Effect of Total Choice Probability Multiple Choice Test Answers Kind of Quality Function Infomasi Three Point Logistics Model Parameters Compiled by Item Response Theory

Budi Susetyo

Universitas Pendidikan Indonesia. Indonesia

Abstract: ***This study examines how the manufacture of measuring instruments that do not have dependency with the ability of the participants in the test, the form of a multiple choice test with three and four answer choices based Three Parameter Logistic Model (L3P). This study aimed at obtaining the form of multiple-choice test based on the function of information quality grain, so it can be used by the compilers of the test including the classroom teacher or field of study and the teaching evaluation. The method used is the application and manufacture of multiple-choice test, three and four answer options based on a model L3P and conducted field trials. The study was conducted in SMP on teaching science with 625 respondents. Assays used compiled L3P models draw upon multiple choice objective test items with three answer choices and four answer choices , and the obtained results of the study showed no effect or difference in the value of the information function grains in both forms of the test . Therefore, educators in the construction of achievement test based on a model L3P no need to question the amount of possible answers (options) on multiple choice tests.***

Keywords: L3P IRT models, the number of answer choices, the value of information items

1. Introduction

The curriculum is a program guide learning activities that should be followed by all education providers. To determine the achievement of learning objectives should be measured with a gauge of quality. Measuring devices that can be used to measure the ability of learners, namely achievement test. Tests commonly used in educational measurement, there are two types; ie form of objective tests and test forms. The division of this test is based on a point scoring exercise test (Saefudin, A, 1996:72). Based on how participants answered the test items are divided into two parts (Popham, W,1981:251) the test items that have been provided the answer, so just choose the answer the test participants (selected response test items) and test items not provided an answer, the test participants need to make their own answer (constructed response test items). In fact many schools do the measurement of learning outcomes using this form of objective tests, especially multiple choice. The use of multiple-choice objective test forms almost done in all subjects and all levels of education. Different uses of objective type tests at every level of education lies in the complexity and number of multiple choice answer options are customizable. Qualification elementary, junior high and high number of answer choices that are often used for no more than four. As for college admission as SBMPTN or Self Exam objective test forms used in general are five answer choices.

In preparing the measuring instrument, there are two types of theory test can be used to measure learning outcomes, namely the classical test theory and the theory of the response item. The division is based on the way the preparation of test items, whereas in the test of physical form there is no difference. Classical test theory exam is usually done on the exam formative and summative exams at school. In classical test theory test items created by the class teacher or subject teachers. Therefore, the test results are local, resulting in differences in the scores obtained on one place to another place. This distinction is possible because of the nature of classical test theory are dependent on the ability of the test taker. High-ability test participants doing the test grains with a low difficulty level, then the test items that do become easier. Instead participants are low-ability test work items with a high degree of difficulty, although with some difficulty trying to work on still wrong. Classical theory test is commonly used in schools as a measure of success in learning to use the form of multiple choice tests, but there is also a description of the test form. To overcome the disadvantage of classical test theory exam experts find another way then called with modern test theory or the theory of the response item (Item Response Theory/IRT). Responsiveness grains removes the dependency theory measuring instrument with the ability of the test taker. Test items which have a high difficulty level unchanged (invariant) and remained high undertaken by anyone, whether by high-ability learners or low-ability. Likewise, the level of difficulty of test items have low fixed low level of difficulty is done by those who are capable are capable of high or low. When examined two forms of the theory test, both theory and classical test theory grains have different responsiveness to the measurement of learning outcomes. Theory responsiveness grains provide better results because the test items no dependence with the ability of learners as occurred in the classical test theory exam or classical.

