

The Application Of Gpcm On Mmc Test As A Fair Alternative Assessment Model In Physics Learning

Edi Istiyono

Physics Education Department, Mathematics dan Natural Science Faculty,
Yogyakarta State University, Yogyakarta, Indonesia
Email: edi_istiyono @uny.ac.id, edi_istiyono_uny@yahoo.co.id

Abstract—This paper aims to evaluate: (1) feasibility scoring of the modified multiple-choice test (MMC Test) in polytomous, (2) conformity of generalized partial credit model (GPCM) as scoring model of the Physics test, and (3) validity of the MMC Test scored generalized partial credit model as a fair assessment model in Physics learning. This paper describes the advantages of the modified multiple-choice test, the application of GPCM on scoring modified multiple-choice test, and the conformity of scoring of GPCM on the modified multiple-choice Physics test. Therefore, it can be concluded that: (1) the modified multiple-choice can be scored in polytomous, (2) generalized partial credit model is a conformable scoring model of the Physics test, and (3) application of the GPCM on the modified multiple-choice test is a fair alternative assessment model in Physics learning.

Keywords: GPCM, modified multiple-choice test (MMC Test), assessment models, and Physics learning

I. INTRODUCTION

Physics lesson conducted according to Standard Process. Standard Process is a criterion regarding the implementation of learning in the educational unit to achieve Graduates Competency Standards [1]. On Curriculum 2013 the task of a teacher is to make Learning Implementation Plan (RPP) and maximize the learning process. Learning to use a scientific approach to learning is the learning activities to adopt measures scientists in building knowledge consists of observing, ask, try, reason and communicate.

To monitor the process, progress, and improvement of learning outcomes of students on an ongoing basis, the necessary assessments. Educational assessment is the process of collecting and processing information to determine attainment of learning outcomes of students. Implementation of the assessment refers to the Education Standards Assessment criteria regarding the mechanisms, procedures, and assessment instruments learning achievement of students [2]. Rate on the Curriculum 2013 is done in the form of authentic assessment that an assessment instrument that assesses the start of the input, process and results (outputs) of learning includes attitudes, knowledge and skills. Assessment technique used relevant by the scientific learning process, being able to assess the ability of learners in the learning process and results. In addition, the assessment is an activity of collecting individual data that result describe its characteristics [3]. Thus, assessment of learning achievement in physics is an assessment of the results of the learning process of physics which is a number that describes the characteristics of individual learners.

Assessment should: (1) adopting a form similar to the type of PISA and TIMSS questions to encourage learning process contribute to the improvement of science literacy of students and at the same time explore scientific thinking skills, critical, creative, and innovative; (2) emphasize the mastery of high and low level concept with variations of assessment (multiple choice, multiple choice reasoned, the description is limited); and (3) introduce the type of questions that tested both nationally and internationally to students and science teachers [4]. Thus educators need to make an assessment using the model and type of questions varies. To monitor the process, progress, and improvement of learning

achievement of students on an ongoing basis, the necessary assessments. Educational assessment is the process of collecting and processing information to determine attainment of learning achievement of students. Implementation of the assessment refers to the Educational Assessment Standards criteria regarding the mechanisms, procedures, and assessment instruments learning achievement of students [2]. Rate on the Curriculum 2013 is done in the form of authentic assessment that an assessment instrument that assesses the start of the input, process and results (outputs) of learning includes attitudes, knowledge and skills. Assessment technique used relevant by the scientific learning process, being able to assess the ability of learners in the learning process and results. In addition, the assessment is an activity of collecting individual data that result describe its characteristics [3]. Thus, assessment of learning achievement in physics is an assessment of the results of the learning process of physics which is a number that describes the characteristics of individual learners.

The test consists of test items. Haladyna [5] says: "A test item is an instruction or question that requires a student response and a rule for scoring the response". Menurut this definition, that the test items in the form of a command or a question that requires a response from learners and require a response menskor rules for such learners. Based on the understanding that the test is a test instrument that provides stimulus in the form of a command or a question that requires a response from the test participants. The response generated by the test participants stated in a score that is easy to interpret.

