INSTRUMENT DEVELOPMENT OF HIGHER ORDER THINKING SKILL IN MATHEMATICS INSTRUCTIONAL ON SENIOR HIGH SCHOOL

Benidiktus Tanujaya
Department of Mathematics Eduacation, University of Papua
Gunung Salju Street, Amban, Manokwari 98314
benidiktus.0903@gmail.com

ABSTRACT

The purpose of this research was to develop an instrument that can be used to measure higher-order thinking skills (HOTS) in mathematics instructional of high school students. This research was conducted using a standard procedure of instrument development, from the development of conceptual definition, development of operational definitions, determination constructs, dimensions, and indicators, to the preparation of blue print, item preparation, expert validation, and testing. Data results from trials be 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 in the five categories of HOTS ability. HOTS grouping results can be used by various interested institutions to evaluate the instructional 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.*

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

Complexity in studying mathematics increases in line with the increase in various fields today, including the increase in the field of mathematics itself. Therefore, the mathematics students need to learn today is not the same as that their parents and grandparents needed to learn. According to the National Research Council (NRC, 2001:1), 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.

Lack of attention to thinking skills in Indonesian students who study mathematics can be predicted as the cause the low of students' thinking skills. This evident can be seen in the PISA report, regarding the mathematical literacy, which is the variable to measure the students' thinking skills in mathematics (Forster, 2004:14). 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, 19). The PISA report shows that Indonesian students' thinking skills in mathematics are currently very low. This fact also shows that the thinking skills Indonesian students, especially in mathematics, get less attention.

The human thinking skills can be classified into two categories; lower order thinking skill (LOTS) and higher order thinking skill (HOTS). According to King, et al., the HOTS of a person will appear when encountered unfamiliar problems, uncertainties, questions, or dilemmas. Furthermore, according to Heong, et al. (2011: 121), HOTS is an important aspect in 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, the HOTS of students studying mathematics need to know by the teachers, so they can perform the qualified mathematics instructional.

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 to the students' HOTS (Murray, 2011: 34), as well as a research on the use of learning models Inquiry-based learning to improve the students' HOTS done by Rooney (2012: 99). In addition, the development of HOTS instrument 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, 2009:462).

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 an instrument that measures the students' HOTS, the first need to know what is HOTS? What indicators that construct the HOTS?

According to Wang and Wang (2011: 209), there are three main components in HOTS, i.e. critical thinking skills, design thinking skill, and system thinking skill, while Miri et al. (2007: 355), states that HOTS consists of three components, namely critical thinking skill, systematic thinking skill, and creative thinking skill. Furthermore, according to Rosnawati (2009: 3) and Yee Mey Hong et al. (2011: 121), critical thinking skill and creative thinking skill are two important indicators of HOTS. Thus, there are at least two indicators in HOTS, so that finding the students' HOTS can be conducted through observation concerning the critical and creative thinking skills.

HOTS is a latent variable that can not be measured directly as the physical variables. In order to measure the characteristics of latent variables, according to Naga (2012: 13), the manifest variables can be used to be measured the latent variables. Measurement of the manifest variables requires a standardized instrument. The problem now is how the teacher can have a standardized instrument and be qualified to measure the students' HOTS.

On the other hand, the limited knowledge and time the senior high school math teachers have in developing a valid and reliable instrument to measure the students' HOTS becomes a constraint for the lack of attention in achieving the fundamental objectives in Mathematics instruction. Therefore, it is necessary to develop a HOTS instrument in 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 construct validity of the HOTS instrument in the mathematics instruction in senior high school, and (3) determine the reliability of the HOTS instrument in mathematics learning in senior high school.

METHOD

This research was conducted in SMA 1 Manokwari, West Papua Province. Development procedure of HOTS instrument was done in eight primary steps, consisted of: theoretical review for building conceptual definition, building operational definition, defining construct, dimension, and indicators, constructing blueprint and items, analyzing readability and social desirability, field testing, and data analysis. Two field trials were conducted, the first was at 208 students, while the second trial followed by 203 students.

