

# Quality Analysis of Test for Diagrams and Mathematical Representation Ability of High School Student in Physics Using Partial Credit Model

Jenny Resti Harjanti  
*Department of Physics, Graduate School*  
 Yogyakarta State University  
 Yogyakarta, Indonesia  
 jenny.resty2016@student.uny.ac.id

Supahar  
*Department of Physics, Graduate School*  
 Yogyakarta State University  
 Yogyakarta, Indonesia

Warsono  
*Department of Physics, Graduate School*  
 Yogyakarta State University  
 Yogyakarta, Indonesia

Jumadi  
*Department of Physics, Graduate School*  
 Yogyakarta State University  
 Yogyakarta, Indonesia

**Abstract**—The quality of learning activities can be determined by the result of assessment. The measuring tool used in the assessment activity is in the form of a test. This study aims to determine the quality of a test for diagrams and mathematical representation ability of High School students in Physics. This research is quantitative descriptive. There are two sets of test that are set A and B, which is each set test consist of 10 items with 20% of the being an anchor items. Test instrument tested to 317 High School students in Physics in Yogyakarta with purposive sampling. Students' response data was analyzed based on the items response theory using Partial Credit Model (PCM). PCM is the development of 1-PL model (Rasch model) which contains one parameter item that is the item difficulty level. The results of test quality analysis show that: (1) items test instrument fit with PCM, (2) item difficulty levels are in the range of values -1.023 to +1.785 which includes the criteria of the good items, (3) the ability of respondents who do the test is in the range of values from -1.9678 to +2.0384, and (4) this test is more appropriate for measuring respondents who have value ability  $-1.85 \leq \theta \leq +1.9$ . Therefore, this test can be used to measure diagrams and mathematical representation ability of High School students in Physics.

**Keywords** — *quality analysis, representation ability, Partial Credit Model (PCM)*

## I. INTRODUCTION

Physics is part of the science that has an important role in life, especially in the field of science and technology. Physics has strong characteristics; because in proving a theory, one must perform

systematic experiments that can be tested, Physics is considered difficult and understandable [1]. Physics aims to study and provide both qualitative and quantitative understanding of the various phenomena or processes of nature and the nature of matter and its application [2]. Therefore, a very important Physics lesson is taught in secondary school.

Through Physics learning, students will get the ability to explore and solve Physics problems. There are several cognitive abilities that play a role in improving students' success in solving Physics problems i.e.: (1) ability to identify and interpret accurately the concepts and principles of Physics and (2) the ability to create descriptions and organize Physics knowledge effectively [3]. These abilities can be known through the use of Physics representation formats. The format of Physics representation is a form that can explain information in solving Physics problems [4]. Form of representation can explain the information in solving the problem. Some forms of representation include the form of textual, mathematical, diagrams, graphs and others [5]. The success of students in solving problems depends on the ability to use representational forms.

Diagram representation is one form of representation that can be used to solve Physics problems related to mechanical problems. Diagram representation is a form that can visualize objects and interactions between objects through the diagrams [6]. The use of diagram representations can reduce the level of complexity of other representations [7]. The diagrams representation can give a more contextual

description of Physics concepts. Therefore, the diagrams representation becomes one of the solutions for students to understand and solve Physics problems.

Another form of representation that can help students to solve Physics problems is mathematical representation. Mathematical representation is the process of determining the mathematical form that will be applied to the concept of Physics [8]. Mathematical form has an important role in the learning of Physics, because usually the form of mathematics is used to describe the physical events in everyday life [9]. One of the important things in solving the problem of Physics is the ability of students in combining symbols and mathematical structures with their intuition and knowledge [10]. Students can use a mathematical representation to express a mathematical idea as a useful form of problem to find a solution.

Encouragement of achievement of diagrams and mathematical representation ability is needed so that students better understand the concepts of Physics learned and can solve Physics problems. Achievement of the diagrams representation that students can describe the diagram and its components [6,11,12] and the achievement of the mathematical representation i.e. (1) students can determine the correct equations and (2) students can operationalize numbers or symbols based on equations [10,13]. To measure the achievement of students' diagrams and mathematical representation, a measuring tool such as a test is required.

