DOI: 10.14421/impulse.2023.42-02

IMPULSE: Journal of Research and Innovation in Physics Education
Volume 4, Issue 2, 79 – 95

© ISSN (p): 2798–1762; ISSN (e): 2798-1754
http://ejournal.uin-suka.ac.id/tarbiyah/impulse

The Three-Tier Diagnostic Test Instrument Development to Identify the Students' Misconceptions about Motion Kinematic Material

Bagus Pratama[1*], Edi Istiyono[2]

[1] Department of Physics Education, Universitas Negeri Yogyakarta, Indonesia
*Corresponding author: bagusprt08@gmail.com

ABSTRACT

This research aims to: (1) determine the construction of the three-tier diagnostic test instrument to identify misconceptions. (2) Find out the quality of the three-tier diagnostic test instrument to identify misconceptions. (3) Know the profile of high school students' misconceptions. (4) Find out the level of practicality of the diagnostic test to identify misconceptions. This instrument development research adapted a method from the Oriondo and Antonio model. The researchers conducted this research at four high schools in Bantul Regency, using 343 students. Content validity was analyzed using Aiken's V formula. The applied data analysis technique is the polytomous item response theory approach according to the PCM. The test characteristics tested include item suitability, reliability, item difficulty level, item characteristic curve, information function, and SEM. Our research shows that (1) the physics research tool we made is made up of 15 questions that have valid and reliable test items. (2) Based on the criteria for the limit of the mean INFIT MNSQ, test items all fit the PCM model. The difficulty level of the test items is in the range between -0.561 and +0.330, which means the test items are in the good category. The total information function of the test is relatively high for abilities between -1.4 and +1.5. (3) The students' misconception profile obtained the highest misconceptions in the rectilinear motion sub-material in the medium category, and (4) The practicality of the instrument based on participants' responses stated that the instrument developed was included in the practical category.

INTISARI

Penelitian ini bertujuan untuk: (1) Mengetahui konstruksi instrumen three tier diagnostic test untuk mengidentifikasi miskonsepsi (2) Mengetahui kualitas instrumen three tier diagnostic test untuk mengidentifikasi miskonsepsi (3) Mengetahui profil miskonsepsi peserta didik SMA (4) Mengetahui tingkat kepraktisan instrumen untuk mengidentifikasi miskonsepsi. Jenis penelitian ini merupakan penelitian pengembangan instrumen menggunakan metode yang diadaptasi dari model Oriondo dan Antonio. Penelitian ini dilakukan di 4 SMA yang ada di Kabupaten Bantul dengan 343 peserta didik. Validitas isi dianalisis dengan formula V Aiken. Teknik analisis data yang digunakan yaitu pendekatan teori respo butir politomus menurut PCM. Karakteristik tes yang diuji antara lain kecocokan butir, reliabilitas, tingkat kesukaran butir, kurva karakteristik butir,

ARTICLE HISTORY
Received: June 11, 2024
Accepted: October 9, 2024

KEYWORDS:

Assessment instruments, Misconceptions, Three-tier diagnostic tests

KATA KUNCI:

Instrumen penilaian, Miskonsepsi, Three-tier diagnostic test

* Corresponding author:
Bagus Pratama, Universitas Negeri Yogyakarta, Indonesia
✉ bagusprt08@gmail.com

fungsi informasi dan SEM. Hasil penelitian menunjukkan bahwa: (1) Konstruksi intrumen penelitian fisika yang dikembangkan terdiri dari 15 butir soal memiliki butir tes yang valid dan reliabel, (2) Berdasarkan kriteria batas mean INFIT MNSQ, butir tes semuanya fit dengan model PCM. Tingkat kesukaran butir-butir tes berada pada rentang antara -0,561 sampai +0,330 yang berarti butir tes dalam kategori baik. Fungsi informasi total tes relatif tinggi untuk kemampuan antara -1,4 sampai +1,5, (3) Profil miskonsepsi peserta didik diperoleh miskonsepsi tertinggi pada submateri gerak lurus dengan kategori sedang, dan (4) Kepraktisan instrumen berdasarkan respon peserta didik menyatakan bahwa instrumen yang dikembangkan ini masuk dalam kategori praktis.

A. Introduction

Physics is a branch of science that studies how matter and energy function and interact. However, Grusche (2019) explains that physics is a field that utilizes mathematical models of scientific activity to help us understand everyday physical phenomena [1]. Understanding these concepts is essential for explaining daily physical phenomena and more complex scientific fields. Personal experiences or daily life can also foster students' thinking [2]. To design learning experiences that support the development of critical thinking and deeper conceptual understanding, educators must understand how ideas correlate with students' thinking.

Students' understanding of certain ideas can vary, and some may differ from experts' comprehension. One of the physics topics requiring conceptual understanding is the kinematics of motion. This concept is critical in physics education because it addresses how objects move without considering the reasons behind their motion. Understanding kinematics is fundamental for studying more complex physics disciplines. Indeed, middle schools have introduced this topic. However, many students still rely on misconceptions, making it difficult for them to grasp linear, projectile, and circular motion. Misconceptions are defined as situations where students have understandings that differ from expert knowledge [3]. The term describes inaccurate or incorrect understandings of concepts or information in a particular field or topic [4]. Often, students make errors in their understanding because the concepts they observe in daily life differ from those taught in the classroom [5], [6].

