RESEARCH ARTICLE

Developing an Assessment Instrument of Mathematical Problem-Solving Skills in Senior High School

Dyah Ayu Fitriana[1]*, Supahar[2]

[1]Graduate Program of Research and Evaluation in Education, Yogyakarta State University, Indonesia

[2]Faculty of Mathematics and Science Physics Education, Yogyakarta State University, Indonesia

*Corresponding author email: dyahayufitriana@gmail.com

ABSTRACT

This study aims to develop an assessment instrument of mathematical problem-solving skills in senior high school. The developed assessment instrument includes instrument outlines, items of mathematical problem-solving skill in senior high school instrument, and scoring guidelines. The development procedures include 1) Planning the test, 2) Trying out the test, and 3) Establishing the Validity Test, 4) Establishing Test Reliability, and 5) Interpreting the Test Score. The trial was conducted on 508 respondents involving 4 experts and 1 practitioner as raters. The interpretation of the results of the trial is intended to find out the instrument construction and the validity and reliability of the assessment instrument. A problem-solving skill assessment instrument has been validated by the raters. Indicators to be measured include understanding problems, formulating strategies, implementing the strategies, and interpreting the results. The reliability coefficient of the results of the trial is categorized as good in which the Cronbach Alpha value is 0.82. Therefore, the developed instrument can be used to measure the mathematical problem-solving skills in senior high school.

This is an open access article under the CC–BY-SA license.

1. INTRODUCTION

Mathematics is a compulsory general science in every level of education. Mathematics is important for students' lives. Learning mathematics at school is the same as educating children to learn about life outside the school (sutherland, 2007). Mathematical science is also important as it becomes the main key in opening up technology developments which are developing rapidly these days (Kennedy, 2008). The objectives of mathematics lessons stated in the regulations of the Government of the Republic of Indonesia are to make students have reasoning abilities, manipulating and generalizing mathematics, compiling evidence, explaining ideas, understanding problems, designing models, solving and interpreting solutions, and communicating ideas.

The era of the 21st century makes problem-solving skills essential for students. It is indicated by the development of learning models which require the improvement of problem-solving skills. It means problem-solving skills are important for students' lives, including arousing curiosity, attention, tenacity, and confidence and fostering interest in mathematics. To improve problem-solving skills, it is necessary to have the right learning method (Cahyani and Setyawati, 2016). It needs teachers' attention towards students to help the development of their problem-solving skills. Mathematics cannot be separated from problem-solving skills (Ulya, 2014). Mathematics teachers have not found any ways to put problem-solving skills in the learning curriculum (NCTM, 2011). Problem-solving skills can be developed by giving mathematics assignments which have the potential to improve the students'

understanding and mathematical development. According to Danoebroto (2008), explains that students prefer learning through problem-solving activities. It is shown by the level of positive confidence, enthusiastic, and students' creativity. State that problem-solving skills can be improved through real practice questions which exist in students' lives (Saiful et al, 2011). Having problem-solving skills means that students can synthesize their previous knowledge to be applied in different situations (Rudnick, 1995). In line with Rudnick, Posamentier and Krulik (2009) states that to train the way of thinking, students can go through problem-solving tasks. Suitable questions for problem-solving are related to the thinking process, reason, and solution to the problem. The problem-solving steps as delivered by G. Polya, the father of problem-solving, include 1) identifying questions, 2) providing solutions, 3) carrying out the proposed plan, and 4) looking back and evaluating the answers.

The students' low problem-solving skills were seen from the mathematics score in the national examination in 2018. The scores of science major decreased by 4.67 which was from 41.92 to 37.25. It is also clarified that there is an improvement of the national examination questions which is shown by the existence of High Order Thinking Skills (HOTS) questions. Problem-solving is one step to train students to get used to facing HOTS problems. According to Ulya (2016), students' problem-solving skills were still low and they needed to practice continuously with various problem-solving questions.

In addition, there is a fact that problem-solving questions still have not developed. The teachers are also incapable to construct problem-solving questions. According to Sunendar (2017), explained that since problem-solving skills are important, teachers are asked to make questions according to the problem-solving criteria. Therefore, it is necessary to have a problem-solving instrument which can be used as an example for practitioners. According to the background and a number of problems, the objectives of the study are (1) Constructing an instrument for mathematical problem-solving skill questions in senior high school. (2) Validity and reliability of instruments for senior high school mathematical problem-solving skill questions.

