunaan instrumen three-tier multiple choice berbantuan microsoft office excel," *EduChemia J. Kim. Dan Pendidik.*, vol. 1, no. 2, hlm. 204–214, 2016.
- [11] N. B. Haka, A. Hamid, N. Nurhidayah, A. D. Kesumawardhani, M. Rudhini, dan R. A. Riski, "Pengembangan Instrumen Evaluasi Two-Tier Multiple Choice Terhadap Literasi Sains Berbantuan Personal Computer," *Biosf. J. Tadris Biol.*, vol. 10, no. 2, hlm. 201–214, Des 2019, doi: 10.24042/biosfer.v10i2.5755.
- [12] S. R. Hidayat *dkk.*, "Pengembangan Instrumen Tes Keterampilan Pemecahan Masalah pada Materi Getaran, Gelombang, dan Bunyi," *J. Penelit. Pengemb. Pendidik. Fis.*, vol. 3, no. 2, hlm. 157–166, Des 2017, doi: 10.21009/1.03206.
- [13] N. Lusiana, L. Kurniawati, dan A. B. Mulyanto, "Analisis Miskonsepsi Siswa Pokok Bahasan Momentum Dan Impuls Di Kelas Xii Ipa. 4 Sma Negeri 4 Lubuklinggau Tahun Pelajaran 2015/2016," *J. Penelit. Pendidik. Fis.*, 2016.
- [14] I. Savira, S. Wardani, H. Harjito, dan A. Noorhayati, "Desain Instrumen Tes Three Tiers Multiple Choice Untuk Analisis Miskonsepsi Siswa Terkait Larutan Penyangga," *J. Inov. Pendidik. Kim.*, vol. 13, no. 1, 2019.
- [15] E. Istiyono, "Pengembangan instrumen penilaian dan analisis hasil belajar fisika dengan teori tes klasik dan modern," 2018.
- [16] D. Sugiyono, "Metode penelitian pendidikan pendekatan kuantitatif, kualitatif dan R&D," 2013.
- [17] E. Istiyono, W. B. Dwandaru, dan F. Rahayu, "The developing of creative thinking skills test based on modern test theory in physics of senior high schools," *J. Cakrawala Pendidik.*, vol. 37, no. 2, 2018.
- [18] M. Finnajah, E. S. Kurniawan, dan S. D. Fatmaryanti, "Pengembangan Modul Fisika SMA Berbasis Multi Representasi Guna Meningkatkan Pemahaman Konsep dan Hasil Belajar Peserta Didik Kelas XI IIS 2 SMA Negeri 1 Prembun Tahun Ajaran 2015/2016," *Radiasi J. Berk. Pendidik. Fis.*, vol. 8, no. 1, hlm. 22–27, 2016.
- [19] M. P. Trianto, "Mendesain model pembelajaran inovatif-progresif," *Jkt. Kencana*, 2009.
- [20] M. K. Majdi, B. Subali, dan S. Sugianto, "Peningkatan Komunikasi Ilmiah Siswa SMA melalui Model Quantum learning One Day One Question Berbasis Daily Life Science Question," *UPEJ Unnes Phys. Educ. J.*, vol. 7, no. 1, hlm. 81–90, 2018.
- [21] M. F. Mubarrok, "Penerapan pembelajaran fisika pada materi cahaya dengan media PhET simulations untuk meningkatkan pemahaman konsep siswa di SMP," *Inov. Pendidik. Fis.*, vol. 3, no. 1, 2014.
- [22] N. Nurhidayatulah dan A. K. Prodjosantoso, "Miskonsepsi materi larutan penyangga," *J. Inov. Pendidik. IPA*, vol. 4, no. 1, hlm. 41–51, 2018.

- [23] K. Bashooir dan S. Supahar, "Validitas dan reliabilitas instrumen asesmen kinerja literasi sains pelajaran fisika berbasis STEM," *J. Penelit. Dan Eval. Pendidik.*, vol. 22, no. 2, hlm. 219–230, 2018.
- [24] A. Septiani, "Penerapan Asesmen Kinerja dalam Pendekatan Stem (Sains Teknologi Engineering Matematika) untuk Mengungkap Keterampilan Proses Sains," 2016.
- [25] D. Dewati, D. Hadiarti, dan R. Fadhilah, "Pengembangan Instrumen Penilaian Tes Diagnostik Pilihan Ganda Dua Tingkat Untuk Mengukur Hasil Belajar Siswa Materi Hidrokarbon di SMA 10 Negeri Pontianak," *AR-RAZI J. Ilm.*, vol. 4, no. 2, Sep 2016, doi: 10.29406/arz.v4i2.671.
- [26] D. Susanti, S. Waskito, dan S. Surantoro, "Penyusunan instrumen tes diagnostik miskonsepsi fisika SMA Kelas XI pada materi usaha dan energi," *J. Pendidik. Fis.*, vol. 2, no. 2, 2014.
- [27] D. Natalia, J. Handhika, dan F. Huriawati, "Pengembangan Instrumen Tes Diagnosis Konsep IPA Fisika," *Momentum Phys. Educ. J.*, vol. 1, no. 2, hlm. 103, Okt 2017, doi: 10.21067/mpej.v1i2.1966.