Journal for the Education of Gifted Young Scientists, 8(4), 1439-1456, Dec 2020
e-ISSN: 2149- 360X
jegys.org

Young Wise Publishing
youngwisepub.com
© 2020

Research Article

The development of diagnostic test instrument for mathematical representation ability (PhysDTRA) in high school physics learning

Awal Mulia Rejeki Tumanggor[1]*, Supahar[2]

Department of Physics Education, Yogyakarta State University, Indonesia

Article Info

Received: 22 August 2020
Revised: 02 October 2020
Accepted: 03 December 2020
Available online: 15 Dec 2020

Keywords:
Development of diagnostic test
Diagnostic instrument
Mathematical representation
Validity and reliability
Work and energy

Abstract

Assessment in education is part of collecting and processing various information related to student achievements during learning. Good learning quality can be seen from the quality of the assessment. Assessment activities that help educators to find out the difficulty of students on a learning material so that it can be guided to achieve completeness criteria, namely by diagnostic tests. The assessment carried out must follow 21st-century learning that is integrated with the industrial era 4.0, namely how to diagnose difficulties related to the representation ability of students in physics learning by utilizing assessments that have advantages in detecting student difficulties and giving suggestions for appropriate improvement. The purpose of this research was to develop a test instrument (*PhysDTRA*) that could be used to diagnose students' mathematical representation abilities in high school physics learning. The results of the research were analyzed quantitatively and qualitatively using Item Response Theory (IRT). Based on the content validity, the *PhysDTRA* instrument was declared to be valid according to the expert judgments who were analyzed using the Aiken's V equation. All items in the test instrument were valid based on the Rasch, INFIT MNSQ, and INFIT *t* models. The *PhysDTRA* instrument has also been relied upon based on the reliability of the estimated items and TIC curves so that it can be used to diagnose and determine the profile of students' mathematical representation abilities. Thus, the *PhysDTRA* instrument developed has fulfilled the test characteristics that are feasible of its content, empirical evidence, validity, and reliability.

To cite this article:
Tumanggor, A.M.R., Supahar. (2020). The Development of Diagnostic Test Instrument for Mathematical Representation Ability (PhysDTRA) in High School Physics Learning. *Journal for the Education of Gifted Young Scientists*, 8(4), 1439-1456. DOI: http://dx.doi.org/10.17478/jegys.777425

Introduction

Education carried out through the learning process will increase student competency if held interactively, fun, and motivating. Education requires feasible planning in learning, implementing the learning process, evaluating the learning process, and monitoring the learning process. The best learning process will form intellectual ability, think critically, and stimulate creativity and behavior or personality changes based on specific practices or experiences. Everything aims to improve the good quality, efficient, and effective education. This is supported by Minister of Education and Culture Regulation No. 22 of 2016, which explains that each learning process aims to achieve graduate competencies, which means the graduate competency standards provide a conceptual framework regarding the learning objectives that must be achieved including the aspects of student competencies.

The low quality of education has become a big problem in Indonesia, therefore it is necessary to have a decent education that is of good quality and can compete internationally (Tumanggor et al. 2020). Learning activities are not just ordinary interactions with students, but through procedures that have been prepared from various aspects and planned according to the basis and rules of learning (Liang et al. 2012; Wilcox & Pollock, 2015). Thus, students will get a good education and references from various types of questions or materials that have been prepared by the teacher as well.

[1] Department of Physics Education, Graduate Program, Yogyakarta State University, Indonesia. E-mail: awalmrt94@gmail.com, Orcid No: 0000-0002-2365-1581
[2] Department of Physics Education, Graduate Program, Yogyakarta State University, Indonesia. E-mail: supahar@uny.ac.id, Orcid no: 0000-0002-2486-5549

Learning the Industrial Revolution 4.0 can assemble the latest approaches to learning physics in combining the three important things that are pedagogic, applied science (*Technology*), as well as synergistic physics content in the utilization of information known as TPACK. The use of technology has a deep influence on students because technology is an effective intellectual tool for example is learning that is supported by the web or online that is easily accessible and accepted by students (Sadaghiani, 2011; Yeh et al. 2016).

Learning success is measured by assessment activities. Assessment can be used as a sign or evidence of an ongoing process so that it is comprehensive (Atkin & Coffey, 2003; Darmawan et al. 2020). Assessment is closely related to evaluating the ability of students to achieve learning objectives. Assessment instruments are used to measure the abilities of students, one of the assessments is to use tests (Loewenthal & Lewis, 2018). Assessment instruments such as test instruments must be developed with exact planning as a strategic step to solving physics problems (Adams et al., 2015; Gurcay & Gulbas, 2015; Kirschner et al. 2016). One way educators measure how students' level of understanding of a subject is by providing a diagnostic assessment (Gurel et al. 2015; Tumanggor et al. 2020; Volfson et al. 2018). The application of diagnostic assessments in learning will be the teacher's reference to ensure teacher achievement and progress in explaining the material that has been taught (Nitko & Brookhart, 2011; Pujayanto et al. 2018). Facts on the field show that there are still many teachers who do not understand the diagnostic assessment instrument model (Pujayanto et al. 2018). Based on observations at several high schools in the Special Region of Yogyakarta, diagnostic assessments are seldom carried out in the learning process because standardized test instruments to measure ability and detect student difficulties are not available. There is also an assessment only limited to ordinary tests without any improvement process.

The important ability in learning in the 21st century as a step in solving physics problems is the ability to use a representation format (Ibrahim & Rebello, 2012). Format representation is a way that can explain information during the problem-solving process (Docktor & Mestre, 2014; Krawec, 2014; Tumanggor et al., 2019). One format of the representation that can be applied is a mathematical representation (equation) in the concept of physics (De Cock, 2012). Mathematics is considered as the language of physics, many physics concepts are expressed in mathematical equations (Pospiech et al. 2016). Physics learning needs to include the conversion of physics modeling into mathematical modeling (for example, functional relations) and interpretation of mathematical models for physics (Treagust et al. 2017; Uhden et al. 2012). Mathematical representation is generally used by students in solving physics problems. Therefore mathematics and physics have a close relationship in learning.

The results of the PISA 2018 evaluation (*Programme for International Student Assessment*) explain that the Indonesian state has decreased the level of student ability for three aspects measured when compared to the PISA 2015 evaluation, and one of three aspects is the mathematical ability (OECD, 2019). This evaluation is in line with the empirical experience in the field, students' ability to solve problems mathematically is still not good enough in applying mathematical equations to physics problems. Students have obstacles in understanding and converting mathematical language to physics, barriers to understanding the value of mathematical representations, and barriers combining the relationship between mathematical representations to physics concepts (Zhe, 2012). This is similar to research conducted by Adlina & Supahar (2019) which states that students have difficulty understanding mathematical symbols and equations in physics. One of the important things in solving physics problems is the ability of students to combine mathematical symbols and structures with their intuition and numeracy knowledge (Bing & Redish, 2014; Bollen et al. 2017; Ceuppens et al. 2018; Docktor et al. 2016). Therefore, it is necessary to develop an assessment in the form of a test instrument to diagnose the mathematical representation ability of students so that appropriate treatment can be given to improve it. The operational framework for developing a diagnostic test instrument for mathematical representation ability can be seen in Table 1

Table 1.

Synthesis Results for The Development of Test Instruments

Theoretical Framework	Conceptual Framework	Operational Framework
Mathematical representation has an important influence on learning physics. The structure of physics uses mathematical models to explain the relationship between variables systematically (Niss, 2016; Treagust et al. 2017). The mathematical representation ability is often used in solving physics problems and is represented in different ways-verbal, sketch, diagram, graph, and equation (Heuvelen & Zou, 2001; Minarni & Napitupulu, 2017). Mathematical representation plays a role in improving students' understanding of concepts and solving abstract problems (Kriek & Koontse, 2017; Pape & Tchoshanov, 2001; Park & Choi, 2013).	Albe et al. (2014) explained that indicators of mathematical representation emphasize the ability of students to determine, compare, and use equations to calculate and solve problems. Bing & Redish (2014) revealed the mathematical representation ability was also measured based on: (1) students' ability to use and determine variables, (2) then combine symbols and variables to form equations, and (3) solve problems using mathematical operations. Angell, Kind, Henriksen, & Guttersrud (2008) stated that there are several indicators in measuring the mathematical representation ability related to physics learning namely; how students connect between variables and determine the right equation, and carry out mathematical operations to develop or simplify equations.	Mathematical representation is a format that describes equations or numbers related to problems. Based on the results of the synthesis and adapted to the Work and Energy material, the test instruments developed in this research using the following indicators: ✓ Interpret variables according to concepts in the form of pictures, tables, diagrams, and graphs to solve problems. ✓ Linking the variables contained in a problem in the form of pictures, tables, diagrams, and graphs to solve problems. ✓ Operate the equation correctly in the form of numbers, symbols, graphics, or images in solving problems. ✓ Conclude the conditions by operating mathematical equations to obtain results.