The data analysis was performed twice according to the number of trials, using factor analysis. The analysis of the first trial data aims to select the items that deserves to be continued in the second trial, while the results of the factor analysis of data followed by a second trial by using the analysis of Structural Equation Modeling (SEM).

There are several requirements in factor analysis, namely: (1) the correlation between the variables. The first stage in the analysis of factors according to Bryman and Cramer (2005: 326), is to calculate the correlation between variables. If the observed variables are not significant, it is not possible formation of one or more factors, (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 ( ). According Widarjono (2012: 242), factor analysis can not be used if the value of has a probability (sig) is greater than 0.05, and (4) examine the Anti-image correlation (AIC) with the criterion measure of sampling adequacy (MSA) ≥ 0.50. Santoso (2012: 66), 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 package 20 was the establishment of several factors as a model which is a linear combination of the items. The model obtained is then analyzed by using the SEM analysis lisrel8.8 program package. By using 8.80 lisrel program, the testing of Second Order Confirmatory Factor Analysis was conducted.

At this stage, three tests was performed, namely: (1) the suitability of the data with the model, (2) the validity and reliability ofthe model, and (3) the significance of the coefficients of the structural model. Hairetal. (1998), as cited by Wijanto (2008: 49), states that the evaluation of the degree of fit of the data to match the model through the entire model (overall model fit), the measurement model fit (measurement model fit), and the suitability ofthe structural model (structural model fit).

The suitability of the whole models was tested by usingseveral measures, as proposed by Wijanto (2008: 61-62), among others: Normed Fit Index (NFI), (2) Non-Normed Fit Index (NNFI), (3) Parsimony Normed Fit Index (PNFI), (4) Comparative Fit Index (CFI), (5) Incremental Fit Index (IFI), (6) Relative Fit Index (RFI), (7) Goodness of Fit Index (GFI), (8) Adjusted Goodness of Fit Index (AGFI), (9) Parsimony Goodness of Fit Index (PGFI), (10) Root Mean Square Residual (RMR), dan (11) Root Mean Square Error of Approximation (RMSEA).

After the match the model and the data are met, then, according to Wijanto (2008: 65-66), to test the measurement model fit, withan evaluation of each constructor separate measurement models through evaluation of the validity and reliability. Reliability measurements performed using CR and VE. A construct as a good reliability when every indicator has a value of CR ≥ 0.70, and the value of VE ≥ 0.50.

RESULTS

Based on the expert opinions, some improvements of the structure and content of the instruments that have been prepared were made, prior to trial several statistical value, which was 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	NFI	0,91	0,86
7	NNFI	0,96	0,90
8	PNFI	0,82	0,77
9	CFI	0,96	0,91
10	IFI	0,96	0,91
11	RFI	0,90	0,84
12	GFI	0,85	0,77
13	AGFI	0,82	0,72
14	PGFI	0,71	0,64
15	RMR	1,12	0,54
16	RMSE	0,051	0,089
17	SLF	0,71 – 1,93	0,24 – 2,40
18	CR	0,80 - 0,97	0,79 - 0,95
19	VE	0,57 - 0,78	0,57 - 0,77

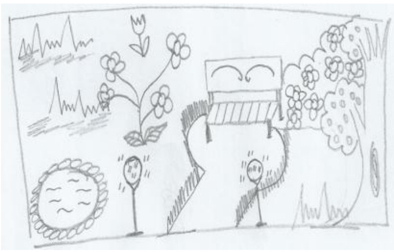
20	Nilai t	2,18 – 37,02	1,77 – 19,53

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

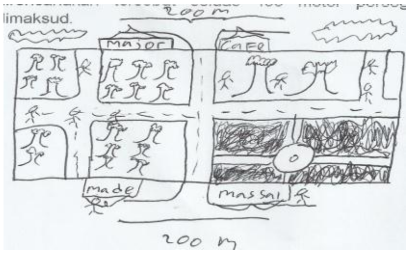
The further results of the factor analysis in the second trial also not showed different ones. Number of factors formed in both analysis are the same, namely 9 factors, the total variance explained, is not too different, which is about 80%. The results of analysis by using SEM also shows that the results did not differ between the two experiments performed. Test of model fit the data used shows that structural equation models were generated in both trials give different results.