One of the biggest challenges encountered by high school students in learning physics is a lack of knowledge and difficulty understanding concepts. This issue can stem from incorrect understanding of physics material taught by their teachers or the influence of an inappropriate learning environment [7]. Students may fail to comprehend concepts taught at school if their learning environment is inadequate. Teacher errors, flawed ideas, and poor understanding can all contribute to students' misconceptions. Because they trust the knowledge their teachers impart, students often find it challenging to correct misconceptions caused by teacher errors [8]. Thus, teachers also play a role in students' misconceptions.

Having tools to identify misconceptions is essential. To uncover and address conceptual errors, concept maps, classroom discussions, practical experiments with Q&A sessions, multiple-choice tests, and written essays are applicable [9]. Addressing students' misconceptions promptly is crucial. One way to do this is by creating assessment tools to evaluate students' understanding of concepts, such as diagnostic tests [10]. Literature research shows that diagnostic tests are an effective method for identifying misconceptions.

A diagnostic test is a systematically and purposefully designed evaluation tool to identify students' strengths and weaknesses in a specific knowledge domain or skill. In education, diagnostic tests aim to assess students' strengths and weaknesses during the learning process ([11]. These tests facilitate educators gather information about students' skills or conditions. Diagnostic tests assist teachers in designing lessons tailored to students' unique needs, such as aligning subject matter with their level of understanding, adapting teaching methods to their learning styles, and providing additional support when needed. Diagnostic tests identify problems or difficulties [12]. These identified issues are then addressed through planned interventions. Therefore, diagnostic tests are essential in decision-making processes regarding individualized learning.

Examples of diagnostic tests include one-tier, two-tier, three-tier, and four-tier diagnostic tests. Two-tier diagnostic tests use a tiered approach to identify students' weaknesses and errors in understanding concepts. However, these tests remain limited and cannot distinguish misconceptions between students who score 1 and 2[13]. Moreover, two-tier diagnostic tests tend to limit the understanding of misconceptions to superficial levels. Without thoroughly grasping the tested concepts, participants might answer correctly based on the options provided. Two-tier tests are often insufficient for identifying complex patterns of misconceptions [14]. For example, participants may frequently give incorrect answers, but these tests lack the flexibility to detect such patterns.

One type of diagnostic test used to identify students' conceptual errors is the three-tier diagnostic test, which evolved from the two-tier diagnostic test. It comprises three levels of questions [9]. The first tier involves standard multiple-choice questions, the second tier presents reasoning options for the first-tier questions, and the third tier assesses students' confidence in their answers. The three-tier diagnostic test provides more accurate results for distinguishing misconceptions from a lack of knowledge [15]. This test enhances conceptual understanding by identifying the depth of misconceptions at each level.

Paper-based tests are traditional assessment methods involving printed test materials distributed to students, who then write their answers directly on the provided sheets. However, paper-based tests require significant effort and time for examiners to collect and manually review materials. Security and confidentiality can also be an issue due to risks of loss or physical damage to the papers [16].

Research on misconceptions shows that high school students may experience them. Therefore, we should conduct studies aimed at identifying misconceptions in kinematics. The title of this research is "Development of a Three-Tier Diagnostic Test to Identify Misconceptions in Kinematics."

B. Method

This quantitative-descriptive research employs the research and development (R&D) method to develop and examine products for educational use [17]. The study used a modified version of the Oriondo and Antonio model, consisting of three stages: (1) design, (2) testing, and (3) test measurement.

The design process included defining the test objectives, identifying the competencies to be assessed, selecting the subject matter, creating a test matrix, drafting blueprints, writing questions, developing scoring guidelines, validating instruments, and revising as needed. At this point, the test goals had to be set, and find out what high school students did not understand about the skills listed in the Kurikulum Merdeka (Independent Curriculum): Learning Outcomes (Capaian Pembelajaran/CP), Learning Progression (Alur Capaian Pembelajaran/ACP), and Learning Objectives Flow (Alur Tujuan Pembelajaran/ATP).

Kinematics of motion serves as the chosen physics topic for this research. The scoring guidelines had to be adapted to the applied question model. Istiyono (2022) explains the scoring guidelines for a three-tier multiple-choice test are as follows [13]:

Table 1. The Three Tier Multiple Choice Scoring

Answer	Reasons	Confidence Levels	Categories	Scores
Correct	Correct	Confident	Understanding the concept	7
Correct	Incorrect	Confident	Misconception	6
Incorrect	Correct	Confident	Misconception	5
Incorrect	Incorrect	Confident	Misconception	4
Correct	Correct	Not confident	Lucky guess	3
Correct	Incorrect	Not confident	Not understanding the concept	2
Incorrect	Correct	Not confident	Not understanding the concept	1
Incorrect	Incorrect	Not confident	Not understanding the concept	0

The trial phase involved determining the subjects, conducting the trial, and analyzing the results. The trial subjects were 11th-grade students from SMA Negeri 3 Bantul. The researchers conducted instrument testing to evaluate the characteristics of the test items. The next step was to look at the trial data to find out about the test instruments' features, such as the standard error of measurement (SEM), the level of item difficulty, and their characteristic curves.

The test measurement stage identified misconceptions. The researchers conducted measurements on high school students in Bantul Regency. The students selected for this study are 11th-grade students in Bantul Regency who have previously studied the kinematics of motion.