The test is one of the measuring tools in the assessment of education for student learning outcomes [14]. It is a planned measurement tool used by educators to determine student achievement and its relation with learning objectives that have been determined [15]. The test is used to determine the achievement of indicators of a subject that must be controlled by students. A good test should have a basic principle of improving learning [16]. This means that tests have an important role in learning. The form of test used in this study is an essay test.

To get a high quality test instrument, in addition to theoretical analysis i.e. items review based on aspects of content, construction and language, it is also necessary to do empirical items analysis. The purpose of items analysis to obtain the characteristics of test items based on quantitative analysis. This quantitative items analysis is based on the items response theory. The results of quantitative analysis are then interpreted qualitatively.

The items response theory is a modern measurement theory commonly used in item analysis. In essence, the items response theory aims to overcome the weaknesses found in classical measurements. The goal of item response theory is to provide both invariant item statistics and ability estimates [17]. This theory uses a mathematical model to relate the item characteristics of the question with the ability of the respondent. That is, respondents who

have high ability will have a probability of responding properly greater than respondents who have low ability [18]. In item response theory, polytomous scoring model is an item response model that has answers more than two categories [19]. There are several models that can be used in the analysis of data polytomous based on item response theory, one of them using Partial Credit Model (PCM). PCM is the development of item response theory model 1 parameter of logistics (1-PL) [18].

Based on the explanation, this research aims to know the test quality for diagrams and mathematical representation ability of High School students in Physics based on item response theory using Partial Credit Model. This paper will mainly focus in goodness of fit items, difficulty level, and test information function.

## II. LITERATURE REVIEW

### A. Item Response Theory

Item Response Theory (IRT) is a way of estimating parameters in a model [21]. Another name of item response theory is latent trait theory (LTT) or characteristics curve theory (CCT). The item response theory aims to obtain data with high accuracy. In the item response theory, the probability of the correct answer given by the students, the characteristics or the parameters of the items, and the characteristics or parameters of the respondents are linked through a model formula that must be adjusted by both the test group and the respondent group [17]. That is, the same point to different respondents must obey the rules of the formula, or the same respondent to a different test item must also adhere to the formula. In this process there is something called invariance between the test item and the respondent. On modern measurements, item difficulty levels are not directly related to the respondent's ability. Therefore, the item response theory serves as an alternative approach that can analyze the test and can interpret the ability of respondents to the test in question.

The development of item response theory is based on two postulates i.e. [20] : (1) the ability of respondents on an item can be predicted by a set of factors called traits, latent traits or abilities. Trait is a person's ability dimension such as verbal ability, psychomotor, cognitive and so forth. (2) The relationship between a respondent's ability to an item and an underlying latent ability device can be illustrated by the item characteristics curve (ICC). In the item response theory, there are supporting assumptions that can indirectly be measured and proven. These assumptions include unidimensional, local independence, and parameter invariant [17, 22].

The first assumption is unidimensional, meaning that each test item measures only one ability. This assumption is very difficult to meet due to the many factors that influence the test, such as cognitive, personality, and test-execution factors, such as anxiety, motivation and tendency to guess. Therefore, unidimensional assumptions can be shown only if the

test contains only one dominant component that measures the respondents' achievement [17].

The second assumption is known as local independence. Respondents' responses on each item are statistically independent. That is, the response of respondents to an item is not related to other items in the test [17].

The third assumption is parameter invariant. Invariant parameter means the item characteristic is not dependent on the distribution of respondent parameters and parameters that characterize the respondent does not depend on the characteristics of the item. A person's ability will not change simply because of testing different levels of difficulty and the test item parameters will not change simply because they are tested on different groups of respondents [22].

### B. Goodness of Fit

There are 3 models that can be used to perform the analysis by using modern theory, namely model 1PL, 2-PL, and 3-PL. The goodness of fit of the IRT model is used to test the item characteristics after being responded by various respondents' abilities [20]. Goodness of fit test IRT model is performed for each item responded.

