

Development of an Instrument to Measure Higher Order Thinking Skills in Senior High School Mathematics Instruction

Benidiktus Tanujaya

Department of Mathematics Education, University of Papua
Gunung Salju Street, Amban, Manokwari 98314, West Papua, Indonesia

Abstract

The purpose of this research was to develop an instrument that can be used to measure higher-order thinking skills (HOTS) in mathematics instruction of high school students. This research was conducted using a standard procedure of instrument development, from the development of conceptual definitions, development of operational definitions, determination of constructs, dimensions, and indicators, to the preparation of blue prints, item preparation, expert validation, and testing. Data results from trials is analyzed using factor analysis and analysis of structural equation modeling (SEM). The data-analysis shows that there are 9 factors HOTS that construct the instrument with good validity and reliability. This instrument classifies high school students into five categories of HOTS ability. HOTS grouping results can be used by various interested institutions to evaluate the instruction of mathematics. These evaluations are used to determine the success of student learning and the success of teachers' teaching.

Keywords: instrument development, higher order thinking skill, mathematics.

1. Introduction

In the Indonesian educational system, mathematics is one of the subjects getting high attention and considered very important. Recognizing the importance of mathematics, whether in structuring thinking skills and students' attitudes, as well as in using mathematics, the teachers' role in improving the mathematics achievement at every level of education should get high attention.

Along with many other fields, the study of mathematics is increasing in complexity. Therefore, the mathematics that students need to learn today is not the same as what their parents and grandparents needed to learn. According to the National Research Council (NRC, 2001), All young Americans must learn to think mathematically, and they must think mathematically to learn. These learning activities should also be applied to students learning mathematics in Indonesia.

Low skill levels in mathematical thinking may be caused by teacher's lack of attention to such thinking skills. Evidence for this can be seen in the PISA report: regarding "mathematical literacy", which is the variable to measure the students' thinking skills in mathematics (Forster, 2004). The results of the PISA survey in 2012 showed that Indonesian students ranks 64th out of 65 countries. The scores achieved by the Indonesian students is 375, while 615, the highest score is obtained by the students in Shanghai, China (OECD: 2012). The PISA report shows that Indonesian students' thinking skills in mathematics are currently very low. This fact also shows that the thinking skills of Indonesian students, especially in mathematics, get less attention.

Human thinking skills can be classified into two categories; lower order thinking skills (LOTS) and higher order thinking skills (HOTS). According to King, *et al.*, the HOTS of a person will appear when encountering unfamiliar problems, uncertainties, questions, or dilemmas. Furthermore, according to Heong, *et al.* (2011), HOTS is an important aspect of teaching and learning. Thinking skills practices are part of the generic skills that should be infused in all technical subjects. Students with higher order thinking skills are able to learn, improve their performance, and reduce their weaknesses. Therefore, teachers must be aware of the HOTS of students studying mathematics, in order to perform well-qualified mathematics instruction.

The importance of the role of HOTS for the students learning mathematics can be seen in the Murray study about the influence of the selection of materials on mathematics learning exercises conducted by the teacher on students' HOTS (Murray, 2011), as well as research on the use of the "Inquiry-Based Learning" model to improve the students' HOTS done by Rooney (2012). In addition, the development of an instrument to consider HOTS is important in learning because the assessment of learning achievements is changing as worldwide reforms, particularly in science education, promote the shift from traditional teaching for algorithmic, lower-order thinking skills, to higher-order thinking skills (Barak and Dori, 2009).

Therefore, it is necessary to conduct a research to develop an instrument that can be used to measure the students' HOTS in mathematics instruction in senior high school. In order to develop such an instrument, we first need to ask what are HOTS? What indicators that construct HOTS?

According to Wang and Wang (2011), there are three main components in HOTS, i.e. critical thinking skills, design thinking skills, and system thinking skills, while Miri *et al.* (2007), states that HOTS consists of three components, namely critical thinking skills, systematic thinking skills, and creative thinking skills. Furthermore,

according to Yee Mey Hong *et al.* (2011), critical thinking skill and creative thinking skill are two important indicators of HOTS. Thus, there are at least two indicators in HOTS, so that measurement of the students' HOTS can be conducted by observing their critical and creative thinking skills.

