Development of Mathematical Problem Solving Tests on Geometry for Junior High School Students

Sri Adi Widodo1, Ibrahim2, Wahyu Hidayat3, Samsul Maarif4, Fitria Sulistyowati5
1,5Department of Mathematics Education, Universitas Sarjanawiyata Tamansiswa
2Department of Mathematics Education, Universitas Islam Negeri Sunan Kalijaga
3Department of Mathematics Education, IKIP Siliwangi
4Department of Mathematics Education, Universitas Muhammadiyah Prof. Dr. HAMKA
*sriadi@ustjogja.ac.id

Abstract

One of the cognitive skills that students must have is problem-solving ability. Therefore, a standardized test is needed to measure this ability. However, until now, the tests used to measure mathematical problem-solving ability were not standardized. This research aims to develop a valid and practical mathematical problem-solving test in the discussion of triangles and rectangles. This research is part of the R&D 4D type with steps: Defined (analysis to develop a test), Design (designing a problem-solving test form), and Develop (expert judgment and small group trials). The instrument in this research was the validation sheet used by the validator to assess each test item. The data analysis technique used the Aiken method for expert judgment and qualitative descriptive for development testing. The result is a valid and practical mathematical problem-solving test because the V-Aiken value is more than 0.532. Based on this, the instrument can be used to measure the mathematical problem-solving ability of junior high school students in the discussion of triangles and rectangles.

Keywords: instrument of research, geometry, problem-solving

Introduction

The instrument is one of the tools used to obtain research data (Onwuegbuzie et al., 2010; Wotruba & Wright, 1975). In quantitative and qualitative research, instruments are important and must be available because they aim to get the necessary research data (Borgman, 2012; Drost, 2011; Taherdoost, 2018). The main instrument in qualitative research is the researcher himself, but tests and non-tests are helpful instruments for obtaining data (Girard & Cohn, 2016; Reeves & Marbach-Ad, 2016). One of the instruments that can measure students' cognitive abilities is a mathematical problem-solving test (Jacobse & Harskamp, 2012; Nguyễn & Nguyễn, 2017; Tanujaya et al., 2017). Generally, this test aims to determine students' ability to solve mathematical problems.

Mathematical problem-solving tests are an urgent and essential need as a mathematical education research instrument and a mathematics learning outcome instrument. Problem-solving tests are needed because they are an effort to improve students' mathematical problem-solving abilities as the main goal of mathematics learning. This is in line with the government's
statement that requires every student to be familiar with Higher Order Thinking Skills problems (Widana et al., 2019). The development of mathematical problem-solving tests is an activity that is relevant to the government's statement.

Mathematical problem-solving ability is a skill that students need to have. It because problem-solving skills can be the basis for students to defend their arguments rationally and logically (National Council of Teachers of Mathematics, 2000), so that students can construct their thinking in solving the problems given (Hidayat et al., 2018). Mathematical problem-solving ability is expected to train students in understanding problems, planning strategies for solving problems, solving problems through planning, and checking the results that have been resolved (Hidayat et al., 2019; Hidayat & Aripin, 2019; Hidayat & Sariningsih, 2018). But, in reality, teachers still have difficulties in designing the learning process from planning to evaluation. Instruments or tests to determine students' cognitive abilities (in this case, mathematical problem-solving ability) must be prepared by the teacher as one of the learning planning activities.

The design and development of learning instruments that can improve students' mathematical problem-solving abilities is one solution to solving these problems (Hendriana et al., 2019; Hendriana et al., 2019; Rohaeti et al., 2019). It like Hartatiana (2014), who developed an argument-based problem-solving test in the discussion of fraction, 2D-Geometry, and 3D-Geometry for fifth-grade elementary school students. Likewise, what was done by Silva et al (2011), and Bidasari (2017), who have developed mathematical problems with the PISA Model to measure students' problem-solving abilities. Thus, the research about the development of mathematical problem-solving ability test is urgent will carried out immediately.

Based on these problems, this study aims to develop a valid and practical mathematical problem-solving test instrument. While the difference between this research and existing research is that this problem-solving test focuses on triangular and rectangular problems. Moreover, a significant difference in the problem-solving stages used in this study are the Polya stages, namely understanding the problem, planning to solve the problem, carrying out the plan, and checking the students’ answers (Polya, 2004).