The use of the form of multiple-choice tests on the response theory item has not been a lot of studies on the user test in school or the policy holder in the evaluation of learning or educational measurement, therefore it is necessary to study in more depth. The study of the weaknesses and strengths of the form of multiple choice tests with a wide selection of answers or multiple choice form of modification has not

been done by the teacher or achievement test developers. The results of this study can provide a real picture of the effectiveness of each number of answer choices on multiple choice tests to measure learning outcomes in the cognitive aspects. While the use of multiple choice tests are generally widely used by teachers at all levels of education to measure learning outcomes, including national level exams. One of the weaknesses in the multiple-choice test that is difficult to select a comparable alternative answers that serve as detractors (EG Norman and Robert L. 1990:120).

The risk of error in the selection of the measuring instrument to measure learning outcomes at every level of education and the low accuracy of measurement is still often the case, so that the measurement results do not inform the real conditions the ability of learners.

In preparing multiple-choice test items, what is the best number of answer choices until now there is no definite answers. Sumadi S. (1987:53) says "things into consideration in determining the number of answer choices is the age of the test participants, elementary school students are not sufficiently able to deal with problems with five or more answer choices". The use of different number of answer choices on objective tests give different probabilities to give the correct answers. The magnitude of the opportunity in giving the correct answer is 1/N, where N is the number of answer choices provided. Item multiple-choice test has four answer options of different opportunities with three answer choices that is equal to one quarter (25 %) and third (33 %). Similarly to form an objective test that provides an answer to that do not provide an answer will have different opportunities to give correct answers. With an assessment of the number of answer choices on the objective test will be known form of multiple choice tests are more appropriate as a measurement of learning outcomes because of unknown function item information on a test developed by Item Response Theory.

In theory responsiveness grain, grain information has an important function, namely as a determinant in selecting test items. This is because the value of item information function is a picture of the relationship between parameters of grain to test the participants' ability parameters are fixed. Three parameter logistic model is one application of the theory of the response item. Three-parameter logistic model has the characteristic parameters and the ability to characterize the grain parameters consisting of grains difficulty level, guesses, and the power difference.

Differences chance to guess the correct answer due to the difference in the number of answer choices that affect the functioning of the information supplied grain, it is necessary to study through research, especially for the three answer choices and four answer options.

2. Research issues

A wide variety of tests that can be used to measure the ability of learners in the cognitive aspects of the learning activities. This study is limited to the issue of multiple choice tests form a three parameter logistic model (L3P) which is based on the theory of the response item (IRT). Shapes under study is an objective test of multiple choice with three answer choices and four answer options are all used to measure the ability of learners in science subjects. Based on the description of the background and limitation issues, the problems examined in this study is formulated as follows;

Are there differences influence the probability of a correct answer to the quality of item information function tests three and four multiple-choice tests on the three-parameter logistic models were developed based on Item Response Theory?.

3. Research Methodology

This section discusses the various things that the methodology used in the study. As for the things covered are;

3.1. Methods Research

The method used in the study is an experimental method to try a multiple choice objective test with four and three choices of answers in science subjects to students. This research is the experiment group participants test that uses a three- choice objective test answers. The control group of this study is that the test participants using an objective test form four answer options.

