Development Of Assessment Instrument Based on Higher Order Thinking Skill (HOTS) To Improve Critical Thinking for Class XI IPS At Sultan Iskandar Muda Private High School Medan for the 2020/2021 Academic Year

Darto Paulus Simanihuruk[1], Indra Maipita[2], Dede Ruslan[3]
{dartopm01@gmail.com}

Program Studi Pendidikan Ekonomi, Program Pascasarjana Universitas Negeri Medan[123]

Abstract. The aim of this research is HOTS-based questions to improve students' critical thinking was developed through the Borg and Gall development model. The results of the material expert validation have an average score of 91 percent in the very good category and the evaluation expert validation has an average of 90 percent in the very good category, teacher response validation scores 93 percent in the very good category, and student response validation scores 95 percent are in the very good category. The effectiveness of HOTS-based questions with the interpretation of the n-gain score category for the experimental group of 52.2% is in the less effective category, while the control group is 28% in the ineffective category. Therefore, HOTS-based questions are considered less effective in improving students' critical thinking in economics subjects.

Keywords: Assessment Instruments, HOTS, Critical Thinking

1 Introduction

Assessment is the evaluation and results of the study of information to determine the achievement of student learning outcomes (Suwartini, Samsi Haryanto, 2017). In conducting evaluations, educators need to make a study of good questions to test the ability of knowledge, attitudes and skills. The results of the evaluation carried out will have an impact in making a decision. That is, the results of the evaluation will provide a better picture of students with their achievements in learning. In the 2013 curriculum, students are required to have the ability to think highly through the learning process and evaluation of learning, for example in working on assessment instruments that are HOTS.

Meanwhile, teachers are required to be able to develop assessment instruments in the form of better and good quality HOTS so that they can be used as good references in evaluating learning. Assessment instruments not only evaluate students' abilities but can improve students' critical thinking. One approach taken is to develop questions that contain high reasoning, namely between C4 to C6. HOTS-oriented assessment instruments are a necessity in today's education world. In the world of education, HOTS thinking is more emphasized in working on questions that are able to train students' thinking patterns such as problem solving, criticizing, and improving children's abilities to this level. HOTS is very closely related to the current education curriculum in Indonesia.

So far, schools are still using the K13 curriculum which aims to produce people who are great, critical, innovative and have extraordinary creativity in learning. From the findings, according to the development rules, studies from students at the research site showed that students were not familiar with HOTS abilities, for example working on more difficult questions with full analysis. In fact, students tend to prefer to work on questions that are easier and easier to solve. Therefore, they are not used to it and do not excite children's interest in learning. Students are happier with instant answers from the internet without thinking, only attaching importance to high grades from the teacher without understanding the essence of the learning objectives.

Meanwhile, teachers are only used to making easy questions, not following the rules for making good and quality assessment instruments such as semester exam questions in economics subjects. Where the questions that have been made have not included a question grid and do not include a description score. Furthermore, the researcher analyzed the midterm exam questions made by the teacher based on the results of the state's questions. It was found that 70 percent was invalid, 76.6 percent was bad questions and 46.7 percent was easy category questions. Based on these problems, researchers are interested and need to develop good and quality HOTS-based assessment instruments to improve students' critical thinking.

ICONSEIR 2021, December 21, Medan, Indonesia

DOI 10.4108/eai.21-12-2021.2317290

Assessment Instrument

The assessment instrument is a tool to measure how well students have improved their learning based on standards (Mangiante, 2013: 222). According to Suryani (2015: 457),an assessment instrument is an assessment tool that will be used by educators to assess the achievement of students through test and non-test techniques. Meanwhile, according to Amirono and Dariyanto (2016), the assessment technique is divided into two forms of tests, namely tests and non-tests.The test is a tool to measure students' abilities in the form of questions, orders or instructions according to the learning objectives, while non-tests are measurements to get results from the form of student behavior. Thus, it can be concluded that an assessment instrument is a tool used to measure the ability of students to achieve learning competencies through tests and non-tests.

According to Matondang (2019), the requirements for quality assessment instruments must be valid, reliable, relevant, representative, practical, economical and have a level of difficulty. According to Arifin (Sofa, 2019), the requirements for quality assessment instruments must be valid, reliable, relevant, representative, practical, discriminatory, specific, and proportional. Meanwhile, Sudijono (2013) states that the requirements of a good assessment instrument are valid, reliable, practical, easy to measure, and easy to administer. Thus, it can be concluded that the requirements for quality assessment instruments are valid, reliable, relevant, representative, practical, proportional, sustainable, and meaningful which can provide an overview and feedback in learning.

