Analysis item test of math problem solving skills in junior high school

A Muzammil¹*, B Budiyono² and H Pratiwi³

¹ Pendidikan Matematika, Universitas Sebelas Maret, Jl. Ir. Sutami No.36 A, Kentingan, Jebres, Surakarta 57126, Indonesia
² Departemen Pendidikan Matematika, Universitas Sebelas Maret, Jl. Ir. Sutami No.36 A, Kentingan, Jebres, Surakarta 57126, Indonesia
³ Departemen Statistika, Universitas Sebelas Maret, Jl. Ir. Sutami No.36 A, Kentingan, Jebres, Surakarta 57126, Indonesia

*ahadmuzammil@student.uns.ac.id

Abstract. An instrument is one of tools has used to collect a data, whether or not a research instrument has needs to be seen characteristics of the instruments compiled. So this research has a purpose is to find out characteristics of instruments test for mathematical problem solving abilities have been compiled in algebraic material for class VII junior high school students. The research method used is a basic research method. Data analysis was done by qualitatively and quantitatively. Qualitative analysis is used to see characteristics of content validity, while quantitative analysis is done to see characteristics of the level of difficulty and distinguishing items, as well as the reliability coefficient of instrument test. The results has showed that all items had content validity, level of difficulty and distinguishing power were good, with high instrument reliability coefficients.

1. Introduction

Test is one of the tools to do a measurements, namely a tool to collect information on the characteristics of an object. This object can be in the form of students' abilities, attitudes, interests, and motivations [1]. In order to produce a good instrument test, there are several steps that must be passed. There are nine steps need to be taken in developing results test or learning achievements, namely compiling test specifications, writing test questions, reviewing test questions, conducting test tests, analysing items, improving tests, assembling tests, carrying out tests, and interpreting test results [2]. A good instrument test can be improve the quality of the assessment results is namely the profile students ability.

To knowing ability to solve mathematical problems, is a student must be faced with mathematical problems (mathematical questions). By facing math problems, students will try to solve problems using all the schemes that are inside themselves. The process for solving problems is called problem solving [3]. Problem solving had involves the interaction between the scheme (knowledge) possessed by students and the application process that was uses cognitive and affective factors in solving problems [4].

Completing mathematical problems are requires a systematic way so that the resolution process becomes easy and directed. Some expert opinions about the steps that can be used to solve mathematical problems, one of them is problem solving from Polya which presents four stages in solving problems is namely Understand the problem, Make a plan, Carry out our plan or implement the plan, Look back at the completed solution or re-examine the answer [5].
The mathematical problem solving ability test instrument besides being used to determine the profile of students’ ability to solve mathematical problems faced, can also be used as a means to train students’ ability to think systematically in solving mathematical problems. To solve mathematical problems, students are required to be able to think systematically [6].

Based on the above, the purpose of this research was to determine characteristics of the instruments test for mathematical problem solving abilities that had been compiled in algebraic material for class VII junior high school students. The characteristics of the instrument in this case are the validity of the instrument’s content, the level of difficulty, item differentiation and instrument reliability.

2. Method
The research method has used is basic research. The subjects has used in the research were class VIII A and VIII B of State Junior High School 2 Dukuhwaru in academic year 2018/2019. The data sources has used in this research are qualitative data in the form of test sheets for content validity and review sheets for items and quantitative data in the form of student answers.

The method has used is to collect data about mathematical problem solving skills was using test methods in the form of essays which amounted to 5 items / questions. Furthermore, this test is called the Mathematical Problem Solving Ability Test. Tests of mathematical problem solving abilities are made referring to material equations and linear variables. Scoring tests of mathematical problem solving abilities refer to indicators of problem solving from the pattern [7], while the guidelines or scoring rubrics for testing the ability to solve mathematical problems are presented in Table 1.

| Poly Stages                        | Score | Scoring Indicator                                                                 |
|-----------------------------------|-------|-----------------------------------------------------------------------------------|
| Understand the Problem            | 3     | Students are able to write (express) what is known and asked of the problems raised clearly |
|                                   | 2     | Students only write (express) what is known or what was asked                      |
|                                   | 1     | Students write down data / concepts / knowledge that are not relating to the problem being proposed so that students do not understand the problem posed |
|                                   | 0     | Students do not write anything down so students do not understand the meaning of the problem posed |
| Make The Problem Solving Plan     | 2     | Students write enough conditions and necessary conditions (formulas) from the problem proposed and use all information that has been collected |
|                                   | 1     | Students tell / write steps for solve the problem but not in sequentially          |
|                                   | 0     | Students do not tell / write steps for solve the problem                           |
| Carry out The Problem Solving Plan| 4     | Students carry out the plans that have been made, use the steps to solve the problem correctly, there is no procedure error, and no algorithm / calculation errors occur |
|                                   | 3     | Students carry out the plan that has been made, use the steps to solve the problem correctly, and there is no procedure error, but an algorithm / calculation error occurs |
|                                   | 2     | Students carry out the plan that has been made, but a procedure error occurs       |
|                                   | 1     | Students carry out the plan that has been made, but a procedure error and algorithm / calculation error occur |
| Look back at the completed solution| 0     | Students are unable to carry out the plans that have been made                     |
|                                   | 1     | Students re-examine the answers                                                   |
|                                   | 0     | Students do not look back the answers                                             |
The data analysis technique has used is the analysis of qualitative and quantitative data. Qualitative analysis is done through a study to find out the validity of the contents of the test instrument is namely the compatibility between the questions in the test with the indicators that have been prepared previously, that an instrument is said to be valid according to content validity if the instrument content is a representative sample of the overall contents measured [8]. While the quantitative analysis is done with the classical test theory approach the calculation of which is aided by Microsoft Excel programs. Several aspects were analysed quantitatively is namely the level of difficulty and distinguishing items, as well as the reliability coefficient of the test instrument.