Problem of Study

The low level of mathematical representation ability in Indonesia based on the results of the 2018 PISA study is the focus of the Indonesian Government to evaluate and improve the quality of education. One of the things that are being reviewed by the Government is the improvement of the assessment system. Diagnostic assessment is an appropriate assessment to detect student weaknesses during the learning process. Besides, based on a summary of interviews and surveys with physics teachers in Yogyakarta, it is stated that almost all teachers have never heard of a diagnostic assessment or used it as an assessment in class. Therefore, this research is focused on developing a diagnostic assessment instrument called *PhysDTRA* (*Physics Diagnostic Test for Mathematical Representation Ability*). Based on the explanation of this information,

- How is the feasibility of the *PhysDTRA* instrument that was developed to diagnose the mathematical representation ability of high school students?
- The feasible instrument developed must meet the valid and reliable characteristics, and be able to determine the profile of student's difficulties in learning physics.

Method

Research Method

This research is research development. The research method used is a combination of the 4-D model and the Oriondo & Dallo-Antonio development model. The stages of development are carried out with a 4-D development model consisting of 4 stages according to Thiagarajan (1974) namely Define, Design, Develop, and Disseminate. Oriondo & Dallo-Antonio (1984) development model which includes 1) Planning the Test, 2) Trying out, and 3) Establishing Test Validity, 4) Establishing Test Reliability, and 5) Interpreting the Test Score. This development procedure produces the *PhysDTRA* instrument shown in Figure 1.

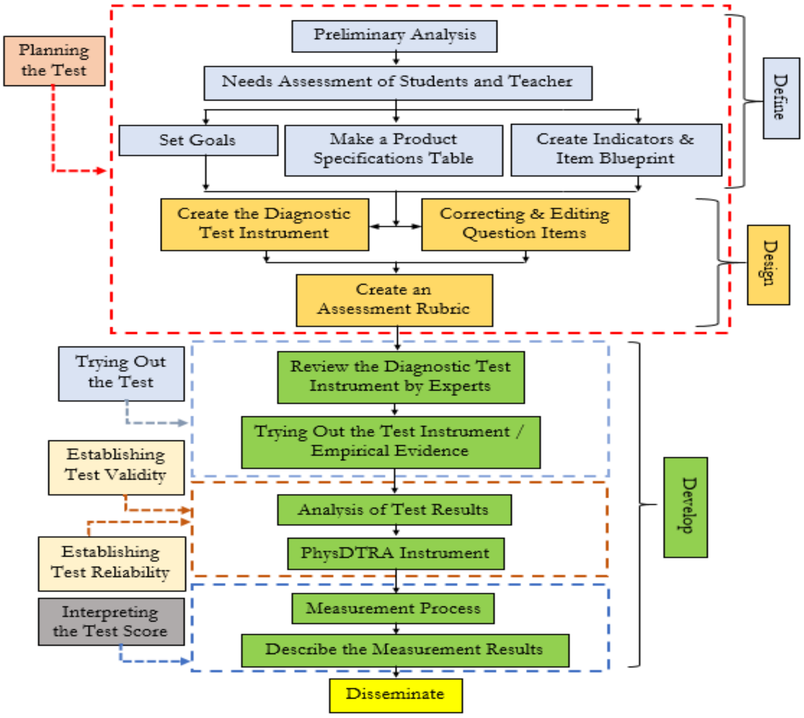


Figure 1.
Test Instrument Development Procedure

The main product of this research is a test instrument that can be used to diagnose mathematical representation ability in high school physics lessons, especially Work and Energy. Data analysis techniques using qualitative and quantitative descriptive analysis. The qualitative analysis aims to see the construction of instruments through expert judgment. Quantitative analysis is used to determine the validity and reliability of the instrument. The draft instrument was created and developed into questions, and the test instrument was given to experts for content validity. The test instrument consisted of 36 multiple choice questions with cognitive levels ranging from C3 (applying) until C4 (analysis). Then, 36 questions were divided into two coded A and B package questions, where each packet was prepared by considering the representation of each indicator measured from the mathematical representation ability. Each package consists of 20 multiple choice questions, including 4 questions as *anchor items.*

Participants

The test instrument was put into a trying out stage in March 2020 involving 296 students of class XI MIPA (*Mathematics and Natural Sciences*) from three high schools located in the Special Region of Yogyakarta, especially Bantul Regency. The three schools are SMA Negeri 1 Banguntapan consisting of four classes, SMA Negeri 1 Sewon consisting of four classes, and SMA Negeri 3 Bantul consisting of three classes.

Data Collection

Mathematical Representation Ability Test

Instruments that have been developed by indicators of mathematical representation ability and item descriptions (Appendix 1) will be validated (content validity) through expert judgment using the instrument item review criteria (Appendix 2) (Pujayanto et al. 2018). The content validity of the test instrument is obtained by providing the test instrument developed to the experts for review. The assessment was conducted by seven raters, namely assessment experts, physics learning experts, physicists, two practitioners evaluating physics education, and two peer reviewers. The instrument readability test was also carried out by giving the instrument to ten students. The results of the content validation test and readability test are used to improve the instrument. The data obtained in this research are the content validation results through expert judgment on the developed test instruments. Then, the validity for empirical evidence is carried out at the trying out stage using the developed test instrument to obtain item *fit* and item reliability results.

Data Analysis

The content validity assessment of the test instruments was analyzed using Aiken's V formula. The coefficient of content validity was based on the judgment of the experts as much as the *n* raters could represent the construct of the measured item. Aiken's V index value is formulated in equation 1.

$$V = \frac{\sum S}{[n(c-1)]} = \frac{\sum[r-lo]}{[n(c-1)]} \quad (1)$$

Description, n is the number of raters, c is the highest validity rating, lo is the lowest validity rating, r is the number given by the rater (Aiken, 1985; Azwar, 2012). Based on equation 1, Azwar (2015) explained that the validation criteria value of the diagnostic test instrument for mathematical representation ability was divided into five criteria as shown in Table 2.

Table 2.
Criteria Value of Aiken's V Validity

Validity Value	Category
$0.8 \leq V \leq 1.0$	Very Good
$0.6 \leq V \leq 0.8$	Good
$0.4 \leq V \leq 0.6$	Acceptable
$0.2 \leq V \leq 0.4$	Bad
$V \leq 0.2$	Very Bad

Validity to show empirical evidence is obtained through analysis of item responses to test results in the form of dichotomous data. Dichotomous data were analyzed using Item Response Theory (IRT) according to the Rasch model or Partial Credit Model (PCM) 1 Parameter Logistics (1-PL). Analysis using the Quest and Parscale programs. The Quest program is used to determine the *goodness of fit*, reliability, and item difficulty index. The Parscale program is used to designate information functions and standard error measurement (SEM) (Supahar & Prasetyo, 2015). Adams & Khoo (1996) explained that items that fit or were categorized fit with the PCM model if the mean and standard deviation of *INFIT MNSQ* were between 0.77 until 1.30 and the *INFIT t* value was between -2.0 until 2.0.

The reliability of the item estimate on the *PhysDTRA* instrument can be determined using the Quest program. The reliability results are known from the output data with the extension *sh* (.sh) in the Summary of Item Estimates section. Subali & Suyata (2011) states that the higher the reliability coefficient the more reliable the instrument and the smaller the possibility of errors. George & Mallery (2020) categorizes the reliability coefficients shown in Table 3.

Table 3.
Interpretation of Reliability Coefficient

Reliability Coefficient	Interpretation
$\alpha \geq 0.9$	Excellent
$0.9 > \alpha \geq 0.8$	Good
$0.8 > \alpha \geq 0.7$	Acceptable
$0.7 > \alpha \geq 0.6$	Questionable
$0.6 > \alpha \geq 0.5$	Poor
$\alpha < 0.5$	Unacceptable

The reliability testing of each item is determined using the Item Information Curve (IIC) and the Total Information Curve (TIC). IIC and TIC are obtained by processing data through the Parscale program.

The category of measurement results of mathematical representation ability (θ) is interpreted in five scales according to Azwar (2012) which are very low, low, medium, high, and very high. The distribution category of mathematical representation capabilities is shown in Table 4.

Table 4.
Category of Mathematical Representation Ability

Ability Range	Category
$\theta \leq -1.5\sigma$	Very Low
$-1.5\sigma < \theta \leq -0.5\sigma$	Low
$-0.5\sigma < \theta \leq +0.5\sigma$	Medium
$+0.5\sigma < \theta \leq +1.5\sigma$	High
$-1.5\sigma < \theta$	Very High

Description, σ is $\frac{1}{6}$(maximum ideal scale – minimum ideal scale). The minimum ideal ability scale is -4 and the maximum ideal ability scale is +4. This is based on a logit scale of -4 and +4. Ability distribution can be known from the output file *PH3* and *SCO* format in the Parscale program.