Goodness of fit test with PCM model is based on the mean value of INFIT Mean of Square (INFIT MNSQ) with standard deviation. If the mean value of INFIT MNSQ is close to 1.0 with standard deviation close to 0.0 then the overall items and cases fit with the PCM model [18]. The test to determine the fit of each item on the PCM model based on INFIT MNSQ value was in the range of 0.77 to 1.30 [18]. The range of INFIT MNSQ is the distribution of calibrated scores which still form leptokurtic curve [20]. If the item fit results indicate that the item fits the model, it can be continued with an analysis of the item parameter estimates and the respondent's ability.

### C. Partial Credit Model (PCM)

Partial Credit Model (PCM) is a model of polytomous scoring which is the development of the 1-PL model (Rasch model). In the PCM assumption, each item has the same discrimination, but the item difficulty level may be different. The difficulty level is a number that indicates the easy or difficulty of an item [23]. The ideal item difficulty level is  $-2 \leq b \leq 2$ , test item is categorizes as very easy if  $b < -2$  and categorizes as very difficult if  $b > +2$  [22].

PCM is a one-item analysis that requires completion steps and has 1-2 or more categories with the total number of final or multiple scores, but the difficulty level of each step is not sequenced [24]. The PCM model is suitable for use on achievement tests. For example, in a matter of Physics count that requires the stage of problem identification until the final solution. However, PCM is also appropriate for analyzing multipoint scale personality response [25].

The Rasch model contains only one location parameter item (difficulty level) which is then

developed by splitting item locations into categories. This development is called the Operating Characteristics Functions (OCF) to be the basic ingredient of making Category Response Functions (CRF). As for the equation to create OCF [26] :

$$P_{i1}(\theta) = \frac{P_{i1}(\theta)}{P_{i0}(\theta) + P_{i1}(\theta)} = \frac{\exp(\theta_n - \delta_{i1})}{1 + \exp(\theta_n - \delta_{i1})} \quad (1)$$

Information:

$\theta_n$  : level trait individual

$\delta_{i1}$  : location parameter item

Equation (1) is a prototype of the development of the Rasch model for the item of polytomous. If  $i$  is an item of polytomous with the number of categories in the item as many as 3 categories with scores of 0, 1 and 2, the CRF equation is as follows:

$$\begin{aligned} &\text{Category 0:} \\ P_{i0}(\theta) &= \frac{1}{1 + \exp(\theta_n - \delta_{i1}) + \exp[(\theta_n - \delta_{i1}) + (\theta_n - \delta_{i2})]} \\ &\text{Category 1:} \\ P_{i1}(\theta) &= \frac{\exp(\theta_n - \delta_{i1})}{1 + \exp(\theta_n - \delta_{i1}) + \exp[(\theta_n - \delta_{i1}) + (\theta_n - \delta_{i2})]} \\ &\text{Category 2:} \\ P_{i2}(\theta) &= \frac{\exp(\theta_n - \delta_{i1}) + \exp(\theta_n - \delta_{i2})}{1 + \exp(\theta_n - \delta_{i1}) + \exp[(\theta_n - \delta_{i1}) + (\theta_n - \delta_{i2})]} \end{aligned} \quad (2)$$

The probability in category 0 shows that there is a 1 in the denominator. This is because in PCM requires the following equation:

$$\sum_{j=0}^0 (\theta - \delta_{ij}) = 1 \quad (3)$$

The category scores on PCM indicate how many steps to correctly complete the item. A higher category score indicates a greater ability than a lower category score.

### D. Information Function

Item information function is a method to explain the strength of an item on the test, the selection of test items and the comparison of some tests. The information function declares the strength or contribution of test items in uncovering the latent trait measured by the test [18]. Mathematically the function of the item information can be written as follows [22]:

$$I_i(\theta) = \frac{[P'_i(\theta)]^2}{P_i(\theta)Q_i(\theta)} \quad (5)$$

Information:

$i$  : 1, 2, 3, ..., n

$I_i(\theta)$  : test information function

$P_i(\theta)$  : probability of respondent with ability  $\theta$  answer correct item- i

$P_i'(\theta)$  : derived function  $P_i(\theta)$  to  $\theta$

$Q_i(\theta)$  : probability of respondent with ability  $\theta$  answer wrong item- i