C. Result and Discussion

The Test Instrument Design

The purpose of the three-tier diagnostic test is to identify high school students' misconceptions more accurately and comprehensively. This research tool attempted to find out how excellent the 11th graders understand the Kurikulum Merdeka's Learning Outcomes (CP), Learning Progression (ACP), and Learning Objectives Flow (ATP).

The test material included physics topics covered in the Kurikulum Merdeka for the odd semester of 11th grade. The researchers developed the test instrument matrix using the competencies and cognitive elements determined in the previous steps. Once the test instrument matrix was complete, the next step involved creating a blueprint.

The test consisted of fifteen multiple-choice questions, structured in three tiers. The first tier contains questions with five answer choices, the second tier required students to provide reasons for their answers, and the third tier assessed students' confidence in their choices.

The subsequent step was instrument validation, carried out by providing validation sheets to expert validators (lecturers from the Physics Education Department at UNY) and practitioner validators (high school physics teachers). Table 2 presents the validity scores of the test items.

Table 2. The Question Item Validity Scores

Question Numbers	V-Value	Criteria
1	1.00	Valid
2	1.00	Valid
3	0.92	Valid
4	0.85	Valid
5	0.96	Valid
6	0.96	Valid
7	0.96	Valid
8	0.96	Valid
9	0.96	Valid
10	0.96	Valid
11	0.92	Valid
12	1.00	Valid
13	0.92	Valid
14	0.92	Valid
15	1.00	Valid

The Trial Test Result Data

The analysis on R-studio program obtained the output values such as assessment of the instrument reliability, observed on the figure.

```
> # Reliabilitas Butir Politomus
> coeff_alpha <- function(responses) {
+   # Get number of items (N) and individuals
+   n_items <- ncol(responses)
+   n_persons <- nrow(responses)
+   # Get individual total scores
+   x <- rowSums(responses)
+   # Get observed-score variance of whole test (X)
+   var_x <- var(x) * (n_persons - 1) / n_persons
+   # Get observed-score variance of each item (Y_j)
+   var_y <- numeric(n_items)
+   for (j in 1:n_items) {
+     var_y[j] <- var(responses[, j]) * (n_persons - 1) / n_persons
+   }
+   # Apply the alpha formula
+   alpha <- (n_items / (n_items - 1)) * (1 - sum(var_y) / var_x)
+
+   alpha
+ }
> coeff_alpha(data_sman_3_bantul_coba)
[1] 0.8613063
> |
```

Figure 1. Reliability Output

The figure indicates that the reliability of the developed instrument is 0.8613063. Efendi & Widodo (2019) explain the reliability value falls within the range of 0.81–0.90, which indicates a satisfactory level of instrument reliability [18].

The item-fit test using Infit Mean Square (MNSQ) is a critical tool in the development and validation of high-quality test instruments. The table presents the analysis results using RStudio software.

Table 3. Infit MNSQ

Item	INFIT MNSQ	Criteria
1	1.134	Fit with model
2	1.035	Fit with model
3	1.173	Fit with model
4	0.828	Fit with model
5	1.067	Fit with model
6	1.047	Fit with model
7	1.024	Fit with model
8	1.153	Fit with model
9	0.933	Fit with model
10	1.255	Fit with model
11	1.134	Fit with model
12	1.023	Fit with model
13	0.910	Fit with model
14	0.828	Fit with model
15	0.843	Fit with model

Table 3 shows the INFIT MNSQ scores are within the 0.828 – 1.255, indicating the fit-criterion of the items toward the model. This figure shows each item fit.

```
****************************************************

Summary outfit and infit statistic
     fit     M    SD
1 Outfit 1.067 0.196
2  Infit 1.026 0.134

****************************************************

Outfit and infit statistic
      item fitgroup Outfit Outfit_t Outfit_p Infit Infit_t Infit_p
1    item 1        1  1.196    1.216    0.224 1.134   1.106   0.269
2    item 2        2  1.212    1.017    0.309 1.035   0.296   0.767
3    item 3        3  1.227    1.734    0.083 1.173   1.361   0.174
4    item 4        4  0.801   -1.531    0.126 0.828  -1.453   0.146
5    item 5        5  1.243    1.324    0.186 1.067   0.538   0.591
6    item 6        6  0.990   -0.025    0.980 1.047   0.428   0.669
7    item 7        7  0.978   -0.084    0.933 1.026   0.243   0.808
8    item 8        8  1.224    1.570    0.117 1.153   1.308   0.191
9    item 9        9  0.872   -0.746    0.456 0.933  -0.577   0.564
10 item 10       10  1.303    2.117    0.034 1.255   2.066   0.039
11 item 11       11  1.196    1.216    0.224 1.134   1.106   0.269
12 item 12       12  1.263    1.743    0.081 1.023   0.226   0.821
13 item 13       13  0.944   -0.307    0.759 0.910  -0.658   0.511
14 item 14       14  0.801   -1.531    0.126 0.828  -1.453   0.146
15 item 15       15  0.755   -1.576    0.115 0.843  -1.308   0.191
>
```