The Development of Assessment Instruments Based on HOTS

According to Thomas & Thome, HOTS is a higher way of thinking than memorizing facts, expressing opinions, or applying rules, formulas, and procedures when solving problems (Nugroho, 2018:16). HOTS is a kind of thinking process that involves psychological activities, trying to explore complex, reflective and creative experiences, these experiences are carried out consciously to achieve the goal of acquiring knowledge including analysis, synthesis and evaluation of thinking levels. Rofiah et al., 2013: 17). Meanwhile, according to Resnick, HOTS is a complex thought process that involves the most basic mental activities, including describing material, drawing conclusions, building representations, analyzing, and establishing relationships (Ariyana et al., 2018). It can be concluded that HOTS are higher thinking skills such as reasoning, analytical skills, problem solving, critical thinking and doing based on facts.

In developing the assessment instrument, Anderson & Krathwohl followed the cognitive process, namely for the LOTS question categories between C1 (remembering), C2 (understanding), C3 (applying), while the HOTS question categories were among C4 (analyzing), C5 (evaluating), C6 (creating). According to Widana (Rodiana & Pahlevi, 2020) the steps to develop a HOTS-based assessment instrument include several steps, namely (1) analyzing basic competencies and indicators to make HOTS questions, (2) creating a grid of questions, (3) using a stimulus that interesting and contextual, (4) writing questions according to the grid, (5) making scoring guidelines and rules and answer keys. Meanwhile, according to Mardapi (Arifin, 2018) the steps in developing the HOTS question instrument include (1) compiling test specifications, (2) writing test questions, (3) studying test questions, (4) conducting test trials, (5) analyzing items questions, (6) improve the test, (7) assemble the test, (8) administer the test, (9) interpret the test results. It can be concluded that the main steps in developing the test in this study are analyzing core competencies and basic competencies for the test, then formulating clear instructional goals that will be assessed, making a grid or blueprint for assessment tools, making question cards, then the questions are reviewed by experts and instruments ready to be tested.

Definition of Critical Thinking

Critical thinking ability is thinking that is capable of high reasoning, reasonable, reflective and able to determine what is done and believed. This pattern of thinking is very important in the current era of education as a necessity in the future. Critical thinking aspect is one part of higher order thinking analysis (HOTS) which can shape students' mindset. In line with Ennis' opinion, critical thinking is a thinking process that analyzes, formulates, or solves problems with reasonable decisions about something. Therefore, it can be concluded that critical thinking is a student's way of thinking in determining something, formulating, and analyzing a problem critically and reliably.

Becoming a critical thinker requires awareness and skills in following the steps of critical thinking properly and correctly, although this step is not a full benchmark in improving critical thinking. The steps for critical thinking are recognizing problems, assessing relevant information correctly, being able to solve problems and drawing conclusions. While the indicators distributed to students in the form of a questionnaire as a benchmark for critical thinking according to Carole Wade are:

a. Activities to formulate questions
b. Limiting a problem
c. Testing the data
d. Analyze various opinions

e. Avoid emotional considerations
f. Oversimplification
g. Interpretation in various ways
h. Tolerance and ambiguity

Economics Learning Materials

The economics subject matter is in accordance with the basic competencies in the 2013 revised 2018 curriculum of economics subject syllabus, class XI IPS semester two containing LOTS:

a. 3.6 Analysing APBN and APBD in economic development.
b. 3.7 Analysing taxation in economic development.
c. 3.8 Describe international economic cooperation.
d. 3.9 Analysing international trade concepts and policies

While the basic competency analysis for HOTS questions is based on basic competencies in accordance with the syllabus for economic subjects in the 2013 revised 2018 curriculum, class XI IPS semester two containing HOTS:

a. 3.6 Analysing APBN and APBD in economic development.
b. 3.7 Analysing taxation in economic development.
c. 3.9 Analysing international trade concepts and policies

2 Research Methods

Development research using the Borg and Gall model with HOTS-based questions on economics class XI social studies. The purpose of this research is to meet whether or not the questions developed are based on HOTS and to find out whether the product is effective or not to improve students' critical thinking patterns. The research was carried out in the even semester of the 2020/2021 academic year. The technique of collecting data was through observation and validation questionnaires, evaluation experts and material expert validation, teacher response questionnaires, and student response questionnaires and following the stages of Borg and Gall development research. After the product is finished, then experimental research is carried out with a sample of 41 students of class XI IPS1 as the control class, and 42 students of class XI IPS2 as the experimental class. The purpose of this study is to improve students' critical thinking in economics learning by using HOTS-based assessment instruments.