3. Result and discussion

3.1. Validity

Content validity is a level of measurement that reflects the expected content domain [9]. The validity of the contents is important for cognitive tests such as tests of mathematical problem solving abilities. A score does not even reflect student learning outcomes if the instrument is not able to comprehensively measure what has been learned by students. To enhance content validity, before making the question items carried out several steps including (1) identifying the materials that have been given along with the instructional objectives, (2) making a grid of test questions to be tested, (3) compiling the test questions along with the answer key, and (4) review the test questions before printing or duplicating. While the indicators that are used as guidelines for measuring content validity in this research, they are (1) conformity with the test grid, (2) conformity with the research objectives, (3) the items are representative samples of a population or sub basic competencies, (4) The items are do not require other knowledge in answering them, and (5) questions have used language that is in accordance with the rules of Indonesian Language [8].

Validation is carried out by three people who are considered experts in the field of mathematics education are namely two mathematics education lecturers and one math teacher at junior high school level. The steps to assess whether the test instrument has high content validity or not will be done by expert judgment (an assessment carried out by experts). In this case the appraiser will be given a validation sheet by the developer to assess whether the grid has been made by test developers have shown that the grid classification represents the content (substance) to be measured. The next step, the assessor will assess that each item has been compiled to be suitable or relevant to the classification of the grid specified.

The results of expert validation has showed that the majority of validators stated that the items were in accordance with the test grid, items were in accordance with the research objectives, items were part of the equation material and one variable linear inequality, items did not require knowledge other than mathematical knowledge, and items the matter of using good and correct Indonesian. These results are in accordance with predetermined indicators is namely the items stated to have content validity if at least 50% of all validators or assessors agree with all indicators that are used as criteria in validation.

3.2. Level of difficulty

The level of difficulty of the item is one indicator that can show the quality of the item. The level of difficulty of the items can be calculated by considering the proportion of the average score of all students compared to the maximum number of scores that can be achieved by all students. The difficulty level index for the description test is formulated as follows [8].

\[
P = \frac{\bar{S}}{S_{\text{maks}}}
\]

With \(P\) is the difficulty level index of a question, \(\bar{S}\) is the mean for the item score, and \(S_{\text{maks}}\) is the maximum score for the item.

Based on the formula, the range of the difficulty index value is \(0 \leq P \leq 1\). With the interpretation of the difficulty index, each item is presented in Table 2.
Table 2. Interpretation of difficulty index questions.

| Index Value (P) | Criteria |
|-----------------|----------|
| 0.0 ≤ P < 0.3   | Difficult|
| 0.3 ≤ P ≤ 0.7   | Average  |
| P > 0.7         | Easy     |

Table 2 was shows that the test items for mathematical problem solving ability to be said to have a good index score if 0.3 ≤ P ≤ 0.7. The results of the calculation of the level of difficulty obtained that on each item in a row has obtained at 0.500; 0.384; 0.406; 0.331; and 0.559. Based on these results it can be concluded that all test items of mathematical problem solving abilities have a good index of difficulty, because it is said that the index of difficulty is good if 0.3 ≤ P ≤ 0.7 with the criteria of moderate difficulty level.

3.3. Differential power
In determining the right item in a research, it must be known that the question has a good distinguishing ability for different students, different times and different places. In other words the question is the ability of a question to distinguish between high-ability students with low-ability students based on certain criteria [10]. The differentiation power of each item is seen from correlation between the scores of the items with the total score. If the instrument is a test of mathematical problem solving ability, then items with a high consistency index can distinguish between smart and less intelligent students. To find out the item differentiation power product moment correlation can be used from Karl Pearson [8].

\[ D = r_{pbis} = \frac{n\Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{(n\Sigma X^2 - (\Sigma X)^2)(n\Sigma Y^2 - (\Sigma Y)^2)}} \]

With D is the discrimination power, \( n \) is the number of subjects to the test (instrument), \( X \) is the item score, and \( Y \) is the total score.