Figure 2. Infit MNSQ

The difficulty levels of the items are based on the R-studio program as found in Figure 3.

```
Item Parameters -A*Xsi
      item   N     M xsi.item AXsi_.Cat1 AXsi_.Cat2 AXsi_.Cat3 AXsi_.Cat4
1   item 1 123 4.423   -0.284      1.477      1.501      0.529     -1.869
2   item 2 123 5.374   -0.561      0.952      0.059     -1.646     -1.855
3   item 3 123 3.024    0.055      0.403      1.413      2.215     -0.410
4   item 4 123 2.382    0.291     -0.036      0.287      3.002     -0.128
5   item 5 123 4.252   -0.058      2.170     -1.298     -0.425     -0.836
6   item 6 123 4.106   -0.231      0.655     -0.723     -1.370     -0.658
7   item 7 123 4.504   -0.299      1.777      0.472     -1.039     -0.584
8   item 8 123 3.138    0.136      0.976      0.044      0.199      0.012
9   item 9 123 2.350    0.330      0.552      1.143      3.288      0.340
10 item 10 123 3.423   -0.035      0.691      0.225     -0.371     -0.489
11 item 11 123 4.626   -0.371      1.526      0.429     -1.198     -2.054
12 item 12 123 3.675   -0.090      0.998      0.460      1.376     -1.076
13 item 13 123 4.081   -0.195      0.629      0.432      0.072     -0.682
14 item 14 123 3.577   -0.040      1.560      0.024      1.099     -0.850
15 item 15 123 3.724   -0.131      0.590      1.100      0.586     -0.597
   AXsi_.Cat5 AXsi_.Cat6 AXsi_.Cat7 B.Cat1.Dim1 B.Cat2.Dim1 B.Cat3.Dim1
```

Figure 3. The Question Item Difficulty Level Output

The figure shows the difficulty levels of the items are within the range of -0.561 and 0.330 based on the categorization of Qomariyah [19].

Table 4. The Item Difficulty Level Criteria

Item	Difficulty Levels	Categories
1	-0.284	Moderate
2	-0.561	Moderate
3	0.055	Moderate
4	0.291	Moderate
5	-0.058	Moderate
6	-0.231	Moderate
7	-0.299	Moderate
8	0.136	Moderate
9	0.330	Moderate
10	-0.035	Moderate
11	-0.371	Moderate
12	-0.090	Moderate
13	-0.195	Moderate
14	-0.040	Moderate
15	-0.131	Moderate

Figure 4 shows the characteristic curve of the easiest items.

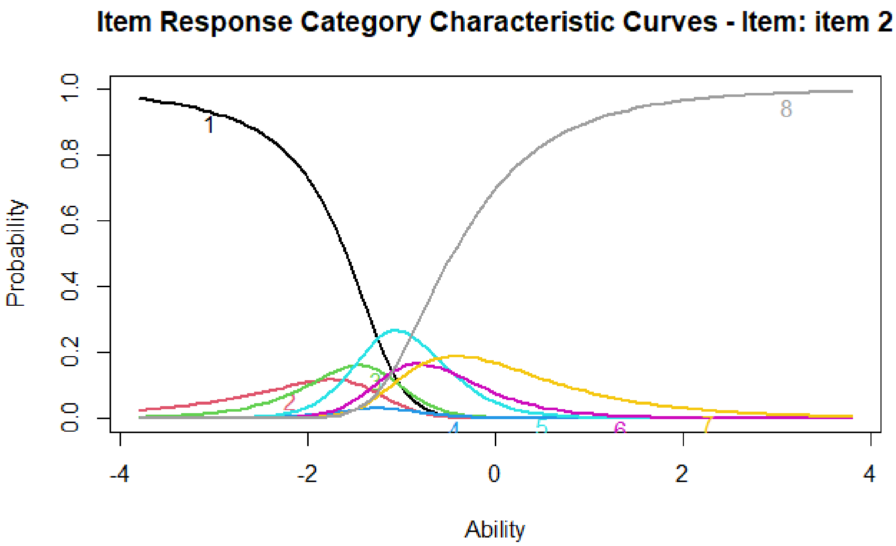

Figure 4. The Characteristic Curve of the Easiest Item Questions

Table 4 illustrates the types of tests that can differentiate the likelihood of responding correctly based on ability levels. The higher the students' ability score, the greater the likelihood of correctly answering higher-category questions. Conversely, the lower their ability score, the smaller the likelihood of correctly answering questions of higher value.

According to the ICC graph for Item 2, students with an ability level of -4 have a high probability of obtaining a score of 0 (Category 1). Students with an ability level of -1.8 are more likely to achieve a score of 1 (Category 2); those with an ability level of -1.5 are more likely to obtain a score of 2 (Category 3); students with an ability level of -1.2 are more likely to achieve a score of 3 (Category 4); and students with an ability level of -1 have a high likelihood of scoring 4 (Category 5). Similarly, students with an ability level of -0.8 have a high probability of obtaining a score of 5 (Category 6), while those with an ability level of -0.5 are likely to achieve a score of 6 (Category 7). Finally, students with an ability level of 4 have a high likelihood of obtaining a score of 7 (Category 8).

Figure 5 displays the information function and standard error of measurement (SEM) output from the RStudio program.