2. RESEARCH METHOD

This study is a research and development study. It used the development model of Oriondo and Antonio which includes 1) Planning the test, 2) Trying out the test, and 3) Establishing the Validity Test, 4) Establishing Test Reliability, and 5) Interpreting the Test Score. As presented in Figure 1, the main product of this study is an instrument that can be used to measure problem-solving skills. This study used two to three classes in nine state schools in Kota Yogyakarta with a total subject of 510 students from class XI science. There were two question packages with 16 items for each package and 8 of them are anchor items. The data analysis technique used qualitative descriptive and quantitative analysis. Qualitative analysis was aimed to see the construct of the instrument through expert judgment. Quantitative analysis was used to determine the validity and reliability of the instrument. The validity was seen empirically and from the content. It was seen through the results of expert judgment which was measured by using V Aiken. According to Retnawati (2017), if the Aiken value index is less than or equal to 0.4, it is said to have lower validity. If it is on the range of 0.4 to 0.8, the validity is moderate. If it is greater than 0.8, it is said to be very valid.
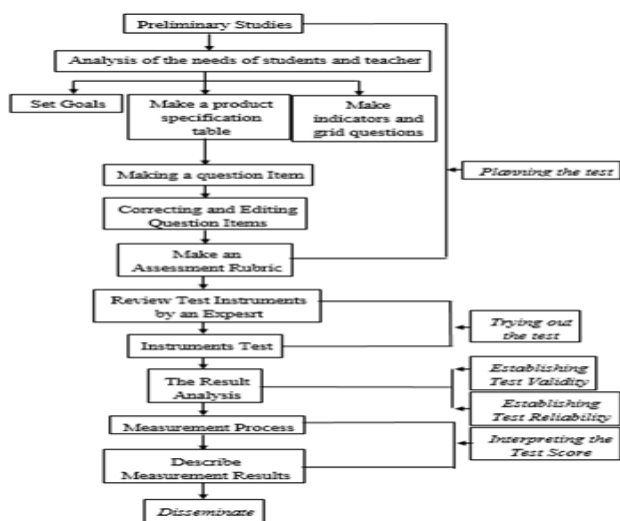

Figure 1. Instrument Development Procedure

3. RESULT AND DISCUSSION

Doing an assessment must be preceded by identifying indicators which are relevant to the learning model used (Supahar, 2015). The indicators used in measuring problem-solving skills are the syntheses of several existing theories. Questions on this instrument are non-routine questions. Each aspect is represented by at least one problem-solving question.

3.1 Aspects of Problem-Solving Skills

According to Polya (1985), NCTM (2000), Adams & Hamm (2010), and Posmentier & Krulik (2009), there are four aspects used in solving problems. They are stated below:

Table 1. Indicators of Problem Solving

Aspect	No.	Indicators	Descriptions
Mathematical Problem-Solving Skills in Senior High School	1	Understanding the problem	Knowing the information on the question and knowing the problem
	2	Formulating strategies	Analyzing the appropriate material by making drawings, tables, mathematical models, and sketches
	3	Implementing the strategies	Working on the problems with the determined strategy
	4	Interpreting the results	Giving opinions or interpreting the results which are found

The four indicators are used to make questions for linear and matrix program materials. Each indicator has two items. Questions on this instrument are in the form of descriptions and the time allocation is 90 minutes (2x school hours).

3.2 Content Validity

Content validity is determined based on expert judgment from experts, practitioners, and peer reviews. The expert judgment results were analyzed by using the V Aiken equation and the results were in the range of 0.7 - 0.8. According to Retnawati (2017), if the Aiken value index less than or equal to 0.4, it is said to have lower validity. If it is in the range of 0.4 to 0.8, it is said that the validity is moderate. If it is greater than 0.8, it is said to be very valid. Based on the Aiken validity calculation, the 16 main questions and 8 anchor questions for packages A and B met the valid criteria. The items had been revised according to the experts' advice. Therefore, the instrument can be used to obtain data in order to empirically get validity and reliability. The calculation results are presented in Table 2.