Test information function is the sum of all information functions of the item. The test information function will be high if the information function of the item is also high. Mathematically, the function of test information can be written as follows [21]:

$$I(\theta) = \sum_{i=1}^n I_i(\theta) \quad (6)$$

Item parameter values and the ability of the respondent is the result of estimation, so the truth is probability and cannot be separated from the measurement error. Standard error of measurement (SEM) in the item response theory is closely related to the information function. Information function and SEM has a correlation which is inversely proportional to the square, the greater the information function the smaller the SEM or otherwise [22]. The SEM equation can be written as follows:

$$SEM(\theta) = \sqrt{\frac{1}{I(\theta)}} \quad (7)$$

Information :

$I(\theta)$  : Information function

$SEM(\theta)$  : SEM estimation value

$\theta$  : ability of respondent

### III. MATERIALS AND METHODS

#### A. Data

There were two sets of tests, namely set A and set B. Each test set consisted of 10 items with 2 items being the anchor item. Research subjects included High School students in Yogyakarta class XI MIA who were present at the time of testing. Test work involved 317 respondents with purposive sampling.

The data used in this research were the result of students' response to the test. Each item test had 3 categories of response, i.e. 1, 2, and 3. In order to reach a certain category-i, the respondent must go through the preceding category (i-1). For example, the respondent must go through the 2nd answer category to get to the 3rd answer category. The items for each test set of diagrams and mathematical representation

ability consisted of 2 parts i.e. (1) five items related to diagrams representation ability and (2) five items related to mathematical representation ability.

The data obtained were polytomous data, which would be analyzed based on item response theory with Partial Credit Model. Data analysis involved: (1) goodness of fit with PCM, (2) estimating the item difficulty level, (3) estimating the respondent ability, and (4) total information function with standard error measurement (SEM).

#### B. Method

This research was a kind of quantitative descriptive research with polytomous data which came from the result of student's answer of the test. This quantitative descriptive study was conducted to find out the test quality of diagrams and mathematical representation ability of High School students in Physics.

Stages were performed to determine the test quality for diagrams and mathematical representation ability: (1) preparation of data in the form of input respondent answer (response) and (2) data analysis based on item response theory with Partial Credit Model.

Analysis of data to be performed involved: (1) calculating item parameters i.e. item difficulty level, (2) examining the fit of item to the PCM, (3) calculating the respondent's parameters (ability) and (4) obtaining the information function.

### IV. RESULTS AND DISCUSSION

#### A. Results

##### 1) Goodness of fit and parameter estimation

The goodness of fit test was performed to find out the fit of diagrams and mathematical representation test on partial credit model (PCM). Goodness of fit test on PCM could be seen from INFIT MNSQ and standard deviation presented in table 1. Table 1 showed that the test instrument for diagrams and mathematical representation ability of High School students in Physics fit with PCM. Analysis result showed that 18 items including 2 anchor items (item 9 and 10) from all set test had INFIT MNSQ value between 0.77 and 1.30, meaning that all items fit with model PCM presented in Figure 1.

TABLE 1. FIT STATISTIC WITH PCM MODEL ON PROBABILITY LEVEL = 0,5

Test Parameter	Item Estimates	Case Estimates
INFIT MNSQ	0.99 ± 0.13	0.97 ± 0.44



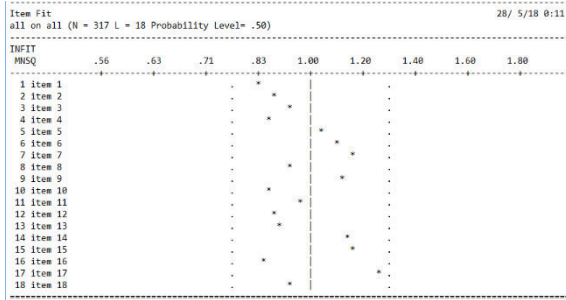


Fig. 1. Item Fit with PCM Model

This analysis uses the Partial Credit Model (PCM) which is the development of the 1-PL model. The result of partial credit model analysis involves one item parameter that is the difficulty level of item (b). The results of the items difficulty level on partial credit model are shown in Figure 2. Figure 2 suggests that the difficulty level of items in test of diagrams and mathematical representation ability is 0.072 with the range of -1.023 to +1.785.