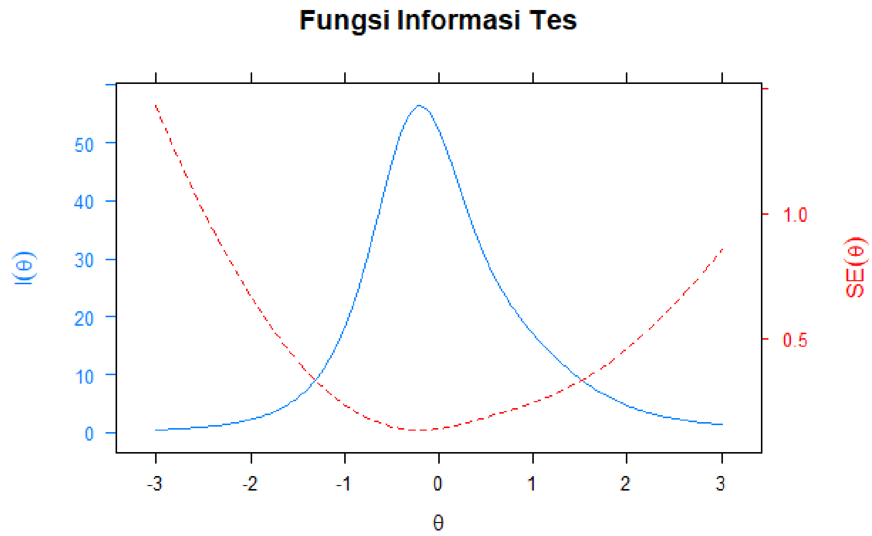

Figure 5. The Information Function and SEM

Based on the analysis results using the Partial Credit Model (PCM), the highest value on the information function curve was found to be 9, with an SEM value of 0.3. Additionally, the total information function curve shows that the first intersection between the information function curve and the SEM curve occurs at -1.4. This suggests that students with extremely low abilities can utilize the instrument.

The total information function curve also shows that the second point of intersection with the SEM curve is at 1.5. This means that the test questions are good for students and respondents with a range of abilities. This demonstrates that the test items can be effectively used for individuals across a wide spectrum of abilities.

The results indicate that this test instrument aligns well with the participants' abilities [20]. The reliability coefficient further confirms that the test instrument and its information function are highly reliable and stable

Measurement Results

The trial test found the developed instrument meet the requirement criteria such as validity, reliability, and fitness of each item toward the PCM. Figure 6 shows the physics misconception level percentage of students on each item.

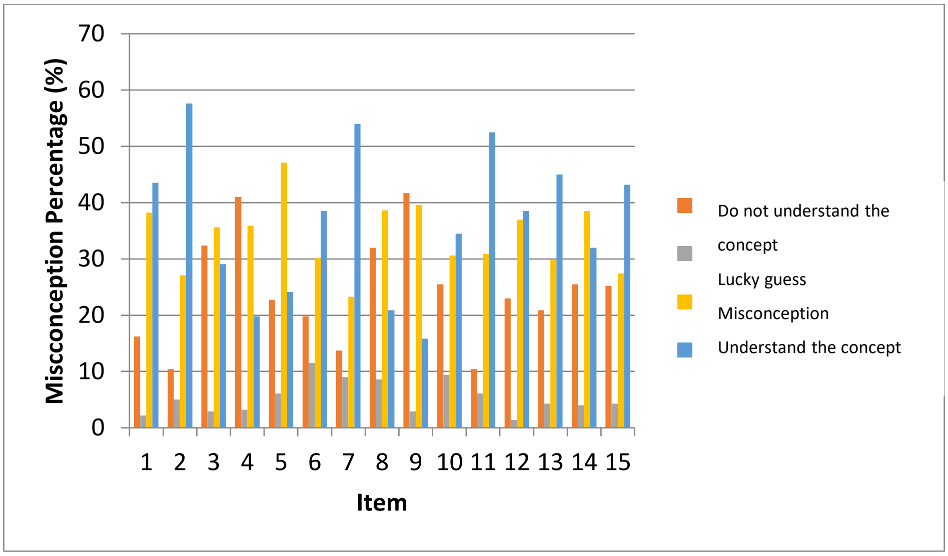

Figure 6. The Misconception Levels of Students on Each Item

Figure 7 shows the overall misconception level percentage.

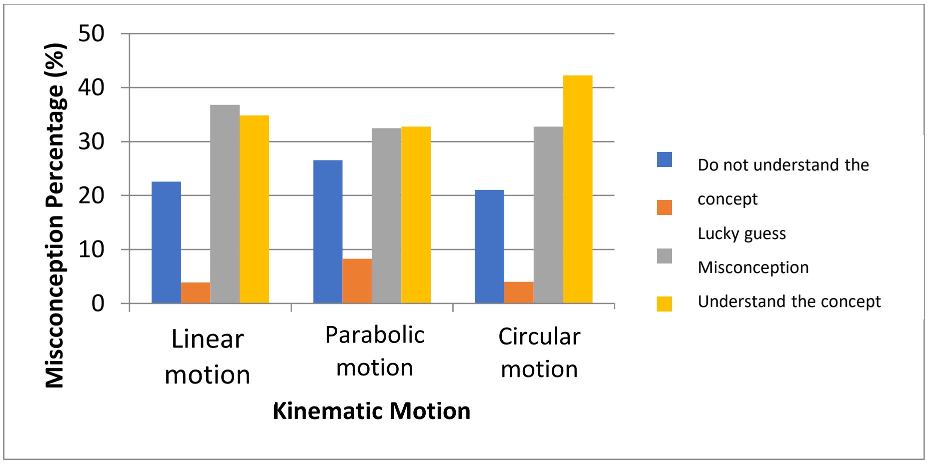

Figure 7. The Overal Misconception Percentage

Figure 7 presents the percentage of misconceptions among students in physics, specifically in the topic of motion kinematics. The results indicate that the highest percentage of misconceptions is found in the sub-topic of linear motion at 36.78%, followed by circular motion at 32.74% and parabolic motion at 32.46%.