DISCUSSION

The results of research showed that the instrument HOTS has good validity and reliability, so that it is qualified tomeasure the high school students' HOTS in learning mathematics. The diversity ofthe students' work results showed this evident. For example, considerone of the following questions:

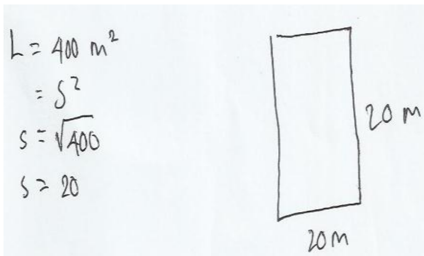


a. answer with score 1



b. answer with score 6

c. answer with score 8



d. answer with score 12

Figure 1. Some Sample of Student's Answer

In accordance with Governor Jokowi 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 very various. The diversity of this response indicates that the instrument can be used to measure the students' HOTS ability well. Some examples of the students' answers are presented below:

Answer in Figure 1 section a, shows that students do not understand the purpose of the question either. Students have sketched garden square shaped, but the problem in questionis not answered properly. Students have tried to give an answer, but the answer is wrong. There are two mistakes made that answer without value and unit to indicate broad question.

Answer in Figure1part b, indicates that the student has understood the question that is sketching a garden. Mistakes made in setting the numbers to determine the area of the park, as requested in the question.

In Figure 1 section c, it appears that students have understood the intention of the question that is sketching a garden. Students have established a unit to determine the number and size as required in the matter, but did not specify the length o rwidth of each side.

The next student, as his work is presentedin Figure 1 part d, shows considerable work perfectly. The student's work shows that in addition to sketching a garden properly, students are also correct in setting the unit to determine the number and area of the park. The advantages of these students compared to students in advance, a number that represents the length and width of the side, is obtained by using a specific calculation.

The variety of the answers given students showed the variety of students' thinking skills, which also shows the variety of the sample used. The variety of characteristics of the sample because the data obtained in the trial also varied. The variety of the resulting data is suspected as the cause of the statistical analysis performed on both trials provide maximum results. This is in accordance with the opinion Tanujaya (2013: 6), which suggests that the variety of data instatistics is an important factor in the analysis of research data, both the estimation and testing of the population parameter.

In addition to the variety of the sample, the following will be presented briefly how the preparation of instruments HOTS development so as to obtain maximum results.

First, this HOTS instrument is based on standard procedures that have been put forward by various experts of measurement. Determination of dimensions and indicators carried out by standard procedures to understand the concept of HOTS theory is good and true. According to Azwar (2012: 11), less understood definition of an attribute being measured results in the desired measuring region in to a measuring region over lapping with other attributes. This causes the instrument obtained be comprehensive enough to revealthe desired attributes.

Second, after the draft was formed, the instrument was validated by expert. Validation of experts in addition to do in Jakarta and Bandung, also conducted in Manokwari, West Papua. In addition to technical improvements to the substance of mathematics, experts also give advice about the language and content of the material. Topics on derivative exponential removed from the instrument, because the subject has not studied by the high school students of class XI Science Department. Problem stories that are too long are also eliminated because the students will take a long time to do it. Some items that like the questions of mathematics

Olympics also eliminated. All Criticisms and suggestions put forward by the experts are always considered. Expert opinionis an important component in the development of the instrument. This is in accordance with the opinion of various experts of measurement, which always include expert opinion as one of the things that must be done in the instrument development.