Results

A theoretical study was arranged to find out the indicators used in the mathematical representation ability. The development of this test instrument was adjusted to the subject matter of The Work and Energy and then divided into sub-materials namely the work, potential and kinetic energy, and the conservation law of mechanical energy. This research has synthesized indicators of mathematical representation ability of several experts, and the results of synthesis are shown in Table 5. Indicators of the items in each sub-material are adjusted with indicators of mathematical representation, some of which are; interpret the work done by a constant force, linking variables between the work and a varying force, operate the work equation on the block-rope system appropriately, conclude conditions based on the concept of the work in the inclined plane correctly, conclude conditions based on the energy concept and obtain the results of block displacement from mathematical operations using changes in mechanical energy.

Table 5.

Mathematical Representation Ability Synthesis Results

Aspect	Sub-Aspect	Indicator
Interpret Data	Interpret Variables	Students can interpret variables according to concepts in the form of pictures, tables, diagrams, and graphs to solve problems.
	Linking Variables	Students can connect the variables contained in a problem in the form of pictures, tables, diagrams, and graphs to solve problems.
Operationalize the Equation	Operate Data	Students can operate equations appropriately in the form of numbers, symbols, graphs, or images in solving problems.
	Formulate Results	Students can conclude conditions by operating mathematical equations to obtain results.

The content validity of the developed test instruments was reviewed by seven raters with four rating scales before testing the instrument. Based on the standard set by Aiken, the minimum standard of Aiken's V coefficient for this study is 0.76 with a probability of 0.045 (Aiken, 1985). The content validity was analyzed quantitatively and the results of the analysis are shown in Figure 2.

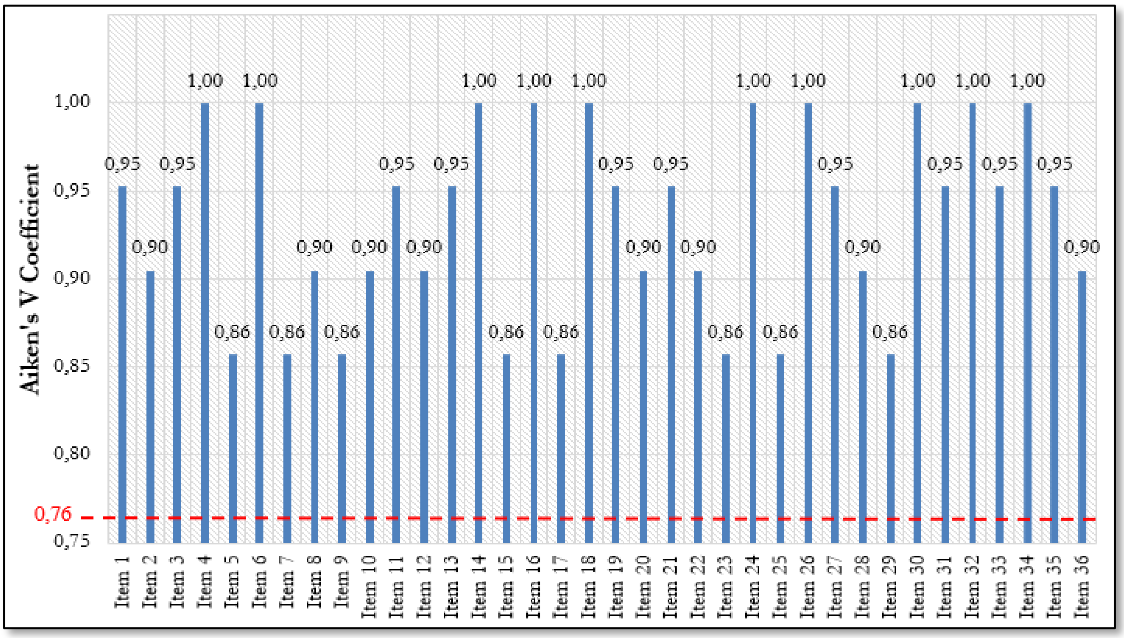


Figure 2.

Expert Validation Result with Aiken's V Coefficient

Figure 2 shows all items between 0.86 to 1.00 and has exceeded the minimum of Aiken's V coefficient limits. All instrument items can be declared valid based on content validation analysis using the Aiken's V coefficient and all items are very good as referred to in Table 2. Results instrument readability test (in terms of language use and suitability of questions presented with physics material) involving 10 students showed good results without the need to be revised. Thus, this test instrument is feasible to use.

Empirical evidence obtained from the trying out stage is in the form of dichotomous data analyzed according to the Rasch model. The trying out stage involved 296 students from high schools in the Special Region of Yogyakarta, especially Bantul Regency. 36 valid items are divided into two packages (Packages A and B). Each package consists of 20 multiple choice questions, including four questions as *anchor items* (four package A items are similar to package B). Quest and Parscale programs are used to analyze empirical test results according to Rasch's IRT model.

The results of the *goodness of fit* analysis can be seen from the INFIT parameters for *Mean Square* (MNSQ) and *INFIT t* showing that the *PhysDTRA* instrument for diagnosing the mathematical representation ability meets the statistical fit criteria according to the Rasch model which is fully presented in Table 6.

Table 6.
Fit Statistics Test Parameter at 0.5 Probability Level

No	Test Parameter	Items Estimation	Cases Estimation
1	Mean & Standard Deviation	0.00 ± 0.84	-1.79 ± 0.59
2	Adjusted Standard Deviation	0.81	0.33
3	Mean & Standard Deviation of INFIT MNSQ	1.00 ± 0.05	1.00 ± 0.12
4	Mean & Standard Deviation of OUTFIT MNSQ	1.01 ± 0.26	1.01 ± 0.45
5	Mean & Standard Deviation of INFIT t	-0.02 ± 0.96	0.08 ± 0.52
6	Mean & Standard Deviation of OUTFIT t	0.03 ± 1.11	0.09 ± 0.71
7	Reliability of Estimate	0.95	
8	Mean Difficulty	0.00 ± 0.84	

The analysis result shows that the item's estimated reliability is 0.95, which means the test sample suitable the item tested and is very good as referred to in Table 6, or the sample provides consistent results and information as expected. The suitability map of 36 items with the Rasch model is shown in Figure 3.

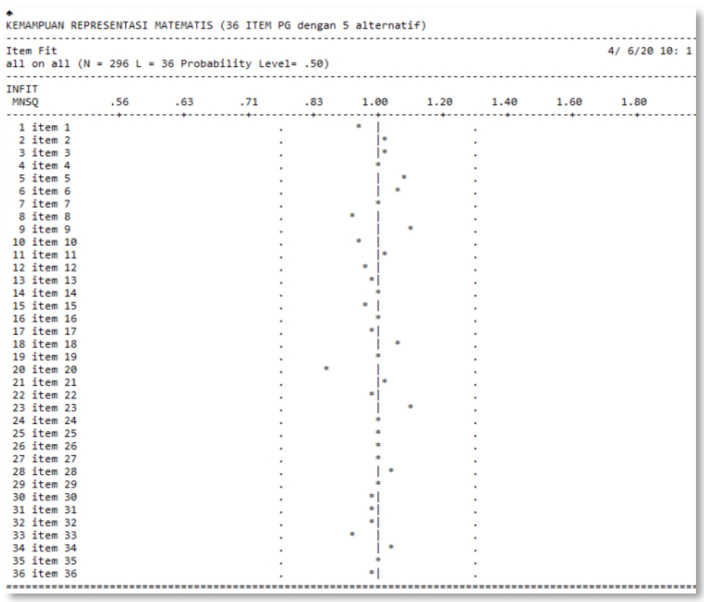


Figure 3.
INFIT MNSQ Distribution Value for each item in the FIT Model

Items or cases that have been tested are fit to the model if the MNSQ INFIT value is between 0.77 until 1.30 (Hambleton et al. 1991). Based on Figure 3, it was stated that 36 items match the Rasch model, with INFIT MNSQ values of items between 0.80-1.20.