Entino (2021) suggests that the data on the level of physics misconceptions among students from three schools in Bantul Regency were then put into levels of misconceptions [21]. Table 5 presents the identified levels of students' misconceptions.

Table 5. The Student Misconception Levels

Sub Materials	Misconception (%)			Overall (%)	Categories
	SMA N 1 Bantul	SMA N 1 Kasihan	SMA N 1 Sewon		
Linear motion	28,4	43,84	38,72	36,99	Moderate
Parabolic motion	23,6	41,46	33,68	32,91	Moderate
Circular motion	16,92	50,62	34	33,85	Moderate

The measurement results indicate that students experienced misconceptions at a moderate level. The grouping of misconceptions by Entino (2021) shows that students have moderate misconceptions about the subtopics of circular motion, parabolic motion, and linear motion [21].

A large-scale trial involving a total of 219 students was conducted across three schools: SMA Negeri 1 Kasihan, SMA Negeri 1 Sewon, and SMA Negeri 1 Bantul. This measurement phase aimed to determine students' levels of understanding during the product development process. Afterwards, the researchers analyzed and classified students' answers and reasoning into four categories of understanding: understanding the concept, misconception, lucky guess, and not understanding the concept.

To identify misconceptions in the topic of motion kinematics, the study employed a three-tier diagnostic test. The researchers chose this method because it offered higher quality and greater precision compared to previous research that used two-tier diagnostic tests.

The two-tier diagnostic test assessed students' understanding of kinematics concepts, whereas the three-tier diagnostic test also evaluated their confidence in the provided answers. This allowed for a more accurate identification of misconceptions, as students were required to explain their reasoning and indicate their confidence levels. Consequently, the three-tier diagnostic test provided more comprehensive and in-depth results regarding students' understanding of motion kinematics.

The misconceptions identified in each question item are: Firstly, 39% of students experienced misconceptions. When interpreting velocity-time graphs, students incorrectly believed that a horizontal line on the graph indicates constant acceleration. The correct concept is that a horizontal line signifies constant or unchanging velocity, meaning no acceleration occurs. Secondly, 28% of students experienced misconceptions regarding acceleration and deceleration in applying uniformly accelerated motion (GLBB). Students mistakenly believed that throwing a ball upward until it reaches its maximum height exemplifies accelerated and decelerated GLBB. However, the correct understanding is that accelerated and decelerated GLBB is observed when a ball is thrown vertically upward and allowed to fall back to the ground. Thirdly, 36% of students experienced misconceptions in calculating acceleration for GLBB. Students incorrectly assumed that the formula for calculating accelerated GLBB is $v_t^2 = v_o^2 - 2as$ leading to incorrect calculation. The correct concept should apply this formula $v_t^2 = v_o^2 + 2as$.

In the fourth question, 36% of students experienced misconceptions. When determining the solution for Aldi to arrive on time in City B, students incorrectly believed that the required speed should remain constant. They assumed that if the speed remained constant, Aldi would not arrive in City B on time. The correct concept is that Aldi's car needs to accelerate.

In the fifth question, 46% of students experienced misconceptions. Students incorrectly classified the speeds of the three trucks. They believed that the truck's speed was inversely proportional to its acceleration, meaning that higher acceleration corresponded to lower speed. The correct concept is the relationship between speed and acceleration in linear motion.

In the sixth question, 31% of students experienced misconceptions. Students incorrectly understood the variables influencing the maximum horizontal position of an object. They believed that the angle of elevation and time variables affected the maximum horizontal position in parabolic motion. The correct idea is that the angle of elevation and the object's initial speed are the factors that affect its largest horizontal position in parabolic motion. The formula for the largest range in parabolic motion should be $X_{MAX} = \frac{v_o^2 sin2a}{g}$.

In the seventh question, 23% of students experienced misconceptions. They incorrectly interpreted the maximum range of parabolic motion, believing that speed is inversely proportional to the maximum range—meaning higher speed results in a shorter range. The correct concept is that speed is directly proportional to the maximum range.

In the eighth question, 39% of students experienced misconceptions. When applying parabolic motion to determine the farthest point an object reaches, students incorrectly assumed that the larger the angle of elevation, the greater the farthest point. The correct concept is that the farthest point occurs when the angle of elevation reaches 45°. When the angle exceeds 45°, the range begins to decrease again.

In the ninth question, 40% of students experienced misconceptions. Students misunderstood how to determine the position of a parabolic motion object on the x and y axes. They incorrectly assumed that the formula for the position of the object on the x-axis was $x = v_o sin\theta t - \frac{1}{2}gt^2$ and the position for two objects in y-axis was $y = v_o cos\theta t$. This misconception led to incorrect calculation. The correct concept for an object in x-axis is $x = v_o cos\theta t$ while for an object in y-axis is $y = v_o sin\theta t - \frac{1}{2}gt^2$.

In the tenth question, 32% of students experienced misconceptions. They misunderstood the relationship between speed and the angle of elevation with the maximum height in parabolic motion. They incorrectly believed that speed and the angle of elevation were inversely proportional to the maximum height, meaning that the greater the initial speed and the angle of elevation, the lower the height. However, the correct concept is - the initial speed and the angle of elevation are directly proportional to the maximum height—so the greater the speed and angle, the higher the parabola.