Assessment in education using two kinds of measurement theory, namely: classical measurement theory and modern measurement theory or item response theory (IRT). Classical test theory (CTT) is also called the True-Score Classical Theory, Classical Test Theory is named for the elements of this theory has been developed and applied for a long time, but still survive today [6]. According to the classical theory of measurement of scoring the test results usually done partially based on the steps that must be taken to correct an answer items. Scoring is done every step and score each item participant adds a score obtained by the students of each step, and the ability estimated by raw scores. Model scoring like this is not necessarily right, because the level of difficulty of each step is not taken into account.

One of the modern measurement theory is GPCM. GPCM is the development of PCM, the PCM discriminant items constant or 1 while the value GPCM discriminant vary. PCM is also appropriate to analyze the response to the measurement of critical thinking and conceptual understanding in science [7]. PCM was developed to analyze the test items that require several steps to resolve. GPCM can be applied to tests done with the steps and clear by the testee. Physics achievement test is a test that is done by exact steps. Therefore, GPCM is expected to be applied properly.

The fact that multiple choice tests are more widely used than other forms of testing. This is because the multiple-choice test used to have advantages, among others: (1) the material being tested can cover most of the learning materials, (2) the students' answers can be corrected easily and quickly, and (3) the answers to each question is certainly true or wrong, so that an objective assessment [8]. Although there are also drawbacks to this test, namely: (1) the possibility of learners to guess the answer is still quite large and (2) the thinking of students can not be seen with the real [8]. In addition, the fact that the scoring results of a multiple choice test with dichotomous models, meaning that if the item response is correct was given a score of 1 and if the response is wrong was given a score of 0. Teachers do not use polytomus scoring models that would be more equitable because it considers the item response measures. With this dichotomous scoring models have yet to appreciate the steps of problem solving, because with different error rate to get the same score of 0. As dichotomous scoring models is certainly less fair.

Based on the description above, the need to study the model instrument in the form of a multiple choice test modified by the model scoring with GPCM. The main issue to be raised in this study were (1) whether the test model of multiple-choice modified can be scored in polytomus four categories according to generalized partial credit models and (2) whether the application of GPCM in multiple-choice modified as an alternative model in the assessment of learning physics effective and fair.

In accordance with the problem to be solved, then the purpose of this paper is to evaluate (1) the feasibility scoring of the modified multiple-choice test (MMC Test) in polytomous, (2) conformity of generalized partial credit models as scoring models of the physics test, and (3) the validity of the modified multiple-choice test scored GPCM as a fair assessment models in physics learning.

II. DISCUSSION

1. The Feasibility Scoring of The Modified Multiple-choice Test (MMC Test) in Polytomous

Structure of items'MMC test consists of stem, the option of stem, reason, and the option of reason. The number of options and choices stem reason each of the five, as Figure 1.

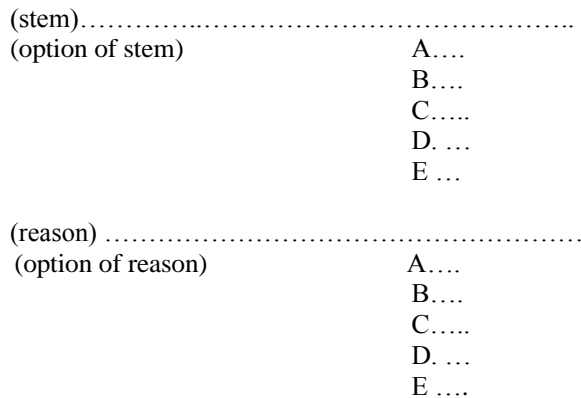


Figure 1. Diagram of items'MMC Test Structure

MMC tests that have these structures, then it can be scored with polytomous scale. The number of selected categories depend on the scoring guidelines. One of the scoring guidelines that can be selected provisions of each category, as follows.

- Category-1: if learners wrong in writing the concepts used and the results are wrong. It is indicated by the students answer questions one and also one of the reasons
- Category-2: if learners wrong in writing the concepts used but the results can be correct. It is indicated by the students answering questions correctly and grounds wrong
- Category -3: if learners correct in writing concepts used but the end result was wrong. It is indicated by learners answer the question wrong and right reasons
- Category-4: if learners correct in writing concepts used and the results are correct. It is indicated by the students answering questions correctly and reason also true

Thus, scaling tests MMC was created in polytomus with 4 categories, ie 1, 2, 3, and 4. Characteristics of MMC item test was scored with polytomous 4 categories can be described as the ICC. Figure 2 illustrates the ICC one item of MMC test.