The smaller the distinguishing index indicates that the items are getting worse to distinguish high-ability students from low-ability students, and vice versa if the differentiation index is getting bigger, the items are better for distinguishing high-ability students from low-ability students. The classification and interpretation of the power of different items for each interval can be presented in Table 3.

Table 3. Classification and interpretation of different grain power for each interval.

| Item Difference Interval | Classification | Interpretation |
|--------------------------|----------------|---------------|
| -1.00 ≤ DB < 0.20       | Bad            | Bad differentiation |
| 0.20 ≤ DB < 0.40        | Excellent      | Has enough differentiation |
| 0.40 ≤ DB < 0.70        | Good           | Has a good distinguishing power |
| 0.70 ≤ DB ≤ 1.00        | Very good      | Has a very good distinguishing power |

Table 3 was shows that the items has used are items have a distinguishing index equal to or more than 0.30 or at least in a satisfactory classification. The results of the calculation of differentiation are obtained that for each item in a row obtained at 0.878; 0.884; 0.791; 0.871; and 0.711. Based on these results it can be concluded that all test items of mathematical problem solving abilities have a distinguishing index of more than 0.30. So that all items have a very good distinguishing ability to distinguish between high-ability students and low-ability students.

3.4. Reliability of the test
Reliability is often referred to as reliability, stability, consistency, meaning an instrument has reliability when the instrument is used to measure repeatedly the results are relatively the same. An instrument is called reliable if the measurement results with the instrument are the same if the measurement is done on the same person at different times or in different people (but have the same condition) at the same time or at different times. To test the reliability of an instrument, one of them can use the Cronbach Alpha formula as follows [8].
\[ r_{11} = \left( \frac{n}{n-1} \right) \left( 1 - \frac{\sum s_i^2}{s_t^2} \right) \]

with \( r_{11} \) is the reliability coefficient of the test, \( n \) is the number of items issued in the test, \( s_i^2 \) is the item variance \( i, i = 1,2, ..., n \), and \( s_t^2 \) is the variance of the total score.

There are several opinions about how to interpret the reliability coefficient. In general, people said that an instrument is called reliable if the reliability coefficient is equal to or more than 0.70.

Based on the results of the calculation of reliability against 5 items that meet the requirements of validity, the level of difficulty and distinguishing power obtained that the reliability index is 0.885. So that the problem solving ability test has met the established criteria, namely the test is said to be reliable if the reliability coefficient is equal to or more than 0.70.

### 4. Conclusion

The results of analysis of content validity, level of difficulty, differentiation and reliability test of mathematical problem solving ability was obtained the final result that more than 50% of validators agree with all indicators that are used as criteria in validation, all items have very good distinguishing power to distinguish between students who high ability with low ability students, and reliability coefficient of 0.885. Based on the results of this research, the test instrument for mathematical problem solving abilities is can be used to collect data related to mathematical problem solving abilities in algebraic material for class VII junior high school students.

### Acknowledgments

The researcher would like to thank you for the reference to improve the quality of this paper and Junior High School 2 Dukuhwaru for the opportunity to conduct research.

### References

[1] Widiyoko E P 2013 *Evaluasi Program Pembelajaran* (Yogyakarta: Pustaka Pelajar)
[2] Mardapi D 2008 *Teknik Penyusunan Instrumen Tes dan Nontes* (Yogyakarta: Mitra Cendikia Press)
[3] Elvina A 2012 *Hubungan antara Self Regulated Learning dengan Kemampuan Memecahkan Masalah pada Pembelajaran Matematika pada Siswa SMUN 53 di Jakarta Timur* [Online] Retrieved from: publication.gunadarma.ac.id
[4] Webb N L 1997 Processes, conceptual knowledge, and mathematical problem-solving ability *Journal for Research in Mathematics Education* 83-93.
[5] Polya G 1973 *How To Solve It: A New Aspect Of Mathematical Method* (USA: Princeton university press)
[6] Pardimin and Widodo S A 2016 Increasing Skills of Student in Junior High School to Problem Solving in Geometry with Guided *Journal of Education and Learning (EduLearn)* 390- 395
[7] Widodo S A and Sujadi A A 2015 Analisis Kesalahan Mahasiswa dalam Memecahkan Masalah Trigonometri *Jurnal Ilmiah Ilmu Sosial and Humaniora* 1 1
[8] Budiyono 2017 Pengantar Metodologi Penelitian Pendidikan (Solo: UNS Press)
[9] Gay L R 1987 *Education Research, Competencies for Analysis and Application* (Columbus: Merrill Publishing Company)
[10] Suwarto 2007 *Tingkat Kesulitan, Daya Beda, dan Reliabilitas Tes Menurut Teori Tes Klasik Jurnal Pendidikan* 16 2