

Development of test instruments to measure high level thinking ability of two linear variable equations system

Anggit Prabowo*, Puspa, Fariz Setyawan

Universitas Ahmad Dahlan, Jl. Ringroad Selatan, Tamanan, Banguntapan, Bantul,
Daerah Istimewa Yogyakarta 55191 Indonesia

*Corresponding e-mail: anggit.prabowo@pmat.uad.ac.id

Abstract

This study aims to develop a test instrument to measure the high-level thinking ability of the two-variable linear equation system material. This research is a Research and Development using the ADDIE development model consisting of steps of analysis, design, development, implementation, and evaluation. The test subjects in this study were students of class VIII of SMP Negeri 3 Payung. The instruments used in this study were validation sheets and documentation. Data collection techniques through expert judgment and tests. This research has succeeded in developing a test instrument to measure the high-level thinking ability of two-variable linear equations system material. The test developed consists of 10 multiple choice items and 5 description items that have been declared valid by expert judgment. The trial results show from the difficulty index analysis of multiple-choice questions was obtained as much as 1 item in the easy category, 9 questions in the moderate category, the analysis of the difficulty index for the essay obtained by 2 items in the medium category, 3 questions in the difficult category. In the analysis of discriminant index of multiple-choice questions, there were 1 question in the bad category, 3 questions with enough categories, and 6 questions in the good category. In the analysis of the difference in power, 2 questions were categorized as bad, 1 question was categorized as sufficient, 2 questions were categorized as good. Based on the analysis of the function of the distractor, it was found that 1 question was not functioning properly, 9 questions were functioning properly. Finally, from the reliability analysis, it was obtained 0.73 for multiple choice questions and 0.716 for essay questions with high reliability interpretation.

Keywords: higher order thinking, systems of linear equations, tests

How to cite: Prabowo, A., Puspa, & Setyawan, F. (2020). Development of test instruments to measure high level thinking ability of two linear variable equations system. *International Journal on Education Insight*, 1(2), 91-100. DOI: <http://dx.doi.org/ije.v1i2.3540>

INTRODUCTION

Mathematics is the science that underlies various sciences as well as sciences that can develop thinking skills. Every development of modern science, social and technology cannot be separated from the language of mathematics. This underlies the importance of mathematics in student learning in schools as a whole and solving problems in their daily lives. Besides its relationship with various scientific disciplines and helping to develop logical and creative thinking, the government, in this case the National Education Office, attaches the objectives of Mathematics to MOEC Regulation Number 58 of 2016 concerning Guidelines for Mathematics Subjects (MOEC, 2016a).

Learning mathematics is not only intended for mastery of mathematical material as a science alone, but to achieve a more ideal goal, namely mastery of mathematical literacy which is needed to understand the world around it and

for success in life. Mathematics learning functions as a means to develop life skills. Life skills that can be cultivated through learning mathematics as stated in the MOEC Regulation Number 58 of 2016 (MOEC, 2016a) concerning Guidelines for Mathematics Subjects and the MOEC Regulation Number 21 of 2016 (MOEC, 2016b) regarding content standards include: (1) use the ability to think and reason in problem solving, (2) communicate ideas effectively, and (3) have attitudes and behaviours that are in accordance with the values of mathematics and learning, such as adhering to principles, being consistent, upholding diversity, respecting differences in opinion, being conscientious, tough, creative, and open.

Changing the curriculum from KTSP to Curriculum 2013 does not change the vision and goals of learning mathematics (Martina, 2017). However, the changes to the 2013 curriculum in schools that have been set by the government have not gone as expected. Because the teacher's role as a source of knowledge is still too dominant (teacher-centered) is not a student-centered, so it does not train students' higher order thinking skills. The thinking ability that develops in individuals as expected in the 2013 curriculum may not happen suddenly. Educational institutions as institutions that are responsible for managing and delivering education.

High Order of Thinking Skill is a student thinking process at a higher cognitive level developed from various cognitive concepts and methods and learning taxonomies such as problem-solving methods, bloom taxonomy, and learning, teaching and assessment taxonomies (Saputra, 2016). These high order of thinking skills include problem-solving skills, creative thinking skills, critical thinking skills, argumentation skills, and decision-making skills. The High Order of Thinking Skill includes critical, logical, reflective, metacognitive, and creative thinking. The 2013 curriculum also requires learning material to be metacognitive which requires students to be able to predict, design, and predict. In line with that the realm of HOTS, namely analyzing, is the ability to think in specifying aspects of a particular context; evaluating is the ability to think in making decisions based on existing facts / information; and creating is the ability to think in building ideas / ideas. So, High Order of Thinking Skill (HOTS) is an ability that students must have, so that when faced with questions that are able to measure higher order thinking skills, students are able to analyze, evaluate, and create questions.