Table 2. The Results of Aiken's Validity

Items	The Value of Aiken's Validity from 7 Raters
A1, A2, A3, A6, A8, 9, 11, 13, B19, B20, B22, B23, B24	0.8
A4, A7, 10, 12, 15, 16, B18	0.73
A5, 14	0.7
B21	0.67

In addition, the content validity is empirically shown by the item compatibility (the fit model based on the partial credit model) as presented in Figure 2. It appears that the INFIT MNSQ was in the range of 0.77 to 1.32.

```
QUEST: The Interactive Test Analysis System
----------------------------------------------------------------
Item Fit                                          2/ 4/19 14: 7
all on all (N = 382 L = 24 Probability Level= .50)
----------------------------------------------------------------
INFIT
 MNSQ     .63     .71     .83    1.00    1.20    1.40    1.60
-------------+---------+---------+---------+---------+-------
   1 item 1          .            |    *        .
   2 item 2          .            |      *      .
   3 item 3          .            |   *         .
   4 item 4          .            |    *        .
   5 item 5          .            |  *          .
```

```
   6  item 6           .          |  *      .
   7  item 7           .          |  *      .
   8  item 8           .          |  *      .
   9  item 9           .        * |         .
  10  item 10          .        * |         .
  11  item 11          .       *  |         .
  12  item 12          .       *  |         .
  13  item 13          .       *  |         .
  14  item 14          .       *  |         .
  15  item 15          .       *  |         .
  16  item 16          .       *  |         .
  17  item 17          .        *|          .
  18  item 18          .         *          .
  19  item 19          .        *|          .
  20  item 20          .        *|          .
  21  item 21          .       * |          .
  22  item 22          .       *|           .
  23  item 23          .       * |          .
  24  item 24          .       * |          .
=========================================================
```
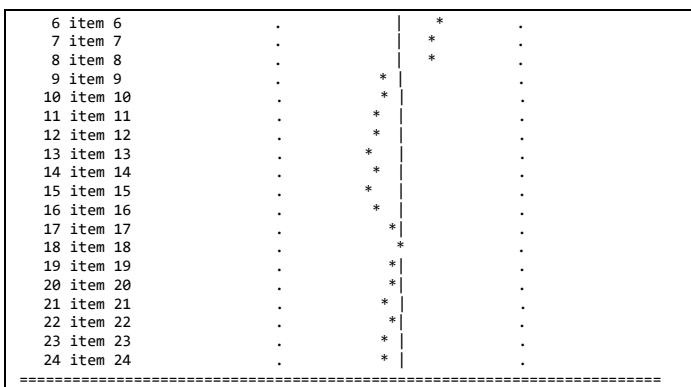
Figure 2. Item Compatibility

According to the partial credit model, the characteristics of an item are indicated by its difficulty level. The analysis of items on the level of difficulty has a range of 0 to 1. If the item is closer to 1, it means that it has a low level of difficulty or it is said to be very easy (Allen and Yen, 1979). Based on the results of the analysis, the difficulty level of the items was in the range of -0.86 to 0.76. The item is said to be good if the level of difficulty is more than -2.00 or less than 2.00. Consequently, all items were said to be good based on the difficulty level as presented in the following fig. 3.
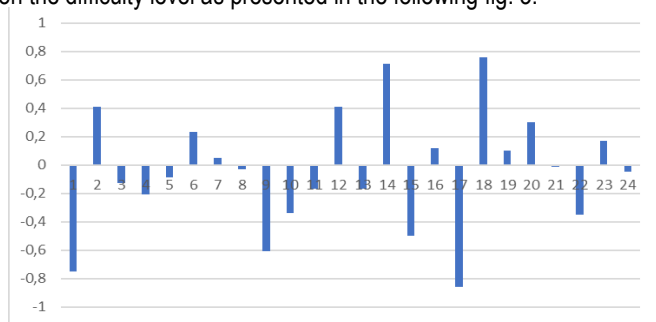