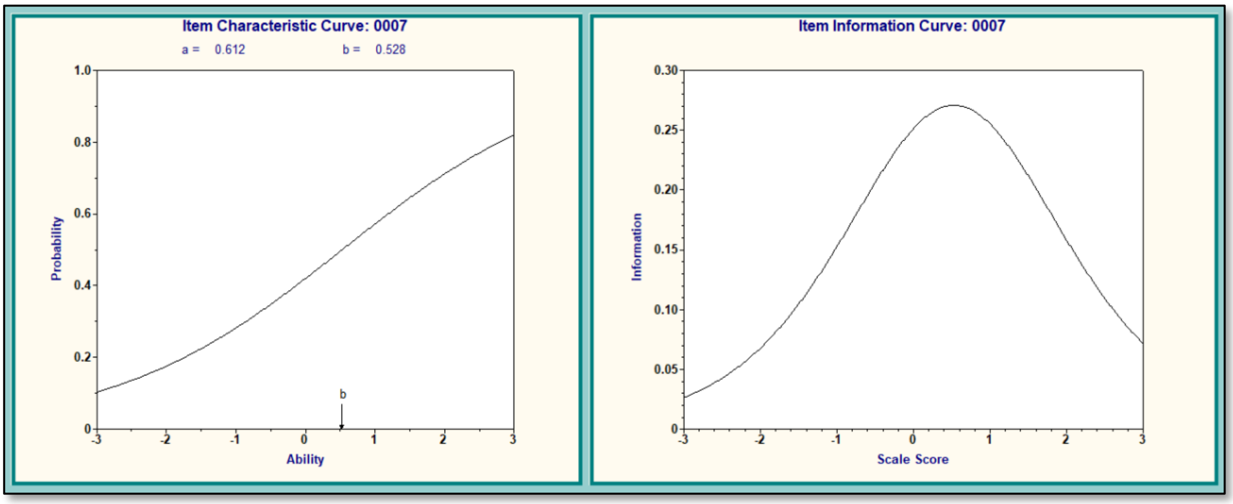


Figure 4.
Item Characteristic and Item Information Curve Number 7

Figure 4 displays the level of validity of each item to measure students' mathematical representation abilities. Item Characteristic Curve explains the value of a = 0.612 as the distinguishing power of items and b = 0.528 as the difficulty level of items. The Information Curve item shows the value of information in item 7 of 0.27 for students who have the ability on a logit scale of 0.5.

Test instrument reliability can be determined based on IRT using the total information function (TIC) curve and SEM. The Parscale program is used to obtain the TIC and SEM curves. The reliability of the instrument to measure students' mathematical representation abilities is shown in Figure 5.

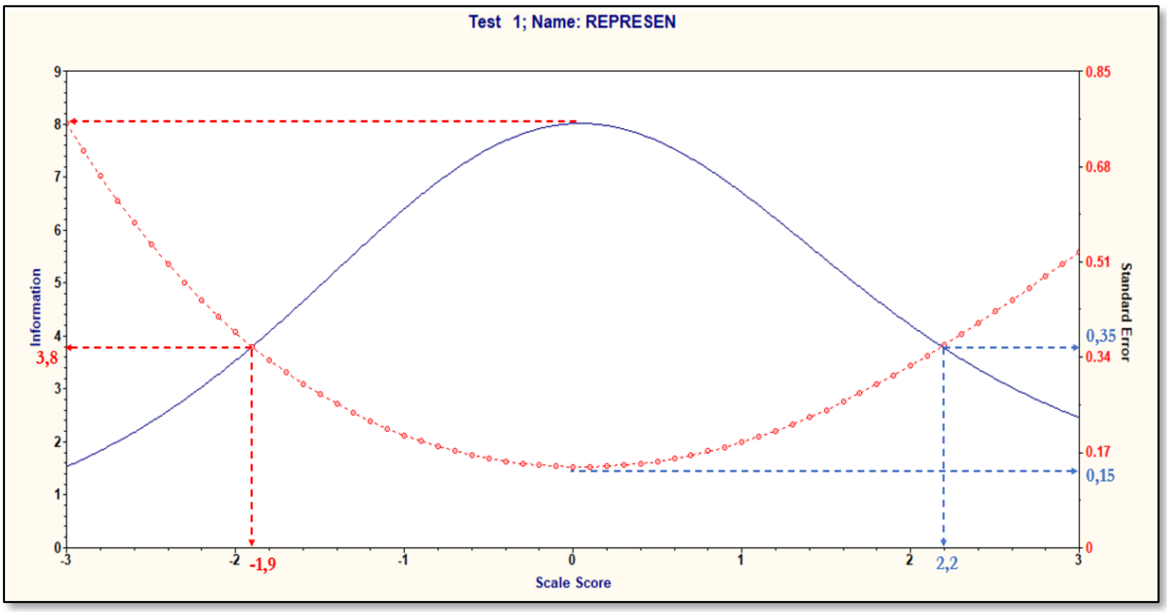


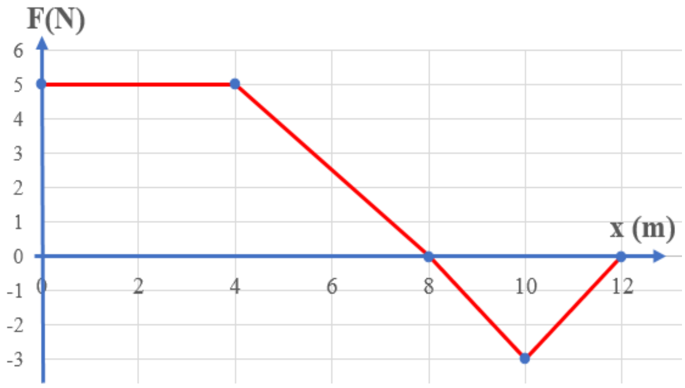
Figure 5.
Total Information Curve (TIC) and Standard Error Measurement (SEM) Test Instrument

Figure 5 explains that the *PhysDTRA* instrument for measuring mathematical representation ability is suitable for students who have the ability (θ) between -2.0 until 1.9 ($-2.0 < \theta < 1.9$). The intersection point between the blue and red line curves shows that the *PhysDTRA* Instrument has a Total Information Function (TIC) as the reliability of 3.8 and Standard Error Measurement (SEM) of 0.35. The greatest reliability is shown at the maximum point (blue curve line) of 8.0 with a standard error of 0.15.

One of the 36 items of the *PhysDTRA* instrument developed to diagnose students' mathematical representation abilities in the Work and Energy material is shown in Table 7. These items follow the cognitive level C4 (Analyzing) of A revision of Bloom's Taxonomy.

Table 7.

Item Test Instrument to Diagnose Mathematical Representation Ability

Item Indicator	Question Item 7
Presented a graph of the relationship between force and displacement, students are able to interpret the graph to determine the velocity of objects that undergo a force at a certain distance.	The force experienced by an object with a mass of 8 kg, as shown in the figure below. If at $x = 4$ m the velocity is 3 m/s, then the velocity at $x = 12$ m is ... m/s  a. $\sqrt{3}$ b. $\sqrt{5}$ c. $\sqrt{10}$ * d. $\sqrt{13}$ e. $\sqrt{15}$

Diagnosis	Suggestion
If you choose A: *Your Answer is Wrong* The student has not been able to determine the variables that affect the Work on the object and has not been able to calculate the Work from the ***F-Δx*** graph so that errors occur when interpreting the Work variables with changes kinetic energy appropriately.	**If you choose A:** You should study the variables that affect the Work in the ***F-Δx*** graph correctly so that you can interpret the Work variables with the change in kinetic energy correctly.
If you choose B: *Your Answer is Wrong* The student do not understand the topics about Work done by changing the force and implemented in graphical form.	**If you choose B:** You should go back to learning about the Work done by changing the force to be able to the Work variables in the form of graphs.
If you choose C: *Your Answer is Right* The student can go to the next topics.	**If you choose C:** no advice for you.
If you choose D: *Your Answer is Wrong* The student has been able to calculate the Work from the ***F-Δx*** graph but has not been able to determine the Work that is positive and negative so that errors occur when interpreting the Work variables with changes in kinetic energy appropriately.	**If you choose D:** You should learn how to determine the Work that is positive and negative from the ***F-Δx*** graph correctly so that you are able to interpret the Work variables with changes in kinetic energy correctly.
If you choose E: *Your Answer is Wrong* The student has not been able to determine the variables that affect the Work on objects and has not been able to determine the greatest positive value of the Work from the graph, so mistakes occur when interpreting the variables correctly.	**If you choose E:** You should learn the variables that affect the Work in the ***F-Δx*** graph correctly, such as calculating the Work from the raster area under a curve/graph, so that you are able to interpret the Work variables correctly.

The difficulty level of each item can be seen from the Quest program output shown in Figure 6.