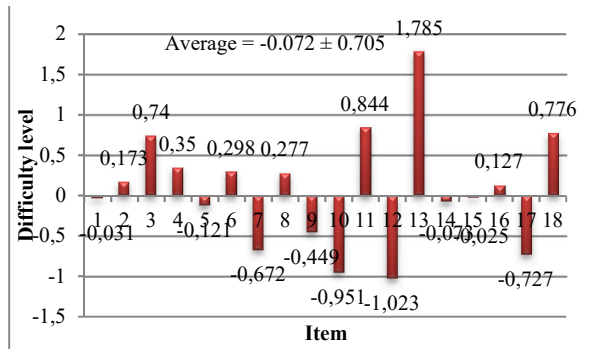


Fig. 2. Items difficulty level

The test items analysis did not only reveal the parameter estimation of the items only, but also the estimation of the parameters of the respondent's ability to do the test. The result of parameter estimation and distribution of respondent's ability in partial credit model are shown in Figure 3. Figure 3 shows that the distribution of ability ranges between -1.9678 and +2.0384 in logit scale between -4 and +4. The average ability of respondents who perform the test of diagrams and mathematical representation ability is 0.000.

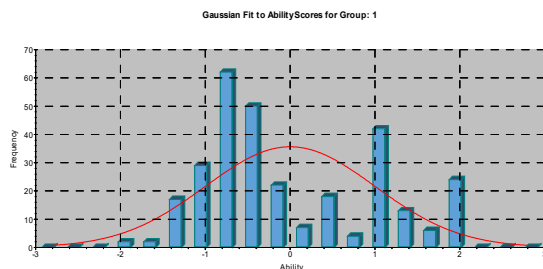


Fig. 3. Value distribution of diagrams and mathematical representation ability

2) Goodness of fit and paramater estimation

Test information function and standard error measurement (SEM) can be known through analysis of the Parscale program. The result of analysis is the

graph of relationship of test information function and SEM can be seen in figure 4. Figure 4 shows that the test of the ability of diagrams and mathematical representations has information function of 4.5 and SEM of ± 0.25. This diagrams and mathematical representation test is suitable for respondents who have ability value  $-1.85 \leq \theta \leq +1.9$ .

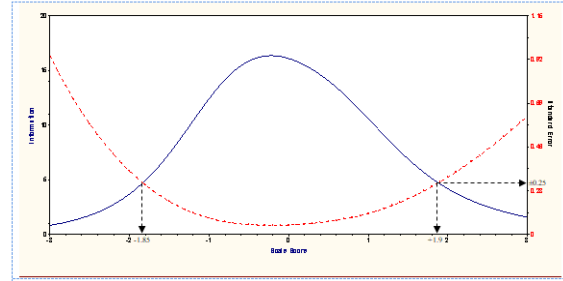


Fig. 4. Curve of information function test and SEM on PCM

B. Discussion

The total test instrument is 18 items divided into two sets of tests, namely A test set and test set B. Each set consists of 10 items with 2 items anchor items. The number of anchor items has met the required minimum number of 20% of the total test items each representing aspects of the diagram and mathematical representation ability [28]. The advantage obtained with such a design, among others, is that cheating in working out the problem can be minimized. The test instrument was tested to 317 High School students in Yogyakarta by purposive sampling. The number has met the number of respondents on the item response theory (IRT) that is 200-1000 respondents [29]. Previously, the test instrument has been reviewed based on aspects of content, construction and language by expert judgment involving expert professors, practitioners (Physics teacher), and lecturers. Then, item analysis is conducted to know more about the quality of test instrument.

Analysis of test instrument for diagrams and mathematical representation ability was done based on items response theory using Partial Credit Model (PCM). Based on PCM, the quality test for diagrams and mathematical representation ability can be determined. The test quality is reviewed based on goodness of fit on PCM, items difficulty level, respondent / case ability, and information function.