Based on the results of the factor analysis, as noted earlier. Then there are nine factors that are formed from the 27 indicators analyzed HOTS constituent instrument. The nine factors are cognitive activity over activity knowing (knowledge) and understanding (comprehension). HOTS is a cognitive activity that is more than just memorize and understand. This is in accordance with the opinions expressed Zohar (2004: 1), which states that know and understand the cognitive activities are grouped into low-level thinking skills (lowerorder thinking), while HOTS is a high-level cognitive activity in the Bloom taxonomy analysis activity,synthesis, and create.

Some examples of cognitive activity HOTS classified according to the Zohar (2004: 1-2), among others: preparing arguments, asking research questions, make comparisons, solve complex problems that non-algorithmic, dealing with controversy, and identify the hidden assumptions. Most scientific research skills, such as formulating hypotheses, planning experiments or draw conclusions, also classified as HOTS. Examples of the Zohar presented in conformity with the indicators HOTS constituent instrument in this study.

Based on the results of research and discussion that has been presented, it can be stated that HOTS instruments developed can be used to measure HOTS high school students in mathematics instructional. Thus the instrument used as a test to measure HOTS students, only consists of nine items. Every item that is used is representative of each factor, namely: (1) the use of the concept, (2) the use of the principle, (3) impact predicting, (4) problem solving, (5) decision-making, (6) working in the limit of competence, (7) trying the new things, (8) divergent thiking, and (9) imaginative thinking.

This instrument still needs to be improved by testing on students with different characteristics with SMA Negeri 1 Manokwari. Tests with broader samples intended that this instrument can be used in the wider population. In addition it is necessary to develop a matter for the whole class and other educational levels.

In this instrument there are groups of about equal previously developed for each indicator on each factor. However, in order to avoid leak age problems, it is necessary to develop similar matter even more in order to develop the bank of item.

HOTS instrument is intended for students of Class XI High School Science Department. The use of this instrument can be extended to all levels of education, after the development of items at each level of education. Development of the items according to the factors obtained. The work of students was evaluated based on the rubric that has been developed. There are three important matters of assessment, namely: (1) understanding the problem, (2) troubles hooting procedures, and (3) the truth of the answer. Each subject has a rating score of 0 to 4, so that each item, scores obtained by students ranged from 0 to 12. Thus, for 9 instrument item is done, students obtain a score ranging from 0 -108.

The range of scores obtained by students can be grouped into several categories. The purpose of categorization according to Anwar (2012: 147) is to put people in groups whose position according to a continuum based on the attribute being measured, and the categorization is based on the assumption that individual score in a population of normal spreads.

Therefore to categorizing HOTS scores, first of all it is assumed that the student scores HOTS have normal distribution. Normally distribution population according to Subanar (2012: 135) has two parameters, namely the average (μ) and standard deviation (σ). A normal distribution populations have measures of central tendency, which includes the average, median, and mode are the same. HOTS known minimum score is 0 and the maximum is108, so that the average score of HOTS is 54, while the standard deviationis unknown, so need a simulation.

The simulation was performed using the Minitab to build a population of normal distribution of data, the data range 0-108, and a total population is 1,000,000. The simulation results show that the standard deviation of the population (σ) is 15.The value of standard deviation and average, according to Anwar (2012: 148) is used as the basis for categorization, in order to obtain classificationof HOTS, as presented in Table 2.

Table 2. Kriteria Penilaian HOTS Siswa SMA

No	Nilai HOTS	Kategori
1	≥ 76	Very High
2	$61 \leq \text{HOTS} < 76$	High
3	$46 \leq \text{HOTS} < 61$	Moderate
4	$31 \leq \text{HOTS} < 46$	Low
5	< 31	Very Low

Based on Table 2, the HOTS of high school students who study mathematics grouped into 5 categories. Each category shows the level of HOTS students in understanding and completing math problems. Categories achieved by students can be used as an evaluation study of mathematics. For example, if most of the students are grouped in categories that are very low, then the mathematics instructional in the class (school) needs to be thoroughly evaluated. On the other hand, if most of the students are categorized into high or very high level, it can be stated that mathematics instructional at the school has been going well, especially in the development of HOTS students.