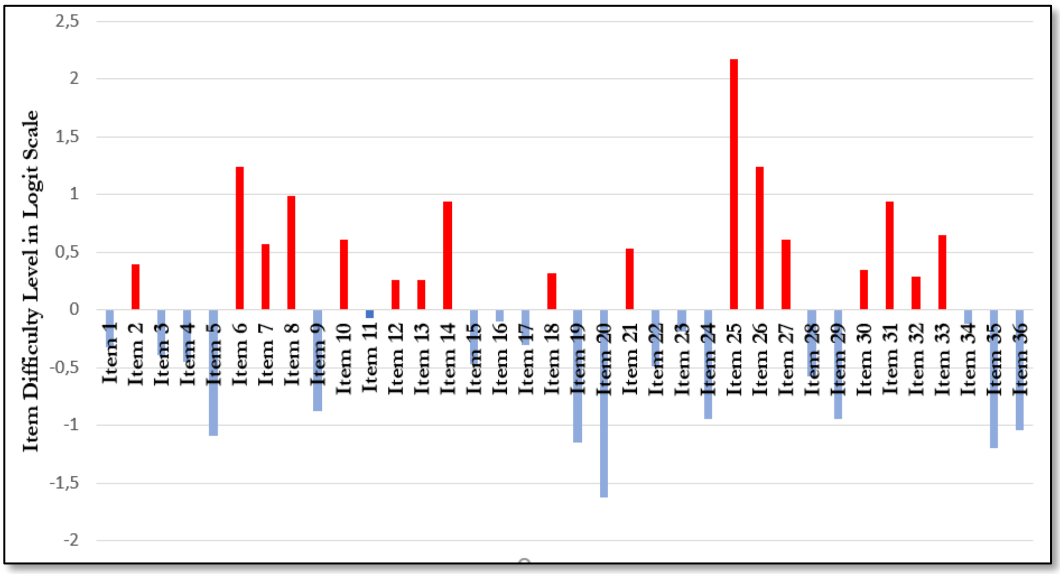


Figure 6.
Item Difficulty Level

Figure 6 shows a graph of the distribution results of student answers with difficulty ranging from easy to difficult questions. Retnawati (2014) stated that the difficulty level (b) of an item is good if it has an item difficulty index between -2.00 until 2.00. Items with a difficulty level of -2.00 indicate that the item is very easy, while a difficulty level of 2.00 indicates that the item is very difficult.

Students' mathematical representation ability is interpreted in five scales according to Azwar (2012) which is very low, low, medium, high, and very high. The student's mathematical representation ability is tetha (θ) and is shown in Figure 7.

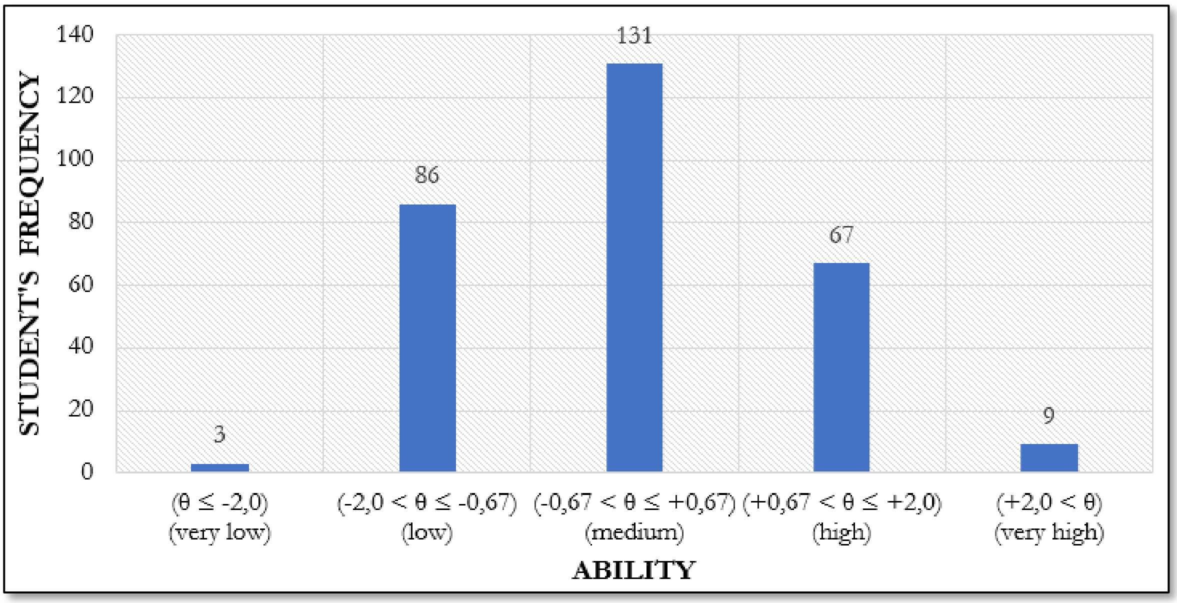


Figure 7.
Students' Mathematical Representation Ability Distributions

Students' abilities can be identified from the PH3 and SCO file format in the Parscale program output. Students' mathematical representation abilities (θ) are presented in the ability column on the logit scale. The measurement results of the mathematical representation ability of 296 students produce a distribution of values between -2.5037 until 2.4977 on a logit scale between -4 until +4.

Discussion

The test instrument which was composed of items in this research was categorized as valid according to the judgment of the experts. The minimum value of the Aiken's V coefficient is 0.76 based on seven raters with four rating scales.

The Aiken's V coefficient for all items is between 0.86 until 1.00 (Figure 2), thus the items developed have met the validity (Aiken, 1985; Azwar, 2012).

The development of the test instrument is in line with research conducted by Wati et al. (2019) which states that all test items having an INFIT MNSQ value are between 0.77 until 1.30. This is evidenced by the map suitability of items that have been fit with the Rasch model (Adams & Khoo, 1996). The *PhysDTRA* instrument also has an item reliability coefficient of 0.95 with a very good reliability category. This is supported by George & Mallery (2020) which states that the test instrument is reliable and excellent based on the interpretation of the reliability coefficient. The TIC curve obtained using the Parscale program (Figure 5) proves that the *PhysDTRA* instrument is very reliable for students with abilities ranging from -2.0 until 1.9 on a logit scale. Item characteristics can be observed from the Item Characteristic Curve (ICC) and Item Information Curve (IIC). Item number 7 represents 36 items on the *PhysDTRA* instrument indicating that the item used in the measurement is feasible and valid because it has a high peak curve.

The lower the item difficulty index, the easier the item to solve, and vice versa, the higher the item difficulty index, the more difficult the item is solved (Subali & Pujiati, 2012). Good items for diagnostic purposes are items that are not too easy and not too difficult (Tumanggor et al. 2020). The level of difficulty items developed to diagnose students' mathematical representation ability is between -1.62 to 2.17 (Figure 6) with an average of 0.00 and a standard deviation of 0.84. Item number 20 has the lowest level of difficulty and item number 25 has the highest level of difficulty. Good items if the item difficulty index is between -2.00 to 2.00, so from 36 items there is 1 item (item 25) that exceeds the highest difficulty level so that it is categorized as an item not good for diagnosing ability student.

Based on the content validity, the *PhysDTRA* was declared to be valid according to the judgment of experts who were analyzed using the Aiken's V equation. All items in the test instrument were valid based on the Rasch model, INFIT MNSQ, and INFIT t. The *PhysDTRA* instrument has also been relied upon based on the reliability of item estimation and TIC curves so that it can be used to diagnose students' mathematical representation abilities. Thus, the *PhysDTRA* instrument developed has fulfilled the test characteristics that are feasible of its content, empirical evidence, validity, and reliability.

The distribution of students' mathematical representation ability is classified in five scales according to Azwar (2012), which are very low, low, medium, high, and very high (Figure 7). The frequency of students at 1.0% has very low representation abilities, and 29.1% with low ability categories. The frequency of students in the mathematical representation ability with a moderate level of 44.3%. Then 22.6% of students have a high level of mathematical representation ability and 3.0% with a very high category.

Conclusion and Recommendations

The diagnostic test instrument (*PhysDTRA*) to detect the mathematical representation ability of high school students in the form of a multiple-choice test equipped with a diagnosis and suggestion has met *Aiken's validity content* based on *expert judgment* with valid criteria and instrument reliability with the excellent category. The validity process is needed in the development of test instruments. Validation of tests used in education should involve analysis of test content and empirical analysis of test scores and data on student responses to test items (Lissitz & Samuelsen, 2007). The suitability test of the model with the data is the benchmark used in selecting the analysis model to be applied to the data. That becomes something important considering that ultimately the analysis executed will be used to estimate individual ability (du Toit, 2003; Kim, 2006; Swaminathan et al. 2007).

The empirical validity was obtained from the analysis of the test responses given to students. The responses were obtained from trying out the test to students. The empirical validity can be determined using Item Response Theory (IRT) (Retnawati, 2016). The Rasch model is part of the IRT which can be done with the help of the Quest program. The items are declared valid if the MNSQ INFIT value is in the range 0.77 to 1.30 (Adams & Khoo, 1996). Some researchers use the MNSQ INFIT limit in the range 0.5 to 1.50 to determine the fit of the item (Planinic et al. 2013; Wati et al. 2019), but this research uses a tighter limit, namely in the range 0.77 to 1.30 to prove the suitability of all items is more fit (Wei et al. 2014).

Reliability (α) of a test is generally expressed numerically within the coefficient range $-1.00 \leq \alpha \leq +1.00$. High coefficient indicates high reliability. Conversely, if the coefficient of a test score is low, the reliability is low. The process of calculating reliability is called estimation. Estimation of the test reliability carried out in this study through composite reliability, specifically, calculating the α of Cronbach (Geldhof et al. 2014). Reliability is related to measurement error. High reliability indicates a small error from the measurement results, and vice versa, the smaller the reliability score, the greater the error of the measurement results. The student's ability score (θ) is between -4 and

+4 according to the origin of the normal distribution. This statement underlines the difficulty level value in item (*b*). An item is good if the item value is in the range of -2 to +2. If the *b* value leads to -2, the item difficulty index is very low, whereas if the *b* value leads to +2, the item difficulty index is very high for a group of test-takers (Hambleton et al. 1991; Retnawati, 2014; Supahar & Prasetyo, 2015).