HOTS is a latent variable that can't be measured directly. In order to measure the characteristics of latent variables, according to Naga (2012), manifest variables can be used to be measured latent variables. Measurement of manifest variables requires a standardized instrument. The problem now is how to provide an instrument to allow teachers to measure students' HOTS.

On the other hand, the limited knowledge and time the senior high school math teachers have to develop a valid and reliable instrument to measure students' HOTS leads to these skills becoming overlooked, which in turn leads to difficulty achieving the fundamental objectives in Mathematics instruction. Therefore, it is necessary to develop a HOTS instrument for mathematics instruction in senior high school.

Operationally, this study aimed to: (1) generate the indicators of HOTS in mathematics in senior high school, (2) determine the validity of the HOTS instrument, and (3) determine the reliability of the HOTS instrument.

2. Method

This research was conducted in SMA Negeri 1 Manokwari, West Papua Province. Development procedure of the HOTS instrument was done in eight primary steps, consisting of: theoretical review for building conceptual definitions, building operational definitions, defining constructs, dimensions, and indicators, constructing blueprints and items, analyzing readability and social desirability, field testing, and data analysis. Two field trials were conducted, the first with 208 students, while the second trial with 203 students.

The data analysis was performed twice according to the number of trials, using factor analysis. The analysis of the first trial data aimed to select the items that deserved to continue in the second trial, while the results of the factor analysis of data in the second trial were performed using the structural equation modeling (SEM).

There are several requirements in factor analysis, namely: (1) the correlation between the variables, (2) the adequacy of the sample size by using Kaiser-Meyer-Olkin (KMO) formula, (3) test whether the observed data is a sample from a multivariate normal population distribution by using the Bartlett test of sphericity (χ^2), and (4) examine the Anti-image correlation (AIC) with the criterion measure of sampling adequacy (MSA) ≥ 0.50 .

The first stage in the analysis of factors according to Bryman and Cramer (2005), is to calculate the correlation between variables. If the observed variables are not significant, it is not possible formation of one or more factors. According to Widarjono (2012), factor analysis cannot be used if the value of χ^2 has a probability (sig) is greater than 0.05. Furthermore, Santoso (2012), states that the MSA item smaller than 0.50 released one by one from the models ranging from the smallest, to the next item remaining factors analyzed again until all remaining items meet the existing requirements.

The results of the factor analysis using IBM SPSS Statistics program was a linear model combining the items identified. The model obtained is then analyzed by using the SEM analysis (Lisrel 8.8 program) in order to conduct Second Order Confirmatory Factor Analysis testing. At this stage, three tests were performed, namely: (1) the suitability of the data with the model, (2) the validity and reliability of the model, and (3) the significance of the coefficients of the structural model. Hair *et al.* (1998), as cited by Wijanto (2008), states that the evaluation of the model can be conducted using overall model fit, measurement model fit, and structural model fit.

The suitability of the whole models was tested using several measures, as proposed by Wijanto (2008), among others: Normed Fit Index, (2) Non-Normed Fit Index, (3) Parsimony Normed Fit Index, (4) Comparative Fit Index, (5) Incremental Fit Index, (6) Relative Fit Index, (7) Goodness of Fit Index, (8) Adjusted Goodness of Fit Index, (9) Parsimony Goodness of Fit Index, (10) Root Mean Square Residual, and (11) Root Mean Square Error of Approximation.

After the match the model and the data are met, then, according to Wijanto (2008), to test the measurement model fit with an evaluation of each constructor separate measurement models through evaluation of the validity and reliability. Reliability measurements were performed using Construct Reliability (CR) and Variance Extracted (VE). A construct is considered reliable when every indicator has a value of CR ≥ 0.70 , and the value of VE ≥ 0.50 .