3.2. Place and Research Subjects

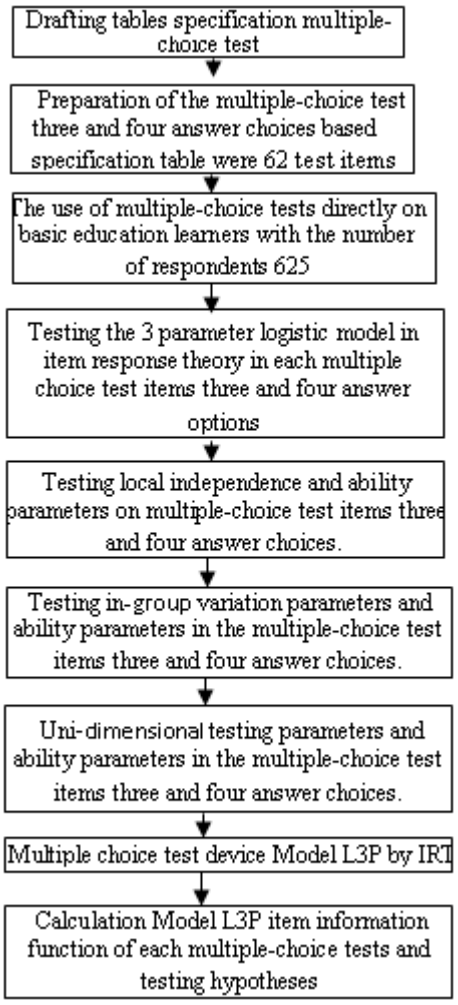
The study was conducted at the level of secondary education in the city of Bandung. The device tests that examined the response prepared by the theory of grain (IRT), therefore in the preparation of the test requires a fairly large number of respondents as belonging to the test standard. According to Mohammad Noer (1987:149) to create a standardized test required minimum number of respondents were "relatively stable minimum size is 200 respondents " . Respondents used in the study are far above the 200 respondents as a minimum. According Nanully, (Mohammad Noer, 1987:149) define the size of the respondent as much as 5 to 10 times the number of grains . Respondents in this study were learners to basic education amounted to 1250 students, with details of each group consisted of 625 respondents. Each group worked on three objective test answer options and a choice of four answers.

3.3. Instruments Research

The dependent variable in this study is the value of the information item multiple-choice test forms three and four answer choices answer choice of science subjects. The instrument used in this study is a multiple-choice test items objective science subjects. Prior research sample test items first analyzed to select the test items that meet the requirements of three parameter logistic model (L3P) on the response theory item (IRT). Test items that meet the requirements of the model L3P many as 42 points for each test form either three or four possible answers.

3.4. Research design

The design of the study briefly looks at the picture below;



3.5. Processing Data

Processing of the data in this study consisted of two phases, namely the testing requirements of the test in accordance with the model L3P on IRT and hypothesis testing.

1. Testing requirements L3P models include; model fit, uni-dimensional, invariance groups, and local independence.
2. Techniques of data analysis to test the hypothesis by comparing the value of the second test information item multiple-choice test forms three and four answer choices.

4. Results

From the results of the testing requirements of the model L3P, of 62 test items that meet the requirements of as many as 42 test items. Stages calculation starting point information item that satisfies the model followed by the calculation of the test participants' ability parameters, parameters grains consisting of level of difficulty, different power, and guesses, then either count item information functions. Based on the results of hypothesis testing on the value of the test is based on the number of information items obtained results an answer choice that item information function values form an objective test three possible answers to the test item information function value objective form four answer options do not give effect to the difference in the value of information items.

Linkages with the item information function test quality can be explained as follows; The theory of the response items used in the measurement gives meaning to the value of the test item information. The lower the value of item information functions within a set of tests, the higher the imprecision of measuring instruments are made with the results of tests measuring the ability of the participants. The calculation of the value of information items for the model parameters L3P3 consists of the participant's ability and item parameters (power difference, level of difficulty, and answers to the riddle). High value item information function, required power difference is quite high, the level of difficulty according to the participant's ability and guessing factor is close to zero. Grain reaches its peak value information if power height difference, the same level of difficulty with test participants ability, and answered by guessing no (zero). These three things must occur simultaneously, the power value of the height difference, but not accompanied by a match between the level of difficulty to the ability of the participants, and the guessing factor is high, it produces a low value item information.