The item information function in IRT is a method for explaining the strength of an item on a test set, selecting test items, and comparing several test sets. The item information function states item strength to reveal the latent trait measured by the test (du Toit, 2003; Johnson & Christensen, 2016; Retnawati, 2014; Tezza et al. 2011). The value of the item parameter index and the student's ability is the result of estimation and has a probability, so it cannot be separated from measurement errors. Standard Error of Measurement (SEM) in IRT is closely related to information functions (Figure 5). The information function is inversely quadratic to SEM, the greater the information function, the smaller the SEM or vice versa (Hambleton et al. 1991; Retnawati, 2014; Wang & Chen, 2016).

This proves that the diagnostic test instrument has met the requirements of the appropriate instrument and is ready to be implemented. The *PhysDTRA* instruments have 36 fit items consisting of 32 main items and four *anchor items* based on empirical validity. The item proportion based on the items' difficulty level varied from easy, medium, and difficult. The information function of the *PhysDTRA* instrument can provide maximum information when given to students with moderate (medium) levels of ability. The suitability of the items based on the ICC graph shows that the *PhysDTRA* instrument developed is following the ability level of high school students.

The application of the *PhysDTRA* instruments should be studied by teachers or educators first before being applied to students so that learning objectives can be achieved. It is recommended to implement the *PhysDTRA* instrument online with web assistance to make it more effective and efficient when held or implement to students in large numbers, as well as add animation and video to the diagnostic questions given so that students can be more visually interested in working on mathematical representation questions. Web-based assessments can automate and personalize feedback to students, as well as provide direct suggestion according to student response data, and also produce progress and proficiency reports for teachers and students individually or in groups (Wang & Chen, 2016). Thus, *PhysDTRA* instruments that are integrated with web-based assessments can be carried out effectively and accessed with the help of a computer or smartphone.

For Further Research

The author realizes that the physics content about Work and Energy has a broad scope and is related to other physics content. It is necessary to do an in-depth study of the major branches of physics regarding mechanics and combine it with the representation ability. The author observes that it is important to develop an assessment instrument to diagnose multiple representation ability, namely the combination of mathematical representation ability with several other representation abilities and vice versa.

For Applicants

The results of this research can be used as input for test instrument developers, educational practitioners, especially in the cognitive aspects. It is also necessary to focus on teacher knowledge related to material, pedagogy, and student character. Commonly, this research shows that the main difficulties faced by teachers or educators when providing remediation after assessment. In conclusion, teachers need the ability and extra time to construct diagnostic instruments and arrange each diagnosis and suggestion according to the indicators to be achieved.

Limitations of the Study

The scope of this research is limited to learning physics in the case study of Work and Energy at the high school level. Therefore, researchers in the field of education must carry out further research with different materials and levels of education (class) to detect students' representational abilities.

Acknowledgement

The author would like to thank for funding support from the Directorate of Research and Community Service (DRPM), Deputy for Strengthening Research and Development of the Ministry of Research, Technology, and Higher Education/ National Research and Innovation Agency through Master Thesis Research, No. T/9.65/UN34.21/PT.01.03/2020.

Biodata of Authors

Awal Mulia Rejeki Tumanggor, M.Pd., was born in Medan, North Sumatera, Indonesia. He graduated from Department of Physics Education, Faculty of Mathematics and Science of Medan State University in 2016. He completed his master of Physics Education in Yogyakarta State University, Special Region of Yogyakarta, Indonesia in 2020. His research interest includes E-Assessment and Evaluational Education especially in Physics Education. **Affiliation:** Department of Physics Education, Graduate Program, Yogyakarta State University, Daerah Istimewa Yogyakarta Indonesia. **E-mail:** awalmrt94@gmail.com **Phone:** (+62)85371492401 **Orcid ID:** https://orcid.org/0000-0002-2365-1581 **Scopus ID:** 57217964003 **WoS ID:** -

Dr. Supahar, M.Si., was born in Kulon Progo, Indonesia. He graduated from Department of Physics Education IKIP Yogyakarta in 1992. He completed master of Department of Physics, ITB Bandung in 2000, and then he completed his doctorate in the Department of Measurement and Assessment Physics Education, Yogyakarta State University in 2014. He is a lecturer in Department of Physics Education, Graduate Program, Yogyakarta State University, Daerah Istimewa Yogyakarta, Indonesia. **Affiliation:** Department of Physics Education, Graduate Program, Yogyakarta State University, Indonesia **E-mail:** supahar@uny.ac.id **Phone:** (+62)87734853276 **Orcid number:** https://orcid.org/0000-0002-2486-5549 **ID Scopus:** 57195466555 **WoS ID: -**

References

Adams, R. J., & Khoo, S.-T. (1996). *QUEST: The interactive test analysis system version 2.1*. Australian Council for Educational Research.

Adams, W. K., Wieman, C. E., & Adams, W. K. (2015). Analyzing The Many Skills Involved in Solving Complex Physics Problems Analyzing The Many Skills Involved in Solving complex Physics Problems. *American Journal of Physics*, *83*(459). https://doi.org/10.1119/1.4913923

Adlina, A., & Supahar. (2019). Developing android assisted worked example application on kinematics (weak) to improve mathematical representation ability in high school physics learning. *International Journal of Scientific and Technology Research*, *8*(10), 3790–3793.

Aiken, L. R. (1985). Three Coefficients for Analysing Reliability and Validity of Ratings. *Educational and Psychological Measurement*, *45*, 131–142. https://doi.org/10.1177/07399863870092005

Albe, V., Venturini, P., & Lascours, J. (2014). Electromagnetic Concepts in Mathematical Representation of Physics. *Journal of Science Education and Technology*, *10*, 197–203.

Angell, C., Kind, P. M., Henriksen, E. K., & Guttersrud, Ø. (2008). An Empirical-Mathematical Modelling Approach to Upper Secondary Physics. *Physics Education*, *43*(3), 256–264. https://doi.org/10.1088/0031-9120/43/3/001

Atkin, J. M., & Coffey, J. E. (2003). *Everyday Assessment in The Science Classroom* (Issue c). National Science Teachers Association.

Azwar, S. (2012). *Penyusunan Skala Psikologi (Psychological Scale Compilation)*. Pustaka Pelajar.

Azwar, S. (2015). *Reliability and Validity* (fourth). Pustaka Pelajar.

Bing, T. J., & Redish, E. F. (2014). The Cognitive Blending of Mathematics and Physics Knowledge. *AIP Conference Proceedings*, *883*(February 2007), 26–29. https://doi.org/10.1063/1.2508683

Bollen, L., van Kampen, P., Baily, C., Kelly, M., & De Cock, M. (2017). Student difficulties regarding symbolic and graphical representations of vector fields. *Physical Review Physics Education Research*, *13*(2), 020109. https://doi.org/10.1103/PhysRevPhysEducRes.13.020109

Ceuppens, S., Deprez, J., Dehaene, W., & De Cock, M. (2018). Design and validation of a test for representational fluency of 9th grade students in physics and mathematics: The case of linear functions. *Physical Review Physics Education Research*, *14*(2), 20105. https://doi.org/10.1103/PhysRevPhysEducRes.14.020105

Darmawan, D., Yatimah, D., Sasmita, K., & Syah, R. (2020). Analysis of Non-Formal Education Tutor Capabilities in Exploring Assessment for Science Learning. *Jurnal Pendidikan IPA Indonesia*, *9*(2), 267–275. https://doi.org/10.15294/jpii.v9i2.24025

De Cock, M. (2012). Representation Use and Strategy Choice in Physics Problem Solving. *Physical Review Special Topics - Physics Education Research*, *8*(2), 1–15. https://doi.org/10.1103/PhysRevSTPER.8.020117

Docktor, J. L., & Mestre, J. P. (2014). Synthesis of Discipline-Based Education Research in Physics. *Physical Review Special Topics - Physics Education Research*, *10*(2), 1–58. https://doi.org/10.1103/PhysRevSTPER.10.020119

du Toit, M. (2003). *IRT from SSI: Bilog-Mg, Multilog, Parscale, Testfact*. Scientific Software International, Inc.

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, *19*(1), 72–91. https://doi.org/10.1037/a0032138

George, D., & Mallery, P. (2020). *IBM SPSS Statistics 26 Step by Step: A Simple Guide and Reference* (Sixteenth). Routledge. https://doi.org/10.4324/9780429056765

Gurcay, D., & Gulbas, E. (2015). Development of Three-Tier Heat, Temperature and Internal Energy Diagnostic Test. *Research in Science and Technological Education*, *33*(2), 197–217. https://doi.org/10.1080/02635143.2015.1018154

Gurel, D. K., Eryilmaz, A., & McDermott, L. C. (2015). A Review and Comparison of Diagnostic Instruments to Identify

Students' Misconceptions in Science. *Eurasia Journal of Mathematics, Science and Technology Education*, *11*(5), 989–1008. https://doi.org/10.12973/eurasia.2015.1369a

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory Library*. Sage Publication, Inc.