Figure 3. The Difficulty Level of Items

An example of an ICC curve based on the results of the analysis using the Item Response Theory is presented in Figure 4.
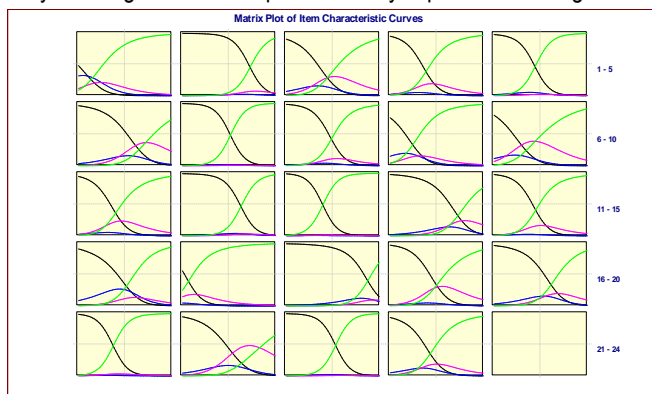

Figure 4. ICC

3.3 Construct Validity

The construct validity used CFA (Confirmatory Factor Analysis). The results showed that there was 1 dominant factor.   Smits, Cujpers, & Straten (2011) state that the output of factor analysis produced by the first factor is able to explain the variance that is more than 20%. The results of the unidimensional of this study show that the

variance which can be explained is 64.8%. It means that this instrument fulfills the assumption of unidimensional as presented in Figure 5.
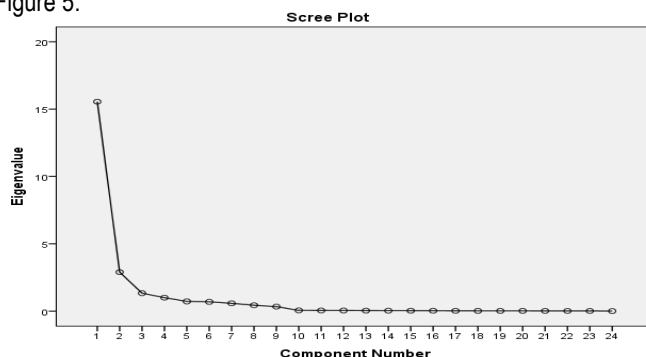

Figure 5. Scree Plot of Factor Analysis

3.4 The Reliability of the Instrument

The reliability was classically estimated by using Cronbach Alpha which obtained a score of 0.83. Based on the Triton category (2006), this instrument was very reliable. Therefore, it can be used as a measuring instrument. The results were analyzed using Item Respond Theory, information functions, and total info. The results show that this instrument is reliable if it is used to measure in the range of -1.8 – 1.8 as presented in Figure 6.
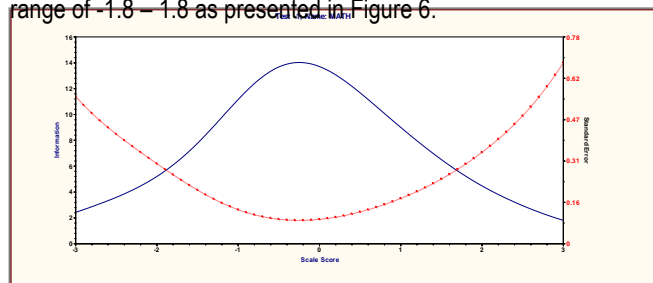

Figure 6. Information Function and Total Info

3.5 An Example of Instrument Items

Here, the researcher attaches an example of questions in the problem-solving skill instrument along with the specifications. Figure 7 presents the item number 8 from package A. The V Aiken score of this item is 0.8, the INFIT MNSQ is 1.08, and the level of difficulty is -0.03. The ICC graph and information functions are also presented in Figure 9. The ICC graph of item 8 (A8) means that the score 0 (category 1) is mostly obtained by students who have very low skill ($\theta = -3$), score 1 (category 2) is obtained by students who have low skill ($\theta = -2$), score 2 (category 3) is obtained by students who have high ability ($\theta = 0$), and score 3 (category 4) is obtained by students who have very high skill ($\theta = 1$). Thus, the level of difficulty is sequential from small to large numbers started from categories 1, 2, 3, and 4.
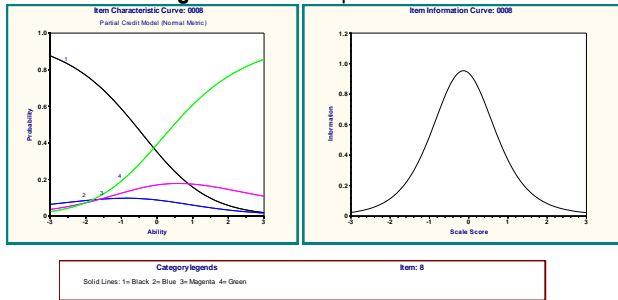
Figure 7. An Example of Items