3 Result and Discussion

The development of HOTS-based questions to improve students' critical thinking was developed through the Borg and Gall development model by following each development stage. In the results of the study conducted by the author, he first analysed the information according to the needs in the field. Based on the analysis of the needs of teachers and students, it was found that it was necessary to develop good and quality HOTS-based questions in accordance with the development procedure. Furthermore, the design is carried out in developing the product, starting from the stage of analysing basic competencies and indicators for each economic lesson in class XI IPS according to the economics syllabus.

After that, design the test grid to be made and the selection of indicators that meet the HOTS and LOTS questions. For the HOTS questions, there are 20 multiple choice questions covering C4 to C6 and the LOTS questions are 20 multiple choice questions covering C1 to C3. After the questions are designed and developed to the stage of question cards. Furthermore, the product can be given to the validator to review each instrument made in the form of suggestions, input or improvements. There are two validators in this study, namely material expert validation and evaluation expert validation. The results of the study from the validation of material experts related to the products developed had an average score of 91 percent in the very good category. And the validator provides suggestions and input in the form of learning objectives to be included in the question card.

Furthermore, the results of the evaluation expert validation related to the products developed have an average of 90 percent in the very good category, with suggestions and inputs given, namely the questions must contain factual data and are in accordance with the realities of everyday life. The small field test stage in the form of validating teacher responses related to the products developed had an average score of 93 percent in the very good category and did not provide suggestions or input. And the validation of student responses related to the product developed had a score of 95 percent in the very good category and there were no student input and suggestions. In a large group test involving a sample of 40 students to find out the results of the quantitative validation, including

validity and reliability tests, differentiating power tests, and testing the level of difficulty of each question. From the analysis of the questions, it was found that the results of the validity test of all questions were in the valid category with the assumption that the r-count was greater than the r-table, which was 0.312.

The results of the reliability test with an average of 0.90 are in very high criteria, the results of the average difficulty level test are in the medium category, and the differentiating power of the questions in the criteria is very good. After going through the large group test stage, so that the questions developed meet the elements of eligibility for HOTS-based questions on economics subjects for class XI Social Sciences that are good and of good quality which are ready to be distributed to students with a product of 20 multiple choice questions. In line with the large group test, the critical thinking questionnaire was first carried out by field testing to determine the feasibility of the instrument. The results of the critical thinking questionnaire test with 20 statements with a 5-point Likert scale are in the valid category.

From experimental research using pretest and posttest. For the pretest, the critical thinking questionnaire was first distributed to the two samples to determine the results of temporary critical thinking skills, then different treatment was given to the experimental class (XI IPS2) on HOTS questions and the control class (XI IPS1) on LOTS questions. The results of the pretest and posttest can be presented in the table below:

Table 1. Pre-test and Post-test Results

Class	Average Pretest	Average Posttest	N-Gain Score(%)
Experiment Class (XI IPS2)	61	83	52,2
Control Class (XI IPS1)	60	73	28

From table 1 above, the results of the experimental class pretest with an average score of 61 while the control class average score of 60. From the two samples, the students' critical thinking at the beginning was not much different for the two classes, both the experimental class and the control class. The results of the posttest with different treatments found the results that the experimental class had an average score of 83 and the control class had an average score of 72. Therefore, it can be seen that the comparison shows that the treatment, namely students who received HOTS questions had higher critical thinking scores than students with LOTS question treatment.

Furthermore, the normality test can be tested using the Kolmogorov-Smirnov technique, the data is declared normal if the probability value or sig > 0.05. The results of the normality test are presented in table 2 below:

Table 2. Normality Test

Class	Data	Sig	α	Keterangan
Experiment	Pretest	0,200	0,05	Normal Distributed Data
	Posttest	0,148	0,05	Normal Distributed Data
Control	Pretest	0,200	0,05	Normal Distributed Data
	Posttest	0,200	0,05	Normal Distributed Data

The normality test data for the pretest and posttest of the experimental group and the control group have a probability value > 0.05, namely, the data is normally distributed. Furthermore, a homogeneity test was conducted to determine whether the two groups of experimental and control pretest samples had the same variance value or not. It is said to have the same variance value if the sig level is > 0.05. The results of the homogeneity test are presented in table 3 below:

Table 3. Homogeneity Test

Levene Statistic	df1	df2	Sig.
1.912	3	156	.130

The results of the homogeneity test using the Levene method, the significance value of the experimental and control posttest results is a sig value of 0.130> 0.05, so it can be concluded that the two samples of the experimental and control groups are homogeneous. To determine the effectiveness of the HOTS-based assessment instrument, it can be calculated by interpreting the n-gain score category for the experimental group, 52.2% is in the less effective category. This is like Widya stated in her research (2019) that the HOTS assessment instrument to improve critical thinking skills is in the less effective category, which is an average of 45.6 percent on a scale of 100.