Heuvelen, A. Van, & Zou, X. (2001). Multiple representations of work–energy processes. *American Journal of Physics*, *69*(2), 184–194. https://doi.org/10.1119/1.1286662

Ibrahim, B., & Rebello, N. S. (2012). Representational Task Formats and Problem Solving Strategies in Kinematics and Work. *Physical Review Special Topics - Physics Education Research*, *8*(1), 1–19. https://doi.org/10.1103/PhysRevSTPER.8.010126

Johnson, R. B., & Christensen, L. B. (2016). *Educational Research: Quantitative, Qualitative, and Mixed Approaches* (Sixth). SAGE Publication.

Kim, S. (2006). A comparative study of IRT fixed parameter calibration methods. *Journal of Educational Measurement*, *43*(4), 355–381. https://doi.org/10.1111/j.1745-3984.2006.00021.x

Kirschner, S., Borowski, A., Fischer, H. E., Gess-Newsome, J., & von Aufschnaiter, C. (2016). Developing and evaluating a paper-and-pencil test to assess components of physics teachers' pedagogical content knowledge. *International Journal of Science Education*, *38*(8), 1343–1372. https://doi.org/10.1080/09500693.2016.1190479

Krawec, J. L. (2014). Problem Representation and Mathematical Problem Solving of Students of Varying Math Ability. *Journal of Learning Disabilities*, *47*(2), 103–115. https://doi.org/10.1177/0022219412436976

Kriek, J., & Koontse, R. D. (2017). First year physics students' expectations of the role of mathematics in physics. *International Journal of Innovation in Science and Mathematics Education*, *25*(2), 1–16.

Liang, L. L., Fulmer, G. W., Majerich, D. M., Clevenstine, R., & Howanski, R. (2012). The Effects of a Model-Based Physics Curriculum Program with a Physics First Approach: A Causal-Comparative Study. *Journal of Science Education and Technology*, *21*(1), 114–124. https://doi.org/10.1007/s10956-011-9287-2

Lissitz, R. W., & Samuelsen, K. (2007). Further Clarification Regarding Validity and Education. *Educational Researcher*, *36*(8), 482–484. https://doi.org/10.3102/0013189x07311612

Loewenthal, K. M., & Lewis, C. A. (2018). *An introduction to psychological tests and scales*. Psychology Press.

Minarni, A., & Napitupulu, E. E. (2017). Developing Instruction Materials Based on Joyful PBL to Improve Students Mathematical Representation Ability. *International Education Studies*, *10*(9), 23. https://doi.org/10.5539/ies.v10n9p23

Niss, M. (2016). Obstacles Related to Structuring for Mathematization Encountered by Students when Solving Physics Problems. *International Journal of Science and Mathematics Education*, *15*(8), 1441–1462. https://doi.org/10.1007/s10763-016-9754-6

Nitko, A. J., & Brookhart, S. M. (2011). *Educational Assessment of Students* (Sixth Edit). Pearson Education, Inc.

OECD. (2019). PISA 2018 Results (Volume II): Where All Students Can Succeed. In *OECD Publishing: Vol. II*. https://doi.org/https://doi.org/10.1787/b5fd1b8f-en

Oriondo, L. L., & Dallo-Antonio, E. M. (1984). *Evaluating Educational Outcomes (Test, Measurement, and Evaluation)* (5th Editio). REX Printing Company, Inc.

Pape, S. J., & Tchoshanov, M. A. (2001). The Role of Representation(s) in Developing Mathematical Understanding. *Theory Into Practice*, *40*(2), 118–127. https://doi.org/10.1207/s15430421tip4002_6

Park, E.-J., & Choi, K. (2013). Analysis of Student Understanding of Science Concepts Including Mathematical Representations: pH Values and The Relative Differences of pH Values. *International Journal of Science and Mathematics Education*, *11*(3), 683–706. https://doi.org/10.1007/s10763-012-9359-7

Permendikbud No. 22 Tahun 2016 tentang Standar Proses Pendidikan Dasar dan Menengah (*Minister of Education and Culture Regulation No. 22 of 2016 on Process Standards for Primary and Secondary Education*), (2016).

Planinic, M., Ivanjek, L., Susac, A., & Milin-Sipus, Z. (2013). Comparison of university students' understanding of graphs in different contexts. *Physical Review Special Topics - Physics Education Research*, *9*(2), 1–9. https://doi.org/10.1103/PhysRevSTPER.9.020103

Pospiech, G., Eylon, B., Bagno, E., Lehavi, Y., & Geyer, M. A. (2016). The Role of Mathematics for Physics Teaching and Understanding. *Nuovo Cimento C Geophysics Space Physics C*, *38*(3), 10. https://doi.org/https://doi.org/10.1393/ncc/i2015-15110-6

Pujayanto, Budiharti, R., Radiyono, Y., Rizky, N., Nuraini, A., Putri, H. V., Saputro, D. E., & Adhitama, E. (2018). Pengembangan Tes Diagnostik Miskonsepsi Empat Tahap Tentang Kinematika (*Developing Four-Tier Misconception Diagnostic Test about Kinematics*). *Cakrawala Pendidikan*, *37*(2), 237–249. https://doi.org/10.21831/cp.v37i2.16491

Retnawati, H. (2014). *Teori Respon Butir dan Penerapannya: untuk Peneliti, Praktisi Pengukuran dan Pengujian, Mahasiswa Pascasarjana (Item Response Theory and Its Application: for Researchers, Measurement and Testing Practitioners, Postgraduate Students)*. Nuha Medika.

Retnawati, H. (2016). *Analisis Kuantitatif Instrumen Penelitian: Panduan Peneliti, Mahasiswa, dan Psikometrian (Quantitative Analysis of Research Instruments: A Guide for Researchers, Students, and Psychometrics)*. Parama Publishing.

Sadaghiani, H. R. (2011). Using multimedia learning modules in a hybrid-online course in electricity and magnetism. *Physical Review Special Topics - Physics Education Research*, *7*(1), 1–7. https://doi.org/10.1103/PhysRevSTPER.7.010102

Subali, B., & Pujiati, S. (2012). *Pengembangan Item Tes Konvergen dan Divergen dan Penyelidikan Validitasnya secara Empiris (Development of Convergent and Divergent Test Item and Empirical Validity Investigation)*. Diandra Pustaka Indonesia.

Subali, Bambang, & Suyata, P. (2011). *Panduan Analisis Data Pengukuran Pendidikan Untuk Memperoleh Bukti Empirik Kesahihan menggunakan program Quest (Guidelines for Educational Measurement Data Analysis to Obtain Empiric Proof of Eligibility using the Quest Program)*. Lembaga Penelitian dan Pengabdian pada Masyarakat Universitas Negeri Yogyakarta.

Supahar, S., & Prasetyo, Z. K. (2015). Pengembangan Instrumen Penilaian Kinerja Kemampuan Inkuiri Peserta Didik Pada Mata Pelajaran Fisika SMA (*Developing Assessment Instrument of The Performance of The Inquiry Abilities in Physics Subject of High School Students*). *Jurnal Penelitian Dan Evaluasi Pendidikan*, *19*(1), 96–108. https://doi.org/10.21831/pep.v19i1.4560

Swaminathan, H., Hambleton, R. K., & Rogers, H. J. (2007). Assessing the Fit of Item Response Theory Models. In *Handbook of Statistics* (Vol. 26, Issue 06, pp. 683–718). Elsevier B.V. https://doi.org/10.1016/S0169-7161(06)26021-8

Tezza, R., Bornia, A. C., & Andrade, D. F. De. (2011). Measuring web usability using item response theory: Principles, features and opportunities. *Interacting with Computers*, *23*(2), 167–175. https://doi.org/10.1016/j.intcom.2011.02.004

Thiagarajan, S., Semmel, D. S., & Semmel, M. I. (1974). Instructional development for training teachers of exceptional children: A sourcebook. *Journal of School Psychology*, *14*(1), 75. https://doi.org/10.1016/0022-4405(76)90066-2

Treagust, D. F., Duit, R., & Fischer, H. E. (2017). *Multiple Representations in Physics Education*. Springer International Publishing. https://doi.org/10.1007/978-3-319-58914-5_7

Tumanggor, A. M. R., Jumadi, J., Wilujeng, I., & Ringo, E. S. (2019). The Profile of Students' Physics Problem Solving Ability in Optical Instruments. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, *5*(1), 29–40. https://doi.org/10.21009/1.05104