Method

The steps of this research are Define, Design, and Develop in R&D with the 4D type (Thiagarajan et al., 1974). The dissemination can't be done because the learning is very different from the usual, which is the impact of COVID-19. The Define stage was a needs analysis to develop instruments, namely determining indicators and assessment rubrics for problem-
Development of Mathematical Problem Solving Tests on Geometry for ...

solving tests. The Design step aims to design a geometry problem-solving test before it is given to the validator (as an expert) to assess its feasibility. There are two activities at the Develop step consists of expert judgment and development testing or small group trials (Thiagarajan et al., 1974). The expert judgment aims to assess the properness of the problem-solving ability test instrument. Development testing is testing the readability of the instruments, such as ambiguous sentences.

In the expert judgment, 14 people had been selected by purposive sampling as subjects (they will be called validators). The validator consists of 4 experts in Mathematics Education and 10 junior high school teachers (education practitioners). In development testing, the subjects selected by purposive sampling were 15 junior high school students. The students consist of 9 girls and 6 boys. The instrument in this study was a validation sheet to assess the validity of the problem-solving ability test. Aspects of the validation sheet include content, construction, and language. The content aspect consists of two items that are conformity with the indicator (A1) and conformity with school level and student level (A2). The construction aspect consists of two items that are completeness of work instructions (B1) and information on each test item (B2). Likewise, the language aspect consists of two items that are conformity of language (C1) and sentences (C2) so that there are no multiple interpretations.

The data analysis technique has been adapted to each step of this research. At the Define and Design step, the data has been analyzed descriptive-qualitative. The data in development testing has been analyzed descriptive-qualitative because to assess the readability of the problem-solving test. Expert judgment data were analyzed using the Aiken method to determine the content validity of the problem-solving test (Aiken & Patrician, 2000; Aiken, 1980, 1999). Development test data has been analyzed descriptive-qualitative because to see the legibility of the problem-solving test that had been developed.

**Results**

In the method section, it has been explained that this research step is only Define, Design, and Develop. The Disseminate was not carried out because the learning conditions were not normal.

**Define**

The Define stage was a needs analysis to develop instruments, namely determining indicators and assessment rubrics for problem-solving tests. To determine the indicator, it is necessary to know the basic competencies in geometry. The basic competencies are: (1) identify the characteristics of triangles based on their sides and angles; (2) identify the characteristics of
the rectangle; (3) calculate the perimeter and area of triangles and rectangles, then use them in problem-solving, and (4) paint triangles, height lines, dividing lines, weight lines, and axes. However, basic competency numbers (1) and (2) are not used as an indicator because the minimum level of problem-solving tests is level 3 (application). Meanwhile, these two basic competencies are at the level of knowledge and understanding. The fourth basic competency cannot be measured using the problem-solving test that had been developed because the test is a cognitive test.

In this regard, the basic competencies that can be measured using problem-solving tests to be developed are calculating the perimeter and area of triangles and rectangles and using them in problem-solving. This basic competency is breakdown into 4 indicator problems, namely solving problems related to perimeter and area: (1) parallelograms and triangles, (2) triangles and rectangles, (3) squares and triangles, and (4) trapezoid and triangle.

After obtaining the basic competencies and test indicators used as a basis for developing a geometric problem-solving test, the next step is to determine the assessment rubric for problem-solving tests. In general, the assessment rubric for problem-solving tests refers to the four steps of solving a problem, namely 1) understanding the problem, (2) planning the solution, (3) carry out our plan, and (4) recheck the answer or look back at the completed solution (Polya, 2004). The detailed score for each step of problem-solving refers to Widodo (2015) and Widodo et al (2017) which can be seen in Table 1.

<table>
<thead>
<tr>
<th>Polya Principle</th>
<th>Score</th>
<th>Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Understand the problem</td>
<td>0</td>
<td>The student does not write anything. It means the student does not understand the meaning of the proposed problem.</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>The student writes the data/concept/knowledge that is not related to the proposed problem. It means the student does not understand the proposed problem.</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>The student only writes or inform what is known or what is asked only.</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>The student can write down or inform what is known and asked about the problem posed clearly.</td>
</tr>
<tr>
<td>Planning the solution</td>
<td>0</td>
<td>Students do not tell or write steps to solve the problem.</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Students tell or write each step to solve the problem but not coherent.</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Students write the required terms and conditions or formula of the problem submitted and use all information collected.</td>
</tr>
<tr>
<td>Carry out our plan</td>
<td>0</td>
<td>Students are not able to implement a plan that has been made.</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Students implement the plan that has been made, but the error of the procedure and error algorithm.</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Students implemented the plan that has been made, but there are procedural errors and algorithm errors.</td>
</tr>
</tbody>
</table>

Table 1. The score for each step of problem-solving

224
3 Students implemented the plan that has been made, using the correct troubleshooting steps, no procedural errors occur, but an algorithm error occurs.

4 Students do the plan that has been made, using the steps to solve the problem correctly, with no procedural errors and algorithm errors.