3. Result

Based on the expert opinions, some improvements of the structure and content of the instruments that have been prepared were made prior to the second trial? Several statistical values generated in the first and second trials, are presented in Table1.

Table 1. Some Statistics on Trial I and Trial II

No	Statistics	Trial I	Trial II
1	KMO	0,772	0,798
2	Chi-Square of Bartlett Test	4397,738	3283,242
3	MSA	0,542 – 0,878	0,657 – 0,897
4	Number of factor	9	9
5	Total Variance Explained	84,230%	78,101%
6	Normed Fit Index	0,91	0,86
7	Non-Normed Fit Index	0,96	0,90
8	Parsimony Normed Fit Index	0,82	0,77
9	Comparative Fit Index	0,96	0,91
10	Incremental Fit Index	0,96	0,91
11	Relative Fit Index	0,90	0,84
12	Goodness of Fit Index	0,85	0,77
13	Adjusted Goodness of Fit Index	0,82	0,72
14	Parsimony Goodness of Fit Index	0,71	0,64
15	Root Mean Square Residual	1,12	0,54
16	Root Mean Square Error of Approximation	0,051	0,089
17	Standardized Loading Factor	0,71 – 1,93	0,24 – 2,40
18	Construct Reliability	0,80 - 0,97	0,79 - 0,95
19	Variance Extracted	0,57 - 0,78	0,57 - 0,77
20	t-value	2,18 – 37,02	1,77 – 19,53

Table 1 shows that all 20 values in the first and the second test statistic are relatively equivalent. The conclusions based on the values were also not statistically different. The statistical value of the first third, KMO, Chi-Square Test of Bartlett, and MSA in the second test gives same result, so the factor or the formation of factor analysis can be performed.

The further results of the factor analysis in the both trial also showed similar result. The number of factors yielded by the analysis of each trial are the same, namely 9 factors, and each analysis explains about 80% of the total variance. The results of analysis by using SEM also shows that the results did not differ between the two experiments performed.

4. Discussion

The results of research showed that the HOTS instrument has good validity and reliability, and hence it provides a good measure of high school students' HOTS in learning mathematics. The diversity of the students' work provide evidence for this. For example, consider the following questions:

In accordance with Governor Jokowi's program, an area in Rawamangun, Jakarta will be built as an open green park.

If the planned park area of 400 square meters, then describe the area.

The answers given by the students involved in the research are quite varied. The diversity of the responses indicates that the instrument can be used to measure the students' HOTS ability well. Some examples of the students' answers are presented below:

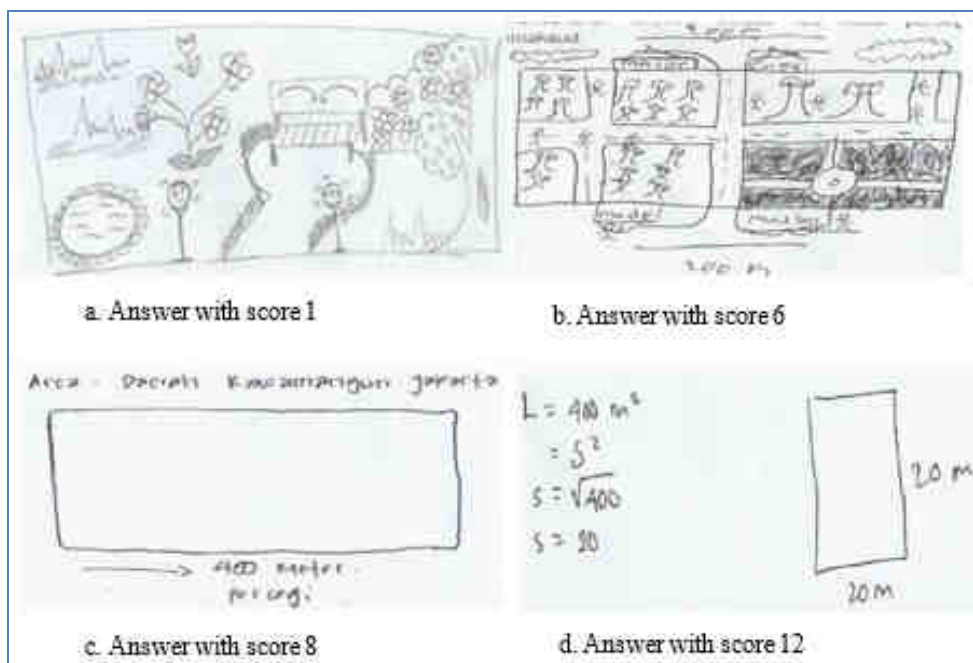


Figure 1. Some Samples of Student's Answer

The answer in Figure 1a shows that students does not understand the purpose of the question. The Student has sketched the garden square shaped, but the problem in not answered properly. The Student has tried to give an answer, but the answer is wrong. The lack of both numerical values and units both call into question the student's understanding of the problem.

The answer in Figure 1b indicates that the student has understood the question is asking for a sketch of the park; however the student has not arrived at numerical values to correctly determine the area of the park as requested in the question.

In Figure 1c, it appears that the student has understood the intention of the question. The students has established units to determine area as required, but did not specify the length or width of each side.

The next student, whose work is presented in Figure 1d a perfect answer. The student's work shows that in addition to sketching the park properly, the student has also correctly set units to determine the area of the park. Compared to the previous students, this student's answer is more correct because it provides numbers for the length and width of each side, obtained by using a specific calculation.

The variety of the answers given by students shows the variety of students' thinking skills, as well as the variance of the sample. This is in accordance with the opinion of Tanujaya (2013), which suggests that the variety of data in statistics is an important factor in the analysis of research data, both in estimation and testing of population parameters.

In addition to the variety of the sample-data, additional factors aid in developing an effective HOTS instruments: First, this HOTS instrument is based on standard procedures that have been put forward by various experts of measurement. According to Azwar (2012), when the definition of an attribute being measured is poorly understood, its measurement may be overlapped or conflated with other attribute. Conversely, when the definition is well-understood, the resulting instrument is comprehensive enough to reveal the desired attributes.

Second, after the draft was formed, the instrument was validated by domain experts based in Jakarta and Bandung, as well as Manokwari, West Papua. In addition to technical improvements to the substance of mathematics, experts also provide advice about the language and content of the material. Topics on derivatives and exponentials were removed from the instrument, because the subject has not been studied by the high school students of class XI Science Department. Excessively long story problems were also eliminated due to time restriction. Culturally-based items, such as questions about the mathematics of the Olympics, were also eliminated. All criticisms and suggestions put forward by the experts were considered. Expert opinion was an important component in the development of the instrument. This is in accordance with the discussion among various experts of measurement, which always include expert domain opinion as an important step in instrument development.

Based on the results of the factor analysis, as noted earlier there are nine factors that are formed from the 27 indicators analyzed by the HOTS constituent instrument. The nine factors describe cognitive activity which goes over and beyond both knowing (knowledge) and understanding (comprehension). HOTS is a cognitive activity that is more than just memorizing and understanding. This is in accordance with the opinions expressed by Zohar

(2004), who states that the knowing and understanding cognitive activities are grouped into low-level thinking skills (lower-order thinking, e.g. LOTS), while HOTS is a high-level cognitive activity in the Bloom taxonomy-based analysis of activity, requiring synthesis and creativity.

Some examples of HOTS cognitive abilities are classified by Zohar (2004), among others: preparing arguments, asking research questions, making comparisons, solving complex, non-algorithmic problems, dealing with controversy, and identifying hidden assumptions. Most scientific research skills, such as formulating hypotheses, planning experiments or drawing conclusions, are also classified as HOTS. The examples presented by Zohar conform well to the HOTS indicators from the constituent instrument proposed in this study.