Analysis result of an overall item and case fit with PCM based on INFIT MNSQ value for items and case. Based on Table 1, that INFIT MNSQ value for items is  $0.99 \pm 0.13$ , and INFIT MNSQ value for case is  $0.97 \pm 0.44$ . This result shows that the overall item and case fit with PCM, because the INFIT MNSQ value is close to 1.0 and standard deviation is close to 0.0 [18]. Based on Figure 1, the result of 18 items fit analysis with PCM reveals INFIT MNSQ value in the range of 0.83 to 1.25. This result shows that all items are declared fit with PCM, because INFIT MNSQ value is in the range 0.77 to 1.30 [18,31]. Based on analysis result in Table 1 and Figure 1, all the test

items of the diagrams and mathematical representation ability are feasible to be analyzed using PCM.

Analysis using PCM involves one item parameter that is item difficulty level. Based on Figure 2, the value of item difficulty level test of diagram and mathematical representation ability is in the range of -1.023 to +1.785. This result shows that items of test for diagram and mathematical representation ability were categorized as good, because the difficulty level ranged from -2.0 to +2.0 [22,30]. Item difficulty level of -1.023 is located on item 12. Item 12 is item that describes the diagram representation. Item difficulty level of +1.785 is located on item 13. Item 13 is item that describes the diagram representation. Anchor items are located on items 9 and 10, while the difficulty level for each anchor item is -0.449 and -0.951. Average difficulty level was  $-0.072 \pm 0.705$ , which can be categorized as good [30].

Next is the result of the parameter estimation ability of the diagrams and mathematical representation of High School students in Physics, in which 317 respondents participated. The x-axis in Figure 3 represents the items being analyzed, while the y-axis represents the items' difficulty level. Figure 3 shows that the respondents' ability score distribution ranges from -1.9678 to +2.0384, in logit scale from -4 to +4, with the average ability of  $0.000 \pm 1$ . Based on average ability of diagrams and mathematical representation, the respondents' ability in diagrams and mathematical representation is medium [31].

Based on analysis result on Figure 4, the curve of information function on PCM was obtained. The x-axis in figure 4 represents the scale score or ability, while the y-axis represents information value and standard error measurement (SEM). Based on Figure 4, the test for diagram and mathematical representation ability has an information function of 4.5. In addition, the test is more appropriate for measuring respondents who have the ability between -1.85 and +1.9 with SEM value of  $\pm 0.25$ . This is because the instrument test can provide high information on the students' diagrams and mathematical representation ability and level of measurement error if tested on the respondents who have the ability between -1.85 and +1.9 [31,32]. In conclusion, the instrument test for the diagrams and mathematical representation ability can be used to measure students' diagrams and mathematical representation abilities in High School Physics material according to partial credit model based on three categories of polytomous data [31].

## V. CONCLUSION

The results of items analysis test of the ability of diagrams and mathematical representation based on item response theory using partial credit model in getting test quality include: (1) Item test instruments fit with partial credit model (PCM); (2) item difficulty levels are in the range of -1.023 to +1.785 which can be categorized as good; (3) the ability of

respondents who do the test is in the range of -1.9678 to +2.0384; and (4) this test is more appropriate for measuring respondents' ability between the values of  $-1,85 \leq \theta \leq +1,9$ .

Based on the results of these analyses, the instrument tests of the ability of diagrams and mathematical representation have good items that can be used to measure the ability of students' diagrams and mathematical representation on High School Physics materials.

PCM model is a fairly good parameter estimation method based on item response theory with polytomous data. Therefore, it is suggested that future studies employ different type of items or use more respondent data.

## ACKNOWLEDGMENT

The authors thanks the Directorate of Research and Community Service for the 9th Higher Education Institution in 2017 2017 through the Postgraduate Research Team for this work.