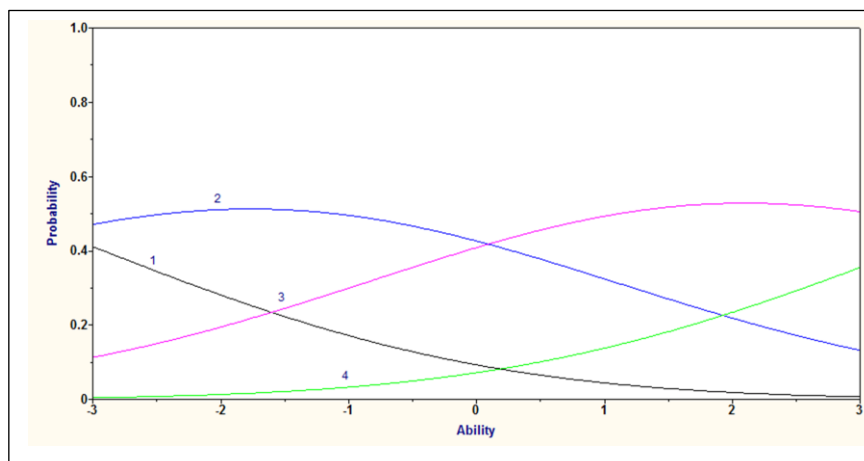


Figure 2. ICC's an item

2. Conformity of GPCM as Scoring Models of The Physics Test

GPCM is the development of PCM. Items' discriminant of The PCM is held exact or 1 while the value of items' discriminant in GPCM vary. PCM is developed to analyze the test items that require several steps to resolve. GPCM can be applied to tests done with the exact steps.

Assessing of tests are based on the steps that can be completed examinees. Although only completed the initial phase alone, examinees are already getting value. The highest value of course obtained when the examinee has completed all phases of the exam questions in that clause. The assessment procedure is virtually identical to how individuals respond to the item in the scale of psychology. For example, an item that provides four categories of responses from 'never', 'rarely', 'often', and 'always' is analogous to the step of completion of such things stated Figure 3. Complete a matter only until the first phase is analogous to the category of 'never' while when it came to the final step, analogous to the category of 'always'. This assumption was later developed into PCM. When it is assumed that a point following the pattern of partial credit the higher the ability of individuals are expected to have a higher score than the individual who has the ability to lower [9]. According to Wright & Masters, PCM is also appropriate to analyze the response to the measurement of critical thinking and conceptual understanding in science [7]. Physics achievement test is a test that is done by exact steps. Thus, the learning achievement tests in physics is perfect score by GPCM.

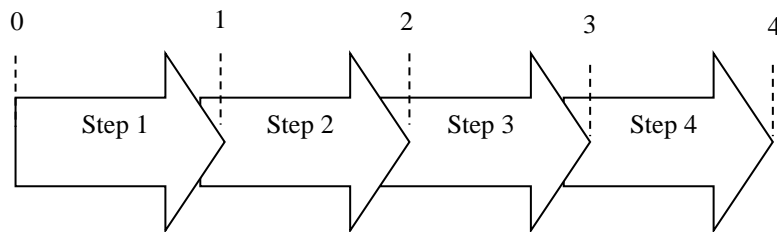


Figure 3. The scale interpreting is as the step of completion of the item

Figure 4 describe that the test is appropriate for measuring students' physics learning achievement that value from -0.4 to 3.2. This is supported by [10] that modified multiple-choice tests can be to measure students' physics learning achievement. GPCM is scoring model has accurate estimates on the item parameter [11]. Based on the graphs total information function and SEM that the modified multiple-choice tests can be used to measure students' physics learning achievement.