While the control group is 28% in the category ineffective. Therefore, HOTS-based questions are considered less effective in improving students' critical thinking in economics subjects.

4 Conclusion

a. The development of HOTS-based questions to improve students' critical thinking was developed through the Borg and Gall development model. The results of the material expert validation have an average score of 91 percent in the very good category and the evaluation expert validation has an average of 90 percent in the very good category, teacher response validation scores 93 percent in the very good category, and student response validation scores 95 percent are in the very good category.
b. The effectiveness of HOTS-based questions with the interpretation of the n-gain score category for the experimental group of 52.2% is in the less effective category, while the control group is 28% in the ineffective category. Therefore, HOTS-based questions are considered less effective in improving students' critical thinking in economics subjects.

Acknowledgement
The researcher would like to thank all those who have helped and supported the researcher in completing this research. Hopefully, the results of this research will be useful for many people, especially the authors and the place of research. The researcher also thanks to the supervisors and validation lecturers who have provided input and suggestions for the perfection of this research.

References

[1] Alec Fisher. 2009. *Berpikir Kritis Sebuah Pengantar*. Terj. Benyamin Hadinata. Jakarta: Erlangga
[2] Amirono, & Daryanto. (2016*). Evaluasi & Penilaian Pembelajaran Kurikulum 2013*. Yogyakarta: Gava Media.
[3] Anderson, L.W., dan Krathwohl, D.R. 2001. A Taxonomy for Learning, Teaching, and Assesing: A Revision of Bloom's Taxonom y of Educatioanl Objectives. New York: Addison Wesley Longman, In.
[4] Arikunto, Suharsimi. 2014. *Prosedur Penelitian Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta.
[5] Arifin, Zainal. 2009. *Evaluasi Pembelajaran*. Bandung: PT Remaja Rosdakarya.
[6] Arifin, Z. (2018). Pengembangan Instrumen Analisis Kesalahan Untuk Mengukur Critical Thinking Skill Berdasarkan Kompetensi Inti Kurikulum 2013. *Eduma : Mathematics Education Learning and Teaching*, *7*(1). https://doi.org/10.24235/eduma.v7i1.2734
[7] Ariyana, Y., Pudjiastuti, A., Bestary, R., & Zamroni. (2018). Buku Pegangan Keterampilan Berpikir Tingkat Tinggi Berbasi Zonasi. *Direktorat Jendral Guru Dan Tenaga Kependidikan*.
[8] Hasan, S. W., Auliah, A., & Herawati, N. (2020). Pengembangan Instrumen Penilaian Kemampuan Berpikir Kritis Siswa SMA. *Chemistry Education Review (CER)*, *3*(2), 185. https://doi.org/10.26858/cer.v3i2.13769
[9] Matondang, Zulkifli. Ely Djulia, Sriadhi dan Janner S. 2019. *Evaluasi Hasil Belajar*. Medan: Yayasan Kita Menulis
[10] Mangiante, Elaine Silva. 2013. Planning Science Instruction for Critical Thinking: Two Urban Elementary Teachers' Responses to a State Science Assessment. Journal Education Science, Vol 3: 222-258. www.mdpi.com/journal/ education.
[11] Mason. J., Leone B., and Kaye S,. 2010. *Thinking Mathematically:* Second Edition. Boston: Pearson Education University of Melbourne.
[12] Nugroho, Arifin. R. 2018. *Higer Order Thinking Skills*. Jakarta: Gramedia Widyasarana.
[13] Rodiana, S., & Pahlevi, T. (2020). Pengembangan Instrumen Penilaian Berbasis Higher Order Thinking Skills (HOTS) Pada Mata Pelajaran Kearsipan Jurusan OTKP di SMKN 1 Sooko Mojokerto. *Jurnal Administrasi Perkantoran (JPAP)*, *8*(1), 82–95. https://journal.unesa.ac.id/index.php/jpap/article/view/8115
[14] Rofiah, Emi., Nonoh Siti Aminah dan Elvin Yusliana Ekawati. 2013. Penyusunan Instrumen Tes Kemampuan Berpikir Tingkat Tinggi Fisika Pada Siswa SMP. Jurnal Pendidikan Fisika, Vol 1 (2): 17-22.
[15] Sudijono, Anas. 2013. *Pengantar Evaluasi Pendidikan*. Jakarta: Rajawali Pers.
[16] Sudjana, N. (2009). *Penilaian Hasil Proses Belajar Mengajar*. Bandung: Remaja Rosdakarya
[17] Sugiyono. 2014. Metode Penelitian Kuantitatif, Kualitatif, dan R&D. Bandung: ALFABETA
[18] Suryani. Y. 2015. *Kompetensi Pedagogik.* Jakarta: AZ-ZAHRA.