## REFERENCES

- [1] Buxton A and Provenzo F 2011 *Teaching Science Elementary & Middle School a Cognitive and Agricultural Approach* (Thousand Oaks: Sage Publications)
- [2] Wospakrik J 1993 *Dasar-dasar Matematika untuk Fisika* (Bandung: ITB Press)
- [3] Mundilarto 2002 *Kapita Selektta Pendidikan Fisika* (Yogyakarta : FMIPA Universitas Negeri Yogyakarta)
- [4] Krawec J 2010 *Journal of Learning Disabilities* **47** 103
- [5] Cock De M 2012 *Physical Review Special Topics – Physics Education Research* **8** 1-15
- [6] Savinen A, Makynen A, Nieminen P and Viiri J 2013 *Physical Review Special Topics – Physics Education Research* **9** 1-11
- [7] Edens K and Potter E 2008 *School Science and Mathematics* **108** 184-196
- [8] Romay E 2004 *Quantum* **10** 1-10
- [9] Redfors A, Hannson L, Hannson O and Juter K 2013 *Learning Science : Cognitive, Affective, and Social Aspects* 376-383
- [10] Bing T J and Redish E F 2013 *In the Physics Education Research Conference* vol 883 (Newyork: AIP Publishing) p 26
- [11] Kohl P B, Rosengrant D and Finkelstein D N 2007 *Physical Review Special Topics – Physics Education Research* **3** 1-10
- [12] Mattson M 2004 *The Physics Teacher* **42** 398-399
- [13] Albe V, Venturini P and Lascours J 2014 *Journal of Science Education and Technology* **10** 197-203
- [14] Widoyoko E P 2014 *Hasil Pembelajaran di Sekolah* (Yogyakarta: Pustaka Pelajar)
- [15] Rahayu Farida 2017 Pengembangan tes creative thinking skills menggunakan computerized adaptive test (CAT) pada mata pelajaran fisika SMA kelas X (Yogyakarta: Department of Physics, Yogyakarta State University) thesis
- [16] Purwanto N 2010 *Prinsip-prinsip dan Teknik Evaluasi Pengajaran* (Yogyakarta: Remaja Rosdakarya)
- [17] Hambleton R K. and Swaminathan H 1985 *Item Response Theory Principles and Applications* (Berlin: Springer Science+Business Media)
- [18] Adams R J and Kho S T 1996 *Acer Quest Version 2.1* (Camberwell, Victoria: The Australian Council for Educational Research)

- [19] Isgiyanto A 2014 *In The Prosiding Seminar Nasional Penelitian dan Penerapan MIPA* (Yogyakarta: Yogyakarta State University) p 43
- [20] Keeves J P and Masters G N 1999 *Advances in Measurement in Educational Research and Assessment* (New York: Springer-Verlag)
- [21] Meyer J P and Zhu S 2013 *Research & Practice in Assessment* **8** 26-39
- [22] Hambleton R K, Swaminathan H and Rogers H J 1991 *Fundamental of Item Response Theory* (Thousand Oaks: Sage Publication)
- [23] Arikunto 2007 *Dasar-dasar Evaluasi Pendidikan* (Yogyakarta: Bumi Aksara)
- [24] Widhiarso W 2010 *Model Politomi dalam Teori Respon Butir* (UGM : Fakultas Psikologi UGM)
- [25] Embretson S and Reise S 2000 *Items Response Theory for Physlogist* (Mahwah USA: Lawrence Erlbaum Associates)
- [26] Engelhard G 2005 *Encyclopedia of Statistics in Behavioral Science* **2** 995-1003
- [27] Lord M L 1980 *Application of Item Response Theory to Practical Testing Problems* (Mahwah USA: Lawrence Erlbaum Associates Publisher)
- [28] Kolen M J and Brennan R L 2014 *Test Equating, Scaling, and Linking: Methods and Practice* (New York: Springer Verlag Inc)
- [29] H S Seon 2009 *Pract. Assessment Res. Eval* **14** 1-8
- [30] Supahar 2014 *Proceeding of International Conference On Research, Implementation and Education of Mathematics and Science* vol 1 (Yogyakarta: Universitas Negeri Yogyakarta) p 137
- [31] Supahar and Prasety Z K 2015 *Jurnal Penelitian dan Evaluasi Pendidikan* **19** 97-108
- [32] Dwi K D 2015 *In Prosiding Seminar Nasional Matematika dan Pendidikan Matematika* (Surakarta: Universitas Muhammadiyah Surakarta)