In the eleventh question, 32% of students experienced misconceptions. They failed to correctly categorize the linear speed in circular motion. Students incorrectly assumed that for concentric wheels, the linear speed is the same. The relationship between wheels A and C, in contact, was misunderstood. Students believed that since wheel C has a larger radius, $V_A < V_C$. However, the correct concept is that for wheels in contact, their linear speeds are the same regardless of the radius difference. For the relationship between wheels A and B, concentric, with radius of A than B, the speed is V_B<V_A.

In the twelfth item, 39% of students experienced misconceptions. When calculating the period of circular motion, students encountered difficulties. According to their understanding, the period is the amount of time required to complete one rotation, leading to incorrect calculations. The correct concept of the period is the number of oscillations or rotations per unit of time.

A total of 31% of students incorrectly understood the thirteenth question. They incorrectly analyzed the differences in angular velocity, frequency, and time when a bicycle moved in a linear path without accelerating. They believed that when the bicycle moved linear without increasing its speed, its linear speed would increase. However, the correct concept is that when the bicycle moved on a flat surface without accelerating, its linear speed remains constant.

In the fourteenth question, 39% of students experienced misconceptions. When dealing with different speed, students misunderstood the analysis of frequency changes in circular motion. Most believed that frequency was inversely proportional to speed, meaning that the higher the frequency, the slower the speed. The correct concept is that frequency is directly proportional to speed, which means that the higher the frequency, the greater the speed.

In the fifteenth question, 29% of students experienced misconceptions. While analyzing the comparison of periods in circular motion, students misunderstood the relationship. They believed that time was inversely proportional to itself, implying that the longer the time required for one rotation, the shorter it becomes. In reality, time is directly proportional to itself, meaning that the longer the time for one rotation, the greater it is.

Subtopics on linear motion reached 37%, parabolic motion 33%, and circular motion 34%, indicating the overall percentage of conceptual errors or misconceptions among students. The results show that the test-takers exhibited a moderate level of misconceptions. The subtopics of linear motion, parabolic motion, and circular motion are considered to have a moderate level of misconceptions among students [21].

The researchers used the questionnaire results collected during the measurement phase to determine the students' reactions to the developed tool. The questionnaire consisted of ten questions addressing the readability of the instrument and its functionality. Students completed the questionnaire after finishing the three-tier diagnostic test. A total of 54 students participated in the survey, and the percentage responses for each question were as follows.

Table 6. The Students' Respons

Item	Respons	Percentage (%)
1	The applied sentences in the questions are readable	75
2	The Indonesian language of the items are excellent and correct.	78
3	The guidelines of the questions are readable.	71
4	The physics phenomena in the given items are understandable.	70
5	The given time is adequate	78
6	The sizes and positions of the figures are readable.	76
7	The graphics to present the items are understandable.	75
8	The products of the three-tier multiple choice test facilitate students to reflect about the physics conceptual understanding	78
9	The implementation of Microsoft Form is suitable for the Three-Tier Diagnostic Multiple Choice Test	73
10	The product motivates me to improve my conceptual understanding	76

The scores from the ten questionnaire items were then averaged, resulting in an overall percentage score of 75%. Nufus & Sakti (2021) explain this score falls within the range of 61% - 80% [22]. Thus, the three-tier diagnostic test instrument developed is practical for identifying physics misconceptions among high school students.

D. Conclusion

The research has led to several conclusions. The designed the physics assessment instrument to measure high school students' knowledge of motion kinematics. The instrument consists of a three-tier multiple-choice format. This physics assessment instrument is highly effective for evaluating high school students' conceptual understanding. The analysis results include validity, reliability, goodness of fit, difficulty level, characteristics, information function, and standard error measurement. The profile of misconceptions among 11th-grade physics students revealed a misconception rate of 37% in the kinematics subtopic, 33% in straight motion, and 33% in parabolic motion. Overall, the level of student misconceptions falls into the moderate category. The assessment tool is practical for identifying students' misconceptions. Student responses regarding the practicality of the instrument showed an mean score of 75%.

References

[1] S. Grusche, "Phenomenon-based learning and model-based teaching: Do they match?," in *Journal of Physics: Conference Series*, 2019, vol. 1287, no. 1. doi: 10.1088/1742-6596/1287/1/012066.

[2] A. Fadllan, W. Y. Prawira, Arsini, and Hartono, "Analysis of students' misconceptions on mechanics using three-tier diagnostic test and clinical interview," in *Journal of Physics: Conference Series*, 2019, vol. 1170, no. 1. doi: 10.1088/1742-6596/1170/1/012027.

[3] N. Mukhlisa, "Miskonsepsi Pada Peserta Didik," *SPEED J. J. Spec. Educ.*, vol. 4, no. 2, pp. 66–76, 2021, doi: 10.31537/speed.v4i2.403.

[4] R. Annisa, B. Astuti, and B. N. Mindyarto, "Tes Diagnostik Four Tier untuk identifikasi pemahaman dan miskonsepsi siswa pada materi gerak melingkar beraturan," *J. Pendidik. Fis. dan Keilmuan*, vol. 5, no. 1, 2019, doi: 10.25273/jpfk.v5i1.3546.

[5] S. Anggrayni and F. U. Ermawati, "The validity of Four-Tier's misconception diagnostic test for Work and Energy concepts," in *Journal of Physics: Conference Series*, 2019, vol. 1171, no. 1. doi: 10.1088/1742-6596/1171/1/012037.