CONCLUSION

Based on the results of research and discussion, it can be concluded that, there are nine factors that make up the instrument HOTS, namely: (1) the use of mathematical concepts, (2) the use of mathematical principles, (3) to predict the impact, (4) problem solving, (5) decision-making, (6) working in the limits of competence, (7) to try new things, (8) divergent thought patterns, and (9) imaginative mindset.

REFERENCES

Azwar, Saifuddin. *Penyusunan Skala Psikologis*, Yogyakarta: Pustaka Pelajar, 2012.

Barak Miri dan Yehudit Judy Dori, "Enhancing Higher Order Thinking Skills among Inservice Science Teachers via Embedded Assessment." *J Sci Teacher Educ.* 20. 2009: 459-474.

Bryman, Alan dan Duncan Cramer. *Quantitative Data Analysis with SPSS 12 and 13: A Guide for Social Scientists.* New York: Routledge, 2005.

Margaret Forster, "Higher Order Thinking Skills," *Research Developments,* Volume 11, Art. 4, Australian Council for Educational Research, 2004.

King, F. J., Ludwika Googson, dan Faranak Rohani, *Higner Order Thinking Skills.* www.cala.fsu.edu (diakses 10 Mei 2013).

Miri, Barak, Ben-Chaim David, dan Zoller Uri. "Purposely teaching for The Promotion of Higher Order Thinking Skills: A Case of Critical Thinking." *Res Sci Educ*, 37, 2007: 353-369.

Murray, Eileen Christina. "Implementing Higher-Order Thinking In Middle School Mathematics Classrooms," *Dissertation*, University of Georgia: 2011.

Naga, Dali Santun. *Teori Sekor pada Pengukuran Mental*. Jakarta: Nagarani Citrayasa, 2012.

National Reseach Council. *Improving Mathematical Education.* Washington: National Academy Press, 2001.

OEDC, PISA 2012 Result: What Students Knows and Can Do, Volume 1, OECD, 2014.

Rooney, Caitriona. "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,* Vol 5(2), 2012: 99-127.

Rosnawati. "Enam Tahapan Aktivitas dalam Pembelajaran Matematika untuk Mendayagunakan Kemampuan Berpikir Tingkat Tinggi Siswa." *Makalah* Seminar Nasional Revitalisasi MIPA dan Pendidikan MIPA, Yogyakarta: UNY, 2009.

Ruseffendi, E. T. *Pengajaran Matematika: Untuk Meningkatkan CBSA*. Bandung: Tarsito, 2006.

Santoso, Singgih. *Aplikasi SPSS pada Statistik Multivariat.* Jakarta: PT Elex Media Komputindo, 2012.

Subanar. *Statistika Matematika: Probabilitas, Distribusi, dan Asimtotis dalam Statistika*. Yogyakarta: Graha Ilmu, 2012.

Tanujaya, Benidiktus. *Penelitian Percobaan.* Bandung: Remaja Rosdakarya, 2013.

Wang, Shouhong, dan Hai Wang. "Teaching Higher Order Thinking in the Introductory MIS Course: A Model-Direct Approach." *Journal Education for Business*, 86, 2011: 208-212.

Widarjono, Agus. *Analisis Statistika Multivariat Terapan*. Yogyakarta: UPP STIM YKPN, 2012.

Wijanto, Setyo Hari. *Structural Equation Modeling dengan Lisrel 8.8: Konsep dan Tutorial.* Yogyakarta: Graha Ilmu, 2008.

Yee Mei Heong, W. Othman, J. Yunos, Tee Tze Kiong, R. Hassan, dan M. M. Mohamad. "The Level of Marzano Higher Order Thinking Skills among Technical Education Students." *International Journal of Social Science and Humanity*, Vol. 1, No. 2, July 2011: 121-125.

Zohar, Anat. *Higher Order Thinking in Science Classroom: Student' Learning and Teacher Professional.* Michigan: Springer, 2004.