5. Conclusion

Based on the results of research and discussion presented here, the proposed HOTS instruments can be used to measure HOTS of high school students in mathematics instruction, because the instrument has good validity and reliability. The instrument consists of nine items. Every item that is used is representative of each factor, namely: the use of the concept, the use of the principle, impact predicting, problem solving, decision making, working in the limit of competence, trying new things, divergent thinking, and imaginative thinking.

This instrument still needs to be developed by conducting the test on students with different characteristics from the students of SMA Negeri 1 Manokwari. Tests with broader samples will be required in order to have higher rate of external validity.

Acknowledgement

The author would like to thank everyone contributed in this research, especially the students, mathematics teachers, and Mr. Lukas Wenno, the head master of SMA Negeri 1 Manokwari. Thanks also to my lovely-wife Jeinne Mumu, my sweet-daughters Aurelia, for all the support, and Mathew Zucker for revision and suggestion.

References

- Azwar, Saifuddin. (2012). *Penyusunan Skala Psikologis*, Yogyakarta: Pustaka Pelajar.
- Barak Miri dan Yehudit Judy Dori, (2009). "Enhancing Higher Order Thinking Skills among Inservice Science Teachers via Embedded Assessment." *J Sci. Teacher Educ.* 20: 459-474.
- Bryman, Alan and Duncan Cramer. (2005). *Quantitative Data Analysis with SPSS 12 and 13: A Guide for Social Scientists*. New York: Routledge.
- Margaret Forster, (2004). "Higher Order Thinking Skills," *Research Developments*, 11(4), Australian Council for Educational Research.
- King, F. J., Ludwika Googson, and Faranak Rohani, *Higner Order Thinking Skills*. [Online] Available: www.cala.fsu.edu (May 10, 2013).
- Miri, Barak, Ben-Chaim David, and Zoller Uri. (2007). "Purposely teaching for The Promotion of Higher Order Thinking Skills: A Case of Critical Thinking." *Res Sci Educ*, 37: 353-369.
- Murray, Eileen Christina. (2011). "Implementing Higher-Order Thinking In Middle School Mathematics Classrooms," *Dissertation*, University of Georgia.
- Naga, Dali Santun. (2012). *Teori Sekor pada Pengukuran Mental*. Jakarta: Nagrani Citrayasa.
- National Reseach Council. (2012). *Improving Mathematical Education*. Washington: National Academy Press.
- OEDC, (2014). PISA 2012 Result: What Students Knows and Can Do, Volume 1, OECD.
- Rooney, Caitriona. (2012). "How am I using inquiry-based learningto improve my practice and to encourage higher order thinking among my students of mathematics?" *Educational Journal of Living Theories*, 5(2): 99-127.
- Ruseffendi, E. T. (2006). *Pengajaran Matematika: Untuk Meningkatkan CBSA*. Bandung: Tarsito.
- Santoso, Singgih. (2012). *Aplikasi SPSS pada Statistik Multivariat*. Jakarta: PT Elex Media Komputindo.
- Tanujaya, Benidiktus. (2013). *Penelitian Percobaan*. Bandung: Remaja Rosdakarya.
- Wang, Shouhong, and Hai Wang. (2011). "Teaching Higher Order Thinking in the Introductory MIS Course: A Model-Direct Approach." *Journal Education for Business*, 86: 208-212.
- Widarjono, Agus. (2012). *Analisis Statistika Multivariat Terapan*. Yogyakarta: UPP STIM YKPN.
- Wijanto, Setyo Hari. (2008). *Structural Equation Modeling dengan Lisrel 8.8: Konsep dan Tutorial*. Yogyakarta: Graha Ilmu.
- Yee Mei Heong, W. Othman, J. Yunos, Tee Tze Kiong, R. Hassan, and M. M. Mohamad. (2011). "The Level of Marzano Higher Order Thinking Skills among Technical Education Students." *International Journal of Social Science and Humanity*, 1(2): 121-125.
- Zohar, A. (2004) *Higher Order Thinking in Science Classroom: Student' Learning and Teacher Professional*. Michigan: Springer.