Look back at the completed solution

<table>
<thead>
<tr>
<th>No</th>
<th>The Mathematical Problems Solving Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A parallelogram ABCD, P, and Q lie on BD so AP and CQ perpendicular BD. If it’s long an AD = 13 cm, BD = 25 cm and the wide area of the parallelogram is 125 cm², count length of PQ!</td>
</tr>
<tr>
<td>2</td>
<td>Rectangle ABCD, like in the picture above. If length AB = 8 cm and BC = 6 cm, then calculate the length of the track AEFC</td>
</tr>
<tr>
<td>3</td>
<td>ABCD is square with the lengths of the sides are 2 cm. E is a midpoint CD, F is the midpoint AD, G is the cut point of BF with AC, and H is the cut point of BE with AC. Determine the area EDFGH!</td>
</tr>
<tr>
<td>4</td>
<td>A square picture frame is rotated 450, with the axis of the intersection of the diagonal-diagonals. If the length of the square side is 1 cm, determine the area of the incision between the photo frame before and after rotating</td>
</tr>
</tbody>
</table>

**Design**

After the define stage obtained basic competencies, question indicators, and the scoring rubric of problem-solving tests, the next step is to design a flat geometry problem-solving test. Problem-solving tests are designed to display instructions for doing questions and questions that must be done by students. Problem-solving tests are designed to display instructions for solving problems that must be done by students. The instruction section gives the steps that students must take to solve problems such as (1) Write down your identity first; (2) Read carefully each problem given and understand the purpose of the problem; (3) Start by working on the easiest problems, then the more complex problems until you can solve all the problems; and (4) Students in solving the following problems must use the steps: understand the problem, planning the problem, carry out the plan, Look back at the completed solution. In the problem-solving section, it consists of 5 problem essays. The problem-solving test can be seen in Table 2.

**Table 2. Test of mathematical problem solving**

<table>
<thead>
<tr>
<th>No</th>
<th>The Mathematical Problems Solving Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A parallelogram ABCD, P, and Q lie on BD so AP and CQ perpendicular BD. If it’s long an AD = 13 cm, BD = 25 cm and the wide area of the parallelogram is 125 cm², count length of PQ!</td>
</tr>
<tr>
<td>2</td>
<td>Rectangle ABCD, like in the picture above. If length AB = 8 cm and BC = 6 cm, then calculate the length of the track AEFC</td>
</tr>
<tr>
<td>3</td>
<td>ABCD is square with the lengths of the sides are 2 cm. E is a midpoint CD, F is the midpoint AD, G is the cut point of BF with AC, and H is the cut point of BE with AC. Determine the area EDFGH!</td>
</tr>
<tr>
<td>4</td>
<td>A square picture frame is rotated 450, with the axis of the intersection of the diagonal-diagonals. If the length of the square side is 1 cm, determine the area of the incision between the photo frame before and after rotating</td>
</tr>
</tbody>
</table>
Mr. Ali has a plot of land shaped like the picture above. Determine the land area of Mr. Ali!

**Develop**

The development stage is the third stage after the defined and design stages, at this stage, two things must be done, namely expert judgment and development testing (Thiagarajan et al., 1974). Expert judgment involves experts in the field of mathematics education and practitioners of mathematics education. Expert judgment aims to determine a good research product (in this case the validity of the content on the geometric problem-solving test instrument). Development testing in this case can be interpreted as small group trials. Development testing aims to test the readability of the instrument to students so that the instruments obtained are not ambiguous or have multiple meanings. The stage of develop is the foundation of a researcher to carry out the next stage, namely, disseminate. The results of the expert's assessment in assessing 5 problem-solving problems in each aspect can be seen in Table 3.

<table>
<thead>
<tr>
<th>No</th>
<th>The Mathematical Problems Solving Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>![Diagram of Mr. Ali's land plot]</td>
</tr>
</tbody>
</table>

Mr. Ali has a plot of land shaped like the picture above. Determine the land area of Mr. Ali!

**Table 3. Validity Summary V-Aiken**

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Validity Item</th>
<th>Item Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Material aspects (A)</td>
<td>A1 0.881 0.952 0.976 0.976 0.905</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A2 0.786 0.976 0.881 0.881 0.929</td>
<td></td>
</tr>
<tr>
<td>Construction aspects (B)</td>
<td>B1 0.857 0.976 0.786 0.736 0.881</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B2 0.786 0.976 0.786 0.952 0.976</td>
<td></td>
</tr>
<tr>
<td>Language aspects (C)</td>
<td>C1 0.952 0.881 0.881 0.905 0.952</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C2 0.786 0.976 0.881 0.857 0.857</td>
<td></td>
</tr>
</tbody>
</table>

Notes from the validator related to problem-solving test instruments including some question symbols that need to be clarified. As in problem 2, the researcher needs to add the right symbols at points E and F. Problem number 5 the image displayed is incomplete. Other than that two validators from experts in the field of mathematics education said that researchers need to be careful in assessing students because there is a possibility that students do not solve
problems using their pattern steps, but students use three steps that have been inherent in student culture, namely being known, asked and answered.