Figure 8. ICC and Total Info of Item Number 8

4. CONCLUSION

The results of the study showed that 1) The measured aspects include understanding the problem, formulating strategies, implementing the strategies, and interpreting the results. Then, they were arranged into 8 main items in package A, 8 main items in package B, and 8 anchor items. 2) The mathematical problem-solving skill instrument meets the valid requirements based on the results of expert judgment analyzed by using Aiken Validity, Factor Analysis, and Compatibility on the 1PL partial credit model. Meanwhile, the reliability was based on classics estimated by using Alpha Cronbach, and Modern estimated using IRT Total Info. Therefore, this assessment instrument can be used by practitioners and academics to find out the mathematical problem-solving skill level of Senior High School students in class XI Science.

Acknowledgements

REFERENCES

Adams, Dennis & Hamm, Mary. Demystify Math, Science, and Technology (Creativity, Innovation, and Pobelm Solving). United State of America: Rownman & Littlefield Education, 2013, pp. 53.

Allen, M. J., dan Yen, W. M. Introduction toe Measurement Theory. Moterey, Mexico: Brooks/Cole Publishing Company, 1979.

Budi, Triton Prawira. Reliabilitas dan Validitas. Yogyakarta: Andi, 2006, pp. 248.

Cahyani, H., & Setyawati, R. W. (2016). "Pentingnya Peningkatan Kemampuan Pemecahan Masalah melalui PBL untuk Mempersiapkan Generasi Unggul Menghadapi MEA". Seminar Nasional Matematika. 10(1), 151-160.

Danoebroto, Sri Wulandari. "Meningkatkan Kemampuan Pemecahan Masalah melalui Penedekatan PMRI dan Pelatihan Metakognitif". Jurnal Penelitian dan Evaluasi Pendidikan, vol. 11, pp. 73-87, 2008.

G. Polya. How to Solve It: a new aspect of mathematical method. United States of America: Princeton University Press, 1988, pp. 5.

Kennedy, Leonard M., Steve, and Johnson, Art. Guiding Children's Learning of Mathematics. United States of America: Thomson Wadsworth, 2008, pp. 3.

Krulik, S., & Rudnik, J.A. The new sourcebook for teaching reasoning and problem solving in elementary school. Boston: Temple University, 1995.

National Research Council. Assessing 21st Century Principles and standars for school mathematics. Washington, DC: The National Academics Press, 2010.

National Research Council. Assessing 21st Century Skills: Summary of Workshop. Washington, DC: The National Academics Press, 2011.

Oriondo, Leonora L. & Dallo-Antonio, E. M.. Evaluating Educatonal Outcomes (Test, Measrument, and Evaluation). Manila: REX Book Store, 1984, pp. 34-53.

Posamentier, Alfred S. & Krulik, Stephen. Problem Solving in Mathematics, Grade 3-6: Powerful Strategies to Deepen Understanding. Universitas Michigan: SAGE Publications, 2009.

Retnawati, H. Validitas Reliabilitas & Karakteristik Butir. Yogyakarta: Parama Publishing, 2017, pp. 31.

Sunendar, Aep. "Pembelajaran Matematika dengan Pemecahan Masalah". Jurnal THEOREMS (Teh Original Research of Mathematics, vol 2, pp. 86-93, 2017.

Supahar. "Pengembangan Instrumen Penilaian Kinerja Penyusunan Laporan Praktikum Fisika SMP Berbasis Inkuiri". Jurnal Pendidikan Matematika dan Sains, vol 1, pp. 23-29, 2015.

Sutherland, Rosamund. Teaching for Learning Mathematics. London: Open University Press, 2007, pp. 32.

Syaiful, etc. "Peningkatan Kemampuan Pemecahan Masalah Matematika melalui Penedekatan Matematika Realistik". Jurnal Pendidikan Matematika dan Sains, vol 1, pp. 9-16, 2011.

Ulya, Himmatul. (2016). "Profil Kemampuan Pemecahan Masalah Siswa Bermotivasi Belajar Tinggi Berdasarkan Ideal Problem Solving". Jurnal Konseling GUSJIGANG. 2(1), 90-96.