Tests are in multiple-choice form provides the option to develop an answer, so just choose the test participants considered the most correct answers. This answer choice tests provide an opportunity for participants to guess the correct answer, especially for those who have low ability. Preparation of multiple choice tests can be performed with classical test theory and modern test theory (Item Response Theory). In the tests are arranged by item response theory, many options are available to answer multiple-choice tests affect the probability of answering correctly. The more the number of answer choices is getting smaller the probability of selecting the correct answers. Therefore item information function is related to the probability of selecting the correct answer on each item on. To achieve a high value item of information required to answer correctly the probability is high, does not depend on the number of answer choices provided , but it really depends on the height capability (ability) test participants. Choose the answer by guessing the answer choice test forms occur, if the test item has a difficulty level and a high level aspects of cognitive domains measured at high while the ability of participants below the level of difficulty of the test item. Participants chose the test by guessing the answer is sometimes also done by blinking their high ability, this is due to the presence of the test items that want to measure the subject matter that has not been taught or learned by the test participants. Therefore, using a form of objective test three possible answers or forms of objective tests of four answer choices do not give effect to the difference in the value of information items. This is not due to differences in the probability of the correct answer on a multiple choice test items by 1/N in choosing the correct answer. But the effect is the parameter tests the ability of participants to the level of difficulty of the grain problem. Participants high-ability test is not affected by the magnitude of the probability of answering correctly in accordance with the number of answer options provided in the form of multiple-choice tests. Factors that influence the answer to the riddle is the ability of the participants, the higher the ability of participants declined to answer with an educated guess (c). High-ability test takers can answer correctly on a particular item even if the item is not available answer choices or altered form of the test. In addition to the ability

of the participants are different power index (b), the lower the power difference, the higher the answer to the riddle (c) and consequently the higher guessed.

5. Conclusion

Based on the results of data processing and the discussion in the previous section can be deduced as follows; The value of information items on the modern theory exam 3 parameter logistic model item objective test three possible answers to the objective test item four answer options in junior high education level showed no effect on the difference in the quality of the test items. This suggests that differences in the number of answer choices objective test items form three answer choices and objective test items form four answer options, the value of the items of information on both forms of the multiple-choice objective test of the same quality or value of item information is not affected by the difference probability of the answer choices provided.

Reference

[1] Azwar, Saifuddin, *Tes Prestasi. Fungsi dan Pengembangan Pengukuran Prestasi.*Yogyakarta: Pustaka Pelajar. 1996.
[2] Azwar, Saifuddin, *Dasar-Dasar Psikometri*. Yogyakarta: Pustaka Pelajar. 2005.
[3] Crocker L dan Algina James.*Introduction To Classical And Modern Test Theory*. New Yoork: Holt, Rinehard And Winston, Inc. 1982
[4] Ebel, L. Robert, *Essentials of Educational Measurement.*New Jersey: Prentice-Hell. INC. 1079.
[5] Hambleton, Ronald K. Swaminathan, H. *Item Response Theory Prinnciples and Applications.*Boston: Kluwer – Nijhoff Publishers, 1985.
[6] Hambleton, Ronald K. Swaminathan, H Rogers, H Jane. *Fundamentals of Item Response Theory***.**London: Sage Publications, 1991.
[7] Hulin L. Charles, Drasgow P. dan Parsons K. Charles, *Item Response Theory, Aplication to Psychological Measurement.* (Illinois:Dow Jones Irwin. 1983.
[8] Naga, Dali S. *Pengantar Teori Skor pada Pengukuran Pendidikan*. Jakarta: Gunadarma. 1992.
[9] Naga, Dali S.*Teori Skor pada Pengukuran Mental* Jakarta: PT Nagarani 2012
[10] Noer, Mohamad. *Pengantar Teori Tes.*Jakarta: Depdikbud.1987.
[11] Popham, James W. *Modern Educational Measurement*. California: Printice-Hill Inc. 1981.
[12] Suryabrata, Sumadi. *Pengembangan Alat Ukur Psikologis*. Yogyakarta: Andi. 2005.
[13] Susetyo, Budi, *Menyusun Tes Hasil Belajar dengan Teori Ujian Kelas dan Teori Responsi Butir*. Bandung: Cakra. 2011.

Author Profile

Budi Susetyo is Doctoral Education in the Field Measurement of Jakarta State University. He is a lecturer in the university education in Indonesia. Of teaching subjects related to the manufacture of the test device or instrument for educational research, evaluation of learning, assessment, psychometric tests and construction of various programs of study and level of education