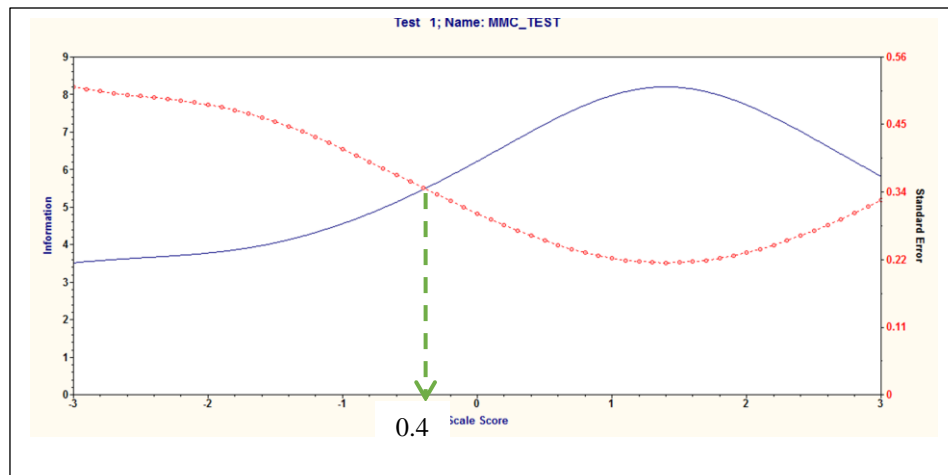


Figure 4 The MMC Test Total Information Function and SEM

3. The Validity of The MMC Test Scored GPCM as a Fair Assessment Models in Physics Learning

The difference classical from modern theory in educational assessment can be illustrated with five students A, B, C, D, and E taking the test as much as 5 items with five response type. To simplify the wrong item was given a score of 0 and a maximum of four given the correct score of 1. Discriminant and difficulty index 5 items found in Table 1. Response of five students is one item wrong and four items

correct. If they were scored in classical and modern give different abilities. If classically abilities the five students at 8 all. It is different if it were scored in a modern (GPCM), five students have different abilities with greater sequence namely A, B, C, D, and E. Students' abilities higher if the student is able to do the item whose difficulty index higher.

Table 1. Discriminant and difficulty index 5 items

Item parameter	Item no-				
	1	2	3	4	5
Discriminant	2	1.5	1	0.5	0
Difficulty	-2	-1	0	1	2

Table 2. Students' abilities are scored by CTT and GPCM

Student's ability	Student's Response				
	A 01111	B 10111	C 11011	D 11101	E 11110
GPCM	-1	-0.5	0	1	4
CCT	8	8	8	8	8

If the score of 1 represents a score of 1, 2, 3, and 4, the five students' abilities would be significantly different. It is supported by Kortemeyer, the technique is increasingly used in Physics Education, for example to examine the validity of concept test [12]. Thus, in a modern scoring GPCM is fairer than classical scoring with a dichotomous scale.

III. CONCLUSION AND SUGGESTION

Based on the analysis, the conclusions that can be drawn are as follows:

1. The modified multiple-choice can be scored in polytomous
2. Generalized partial credit model is a conformable scoring model of the physics test
3. Application of the GPCM on the modified multiple-choice test is a fair alternative assessment model in physics learning.

IV. REFERENCES

- [1] Mendikbud. Permendikbud no. 65 tahun 2013 tentang Standar Proses. Jakarta. 2013
- [2] Mendikbud. Permendikbud no. 66 tahun 2013 tentang Standar Penilaian Pendidikan.
- [3] Mardapi, Djemari Pengukuran, penilaian, dan evaluasi pendidikan . Yogyakarta: Nuha Litera. 2012
- [4] Mendiknas. Peraturan Menteri Pendidikan Nasional No 20 Tahun 2007 tentang Standar Penilaian
- [5] Haladyna, T. M. Devoping and Validating Multiple Choise Test Items. New Jersey: Lawrence Erlbaum Associates, Inc. 2004.
- [6] Suryabrata, S. Pengembangan alat ukur psikologis. Yogyakarta: Andi Offset 2002.
- [7] Van der Linden, W. J & Hambleton, R. K.. Handbook of Modern Item Response Theory . New York: Springer-Verlag New York, Inc 1997
- [8] Sudjana, N. Penilaian Hasil Blajar Mengajar. Bandung: PT Remaja Rosdakarya. 1990.
- [9] Widhiarso, W. Model politomi dalam teori respons butir. Yogyakarta: Psikologi UGM. 2010
- [10] Nitko, A.J & Brookhart, S. M. Educational assessment of students. (6th ed). Boston: Pearson Education, Inc. 2011.
- [11] Abadyo & Bastari. Estimation of ability and item parameter in mathematics testing by using the combination of 3PLm/GLM and MCM/GPCM Scoring Model. Research and Evaluation in Education Journal. Vol 1 Number 1 June 2015 p 55-72.
- [12] Kortemeyer, G. Extending Item Response Theory to Online Homework. Physics-ed.ph January 9, 2014, <http://arxiv.org/pdf/1401.1702.pdf>