In the development testing phase, one of the goals is to see the readability of the mathematics problem-solving test instrument. The readability of the problem-solving test instrument is seen from linguistic aspects such as the meaning of math problems that are ambiguous or not. From the results of development testing conducted on 15 students who were taken purposively, it was found that in problems number 2 and number 5, most students stated that they did not understand the meaning of the questions. This is as stated by the validators at the expert judgment stage. The summary notes and improvements to the geometry problem solving test can be seen in Table 4.

**Table 4. Revised summary of problem-solving problems**

<table>
<thead>
<tr>
<th>Number of problems</th>
<th>Problem Before Repairing</th>
<th>Problem after Repairing</th>
<th>Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td><img src="image1.png" alt="Diagram 2" /></td>
<td><img src="image2.png" alt="Diagram 2" /></td>
<td>Improvements to the image added a right symbol</td>
</tr>
<tr>
<td>5</td>
<td><img src="image3.png" alt="Diagram 5" /></td>
<td><img src="image4.png" alt="Diagram 5" /></td>
<td>Improvements to Pictures</td>
</tr>
</tbody>
</table>

**Discussion**

Whether or not the test instrument can be seen from the results of the content validity index (Pardimin et al., 2018). To measure whether an instrument is good or not, the validity index obtained is compared with the product-moment correlation index for N several validators at the significance level $\alpha = 5\%$. Because this study uses 14 validators, the product-moment correlation index is $(r_{xy}) = 0.532$. So, if the content validity index of Aiken ($V$-Aiken) is more than 0.532, then the mathematics problem-solving test instrument is declared good and can be used to measure students' ability in solving math problems. However, if the content validity index of Aiken ($V$-Aiken) obtained is less than or equal to 0.532, then the mathematical problem-solving test instrument needs to be improved.
Content validity is important for cognitive tests, such as achievement tests and problem-solving skills. So before the test is used to obtain data about the cognitive abilities of the respondent, a researcher needs to do a validity test based on the content first. This is done so that a test instrument can reflect the students’ true cognitive abilities. Some things that a researcher can use to enhance content validity include identifying materials that have been given along with their instructional objectives, making a grid of test questions to be tested, compiling test questions and their answer keys, and examining test questions before they are printed or duplicated (Budiyono, 2003). Based on Table 3, it is obtained that each validation item has a content validity index of Aiken (V-Aiken) of more than 0.738. So it can be concluded that the mathematical problem-solving test instrument is expressed both in terms of mathematics, construction, and language. Content validity is a degree of measurement that reflects the expected content domain (Allen & Yen, 2001; Gay, 1990; Kimberlin & Winterstein, 2008). The results of this research product, in general, can add to existing problem problems such as argument-based problem-solving tests for fraction material for elementary students (Hartatiana, 2014), problem-solving tests with the PISA model (Anisah et al., 2013; Bidasari, 2017; Jurnaidi & Zulkardi, 2014; Silva et al., 2013).

The notes from the validator related to the caution that the researcher needs to take in giving student problem-solving scores, the researcher do not write "Students in solving the following problems must use their pattern steps: understand the problem, planning the problem, carry out the plan, Look back at the completed solution "to the instructions in solving the geometry problem. This is because the problem-solving process is a heuristic process, not a problem-solving process that must be sequenced. This contradicts the results of previous research which stated that in solving problems a student must solve them in an orderly manner (Pardimin & Widodo, 2016), in solving problems a student must understand problems, plan problems, and carry out planning (Polya, 2004). If problem-solving is designed as a heuristic process, then the problem-solving procedure can be ignored by students who implicitly carry out Polya's problem-solving stages. This is like the findings in previous research which state that (1) some students do not make stages of understanding the problem by not writing what is known and asked about the problem, but students can solve the problem correctly; (2) or students in writing do not take the stage of checking answers anymore, but students can provide conclusions in solving problems (Widodo, 2014; Widodo et al., 2019; Widodo & Turmudi, 2017).
Conclusion

Based on the results of this study, it can be concluded that a valid and practical mathematical problem-solving test because the V-Aiken value is more than 0.532. Based on this, the instrument can be used to measure the mathematical problem-solving ability of junior high school students in the discussion of triangles and rectangles.

References