[6] N. J. Fratiwi *et al.*, "Overcoming Senior High School Students' Misconceptions on Newton's Laws: A DSLM with Inquiry Learning based Computer Simulations," in *Journal of Physics: Conference Series*, 2019, vol. 1204, no. 1. doi: 10.1088/1742-6596/1204/1/012023.

[7] W. Wartono, M. N. Hudha, and J. R. Batlolona, "Guided inquiry and PSR in overcoming students' misconception on the context of temperature and heat," in *AIP Conference Proceedings*, 2018, vol. 2014. doi: 10.1063/1.5054433.

[8] A. Busyairi and M. Zuhdi, "Profil Miskonsepsi Mahasiswa Calon Guru Fisika Ditinjau Dari Berbagai Representasi Pada Materi Gerak Lurus Dan Gerak Parabola," *J. Pendidik. Fis. dan Teknol.*, vol. 6, no. 1, pp. 90–98, 2020, doi: 10.29303/jpft.v6i1.1683.

[9] A. M. Rampengan, K. G. Kumonong, and A. T. Rondonuwu, "Analisis Kesalahan Konsep Mahasiswa Calon Guru Fisika Pada Materi Kinematika Menggunakan Tes Diagnostik Three Tier," *Charm Sains J. Pendidik. Fis.*, vol. 2, no. 2, pp. 94–98, 2021, doi: 10.53682/charmsains.v2i2.113.

[10] H. Haerunnisa, P. Prasetyaningsih, and L. T. Biru, "Analisis Miskonsepsi Siswa SMP pada Konsep Getaran dan Gelombang," *PENDIPA J. Sci. Educ.*, vol. 6, no. 2, 2022, doi: 10.33369/pendipa.6.2.428-433.

[11] N. A. K. Asnawi, Tengku Muhammad Sahudra, Dini Ramadhani, Ary Kiswanto Kenedi, Muhammad Rizki Wardana, *Gaya Belajar Siswa Sekolah Dasar dan Tes Diagnostik*. Yogyakarta: Penerbit Deepubliser Digital, 2023.

[12] A. S. A. Wahyuni, A. Wartawijaya, and M. Hasyim, "Pelatihan Membuat Instrumen Tes Diagnostik Four Tier Test pada Guru SMAN 11 Makassar," *PENGABDI*, vol. 3, no. 2, 2022, doi: 10.26858/pengabdi.v3i2.40870.

[13] E. Istiyono, "Diagnostic Tests as an Important Pillar in Today's Physics Learning: Four-tier Diagnostic Test a Comprehensive Diagnostic Test Solution," in *Journal of Physics: Conference Series*, 2022, vol. 2392, no. 1. doi: 10.1088/1742-6596/2392/1/012001.

[14] M. A. Farrosi, W. Siswaningsih, N. Nahadi, and T. Rahmawati, "Profil Miskonsepsi Siswa SMA pada Materi Kesetimbangan Kimia Menggunakan Tes Diagnostik Pilihan Ganda Tiga Tingkat," *J. Ris. dan Prakt. Pendidik. Kim.*, vol. 10, no. 2, 2022, doi: 10.17509/jrppk.v10i2.52243.

[15] E. A. Soru, H. F. Rares, and J. Caroles, "Penggunaan Three-Tier Diagnostic Test Untuk Pemetaan Tingkat Penguasaan Elektrokimia Mahasiswa Kimia Semester II," *Oxyg. J. Chem. Educ.*, vol. 2, no. 1, 2020, doi: 10.37033/ojce.v2i1.132.

[16] R. Rachmawati and A. Kurniawati, "Pengembangan Instrumen Penilaian Tes Berbasis Mobile Online pada Prodi Pendidikan Matematika," *Prima J. Pendidik. Mat.*, vol. 4, no. 1, 2020, doi: 10.31000/prima.v4i1.1891.
[17] A. Maydiantoro, "Model-Model Penelitian Pengembangan (Research and Development)," *J. Metod. Penelit.*, no. 10, 2019.
[18] Y. Efendi and A. Widodo, "Uji Validitas Dan Reliabilitas Instrumen," *J. Kesehat. Olahraga*, vol. 7, no. 2, 2019.
[19] L. Qomariyah, "Analisis Tingkat Kesukaran dan Daya Pembeda Butir Soal TOAFL Universitas Hasyim Asy'ari Tebuireng Jombang," *Lisanan Arab. J. Pendidik. Bhs. Arab*, vol. 6, no. 1, 2022, doi: 10.32699/liar.v6i1.2549.
[20] A. Z. Zain, T. R. Ramalis, and M. Muslim, "Karakterisasi Instrumen Tes Keterampilan Bepikir Kreatif Berdasarkan Analisis Partial Credit Model," *J. Ilm. Pendidik. Fis.*, vol. 6, no. 1, 2022, doi: 10.20527/jipf.v6i1.4806.
[21] R. Entino, E. Hariyono, and N. A. Lestari, "Analisis Miskonsepsi Peserta Didik Sekolah Menengah Atas pada materi Fisika," *PENDIPA J. Sci. Educ.*, vol. 6, no. 1, 2021, doi: 10.33369/pendipa.6.1.177-182.
[22] V. F. Nufus and N. C. Sakti, "Pengembangan Lembar Kerja Peserta Didik Elektronik Berbasis Flipbook Pada Mata Pelajaran Ekonomi Kelas XI," *J. PTK dan Pendidik.*, vol. 7, no. 1, 2021, doi: 10.18592/ptk.v7i1.4633.