Tumanggor, A. M. R., Supahar, Kuswanto, H., & Ringo, E. S. (2020). Using four-tier diagnostic test instruments to detect physics teacher candidates' misconceptions : Case of mechanical wave concepts. *Journal of Physics: Conference Series*, *1440*, 1–8. https://doi.org/10.1088/1742-6596/1440/1/012059

Tumanggor, A. M. R., Supahar, Ringo, E. S., & Harliadi, M. D. (2020). Detecting Students ' Misconception in Simple Harmonic Motion Concepts Using Four-Tier Diagnostic Test Instruments. *Jurnal Ilmiah Pendidikan Fisika Al-Biruni*, *9*(1), 21–31. https://doi.org/10.24042/jipfalbiruni.v9i1.4571

Uhden, O., Karam, R., Pietrocola, M., & Pospiech, G. (2012). Modelling Mathematical Reasoning in Physics Education. *Science & Education*, *21*(4), 485–506. https://doi.org/10.1007/s11191-011-9396-6

Volfson, A., Eshach, H., & Ben-Abu, Y. (2018). Development of a diagnostic tool aimed at pinpointing undergraduate students' knowledge about sound and its implementation in simple acoustic apparatuses' analysis. *Physical Review Physics Education Research*, *14*(2), 20127. https://doi.org/10.1103/PhysRevPhysEducRes.14.020127

Wang, J., & Chen, S. (2016). Development and Validation of an Online Dynamic Assessment for Raising Students' Comprehension of Science Text. *International Journal of Science and Mathematics Education*, *14*(3), 373–389.

Wati, M., Mahtari, S., Hartini, S., & Amalia, H. (2019). A Rasch model analysis on junior high school students' scientific reasoning ability. *International Journal of Interactive Mobile Technologies*, *13*(7), 141–149. https://doi.org/10.3991/ijim.v13i07.10760

Wei, S., Liu, X., & Jia, Y. (2014). Using Rasch Measurement to Validate The Instrument of Students' Understanding of Models in Science (SUMS). *International Journal of Science and Mathematics Education*, *12*(5), 1067–1082. https://doi.org/10.1007/s10763-013-9459-z

Wilcox, B. R., & Pollock, S. J. (2015). Validation and analysis of the coupled multiple response Colorado upper-division electrostatics diagnostic. *Physical Review Special Topics-Physics Education Research*, *11*(2). https://doi.org/10.1103/PhysRevSTPER.11.020130

Yeh, Y. F., Hsu, Y. S., Wu, H. K., & Chien, S. P. (2016). Exploring The Structure of TPACK with Video-Embedded and Discipline-Focused Assessments. *Computers and Education*, *104*, 49–64. https://doi.org/10.1016/j.compedu.2016.10.006

Zhe, L. (2012). Survey of Primary Students' Mathematical Representation Status and Study on the Teaching Model of Mathematical Representation. *Journal of Mathematics Education*, *5*(1), 63–76.

Appendix 1.
Mathematical Representation Ability Scale

Case of Study: WORK

Interpreting variables according to concept in the form of pictures, tables, diagrams, and graphs to solve problems
Presented an image of a block on a plane that has friction force, students are able to:
1. interpret the magnitude of the angle correctly to do the smallest work. (Case 1)
2. interpret the magnitude of the angle correctly to do the greatest work.
3. interpret the magnitude of the angle correctly to do the smallest work. (Case 2)

Connecting the variables contained in a problem in the form of pictures, tables, diagrams, and graphs to solve problems
Presented a graph of the force to the position, students are able to:
4. relate the variables by determining the combination of the greatest work done by a varying force at certain points. (Case 1)
5. relate the variables by determining the combination of the smallest work done by a varying force at certain points.
6. relate the variables by determining the combination of the greatest work done by a varying force at certain points. (Case 2)

Operating equations appropriately in the form of numbers, symbols, graphs, or images in solving problems
Presented an illustrative image of an object on a flat plane, students are able to:
7. Operate the Work equation on the system appropriately. (Case 1)
8. Operate the Work equation on the system appropriately. (Case 2)
9. Operate the Work equation on the system appropriately. (Case 3)

Concluding conditions by operating mathematical equations to obtain results
Presented an illustrative image of an object on an inclined plane, students are be able to:
10. conclude conditions based on the concept by connecting the concept of work with the displacement of objects.
11. conclude conditions based on the concept of an inclined plane correctly. (Case 1)
12. conclude conditions based on the concept of an inclined plane correctly. (Case 2)

Case of Study: KINETIC AND POTENTIAL ENERGY

Interpreting variables according to concept in the form of pictures, tables, diagrams, and graphs to solve problems
Presented a graph of the relationship between force and displacement, students are able to:
13. interpret the graph to determine the velocity of objects that undergo a force at a certain. (Case 1)
14. interpret the graph to determine the velocity of objects that undergo a force at a certain. (Case 2)
15. interpret the graph to determine the velocity of objects that undergo a force at a certain. (Case 3)

Connecting the variables contained in a problem in the form of pictures, tables, diagrams, and graphs to solve problems
Presented a graph of the relationship between velocity and time, students are able to:
16. relate the relationship between work and the resultant force applied to the object. (Case 1)
17. relate the relationship between work and the resultant force applied to the object. (Case 2)
18. relate the relationship between work and the resultant force applied to the object. (Case 3)

Operating equations appropriately in the form of numbers, symbols, graphs, or images in solving problems
Presented illustration case of an object, students are expected to be able to:
19. operate the data so that they find the right mathematical equation. (Case 1)
20. operate the data so that they find the right mathematical equation. (Case 2)
21. operate the data so that they find the right mathematical equation. (Case 3)

Concluding conditions by operating mathematical equations to obtain results
22. Presented a ball is at a certain height, students are expected to be able to conclude conditions based on the concept and obtain the results of the work of weight from mathematical operations.
23. Presented a block that is placed on a flat plane, students can conclude conditions based on the concept and obtain the results of block displacement from mathematical operations using changes of kinetic energy. (Case 1)
24. Presented a block that is placed on a flat plane, students can conclude conditions based on the concept and obtain the results of block displacement from mathematical operations using changes of kinetic energy. (Case 2)

Case of Study: CONSERVATION OF MECHANICAL ENERGY

Interpreting variables according to concept in the form of pictures, tables, diagrams, and graphs to solve problems

Presented an illustrative image of objects experiencing free-fall motion with different heights and final velocities, students are able to:

25. interpret the mechanical energy of objects correctly. (Case 1)
26. interpret the mechanical energy of objects correctly. (Case 2)
27. interpret the mechanical energy of objects correctly. (Case 3)

Connecting the variables contained in a problem in the form of pictures, tables, diagrams, and graphs to solve problems

Presented an image of objects that are in a certain position, students are able to:

28. relate the velocity of the object at each point in the law of mechanical energy conservation. (Case 1)
29. relate the height of the object at each point in the law of mechanical energy conservation.
30. relate the velocity of the object at each point in the law of mechanical energy conservation. (Case 2)

Operating equations appropriately in the form of numbers, symbols, graphs, or images in solving problems

31. Presented illustrative images of objects experiencing free-fall motion from a certain height, students are expected to be able to operate the correct equation from the illustration of objects experiencing free-fall motion.
32. Presented an image and data, students are expected to be able to operate the correct equation from the illustration of objects experiencing free-fall motion.
33. Presented an image and data, students are expected to be able to operate the correct equation from the illustration of objects moving in the track.

Concluding conditions by operating mathematical equations to obtain results

Explained a state of an object at a certain position, students are able to:

34. conclude the mathematical operation results of the object's velocity at a certain position. (Case 1)
35. conclude the mathematical operation results of the object's height at a certain object’s position.
36. conclude the mathematical operation results of the object's velocity at a certain position. (Case 2)

Appendix 2.

The Instrument Item Review Criteria

Aspect		Review Criteria
Material	1	The item matches the indicator in the item blueprint
	2	The item matches the basic competency
	3	The item matches the indicators of achievement
	4	The concept given is correct
	5	The answer key given matches the question
	6	Use the correct units
	7	Write down units correctly
	8	Write down the equation exactly
Construction	9	Item formulated with short, dense clearly and unambiguously
	10	The statement on the item does not provide a clue to the answer key
	11	Items are free from negative statements
	12	The formulation of items and the answer choices are only statements that are needed.
	13	The item does not depend on the answer to the previous question
	14	Pictures, graphs, tables, diagrams, or the like are clear and functional (if provided).
	15	The length of the answer choices is relatively the same.
	16	Homogeneous and logical answer choices.
	17	The answer choices in the form of numbers are sorted according to the size of the number
	18	The answer choices do not use the statement "all the above answers are wrong/true" and the like
Language	19	Use the Indonesian language according to the General Guidelines for Indonesian Spelling (PUEBI)
	20	Use the correct sentence structure.
	21	Be consistent in using terms.
	22	Items arranged communicatively.
	23	The expressions used are unambiguous.
	24	Item does not contain the term "local language"