Measuring students' higher order thinking skills certainly requires an assessment instrument in the form of a written test, in addition to being used to determine the profile of students' ability to think at higher levels, it is also a means of training students' ability to think at a higher level. The questions used as an exercise can be in the form of questions that test students in problem solving, critical thinking and creative thinking. So, to measure high-order thinking skills, an instrument in the form of a written test is needed to train students' thinking skills which include logical, systematic, critical, and creative thinking.

The form of the test, in particular High Order Thinking Skill, is a multiple choice test and an essay test. Multiple choice tests are used because to measure students' higher order thinking skills with a variety of answer choices, whether students will be fooled or not by an answer that is almost closer to the actual answer. While the essay test is used because it measures the ability to think higher and is able to reduce the likelihood of students guessing answers. The quality of the High Order Thinking Skill test is to train students by giving questions that invite students to analyze, evaluate, and be creative. Students'

thinking skills can be trained through student activities by giving various questions.

The competitiveness of Indonesian students, especially in mathematics, is still very low compared to students from other countries. This is shown by one of the international studies, namely the PISA (International Program for Student Assessment) test held by the OECD (Organization for Economic Co-operation and Development). PISA results (2015) show that Indonesia is ranked 69 out of 76 countries. Meanwhile, the results of the TIMSS (Trends in International Mathematics and Science Study) study show that Indonesian students are ranked 36 out of 49 countries in terms of carrying out scientific procedures, Indonesia's score in mathematics is 386, science 403 and reading 397. From these results Indonesia is still far away. left behind from several countries,

However, the importance of higher-order thinking skills is not in line with the habituation of these abilities. In formal education the ability to think at higher levels is more likely to be instilled at the tertiary level. At the junior high school level, schools rarely accustom their students to higher-order thinking. Mathematics learning tends to be specialized in the ability to solve problems with formulas in accordance with the procedures taught by the teacher. Based on the above problems, a study was conducted to develop a test instrument to measure higher order mathematical thinking skills. In this study, the test instrument developed was focused on class VIII students on the material of the two-variable linear equation system.

RESEARCH METHOD

The development model used is the ADDIE development model which stands for analysis, design, development, implementation, and evaluation (Branch, 2009). In accordance with the research model taken by the researcher, the development procedure consists of five stages. First step was analysis. It analyzed the development of the test instrument and analyze the feasibility of the test instrument requirements. The analysis stage was carried out by researchers by conducting interviews with mathematics teachers of SMP Negeri 3 Payung.

Second step was design. Researcher designed the question grid on the test instrument, the test instrument questions for the higher order thinking skills ability and the answer key to the test instrument. Third stage was test development. The test instrument developed was then consulted with the supervisor. Then, the results of the consultation are used as a reference for repair / revision of the test instruments. After that, expert lecturers validated the test instruments. Validation aims to determine the feasibility of the test instruments produced before being used to measure higher order thinking skills. The results of the validation are data to measure the validity of the test instrument, as well as suggestions or input from the validator. The test instrument that has been validated is then revised according to suggestions or input from expert lecturers.

Fourth stage was implementation. At this stage, test instrument that is produced at the development stage and then implemented or tested in the classroom. Fifth stage was evaluation. In the last stage the researcher carried out the evaluation, an evaluation was carried out related to the development of the test instrument. The researcher revised the product according to the evaluation results. The product in the form of an evaluation instrument as a result of this development was tried to 25 students. Data collection instruments were validation sheet and documentation. The data analysis technique was

performed using quantitative descriptive analysis according to the development procedure carried out. The steps to analyze the items characteristics (difficulty index, discrimination index, distractors, and reliability).

RESULTS AND DISCUSSION

This research was conducted in 5 steps.

Analysis

The analysis stage in this study includes needs analysis, curriculum analysis, and analysis of student characteristics.

Needs analysis

The results of this analysis were obtained through observation and interviews. Based on the results of observations and interviews conducted with a class VIII mathematics teacher at SMP Negeri 3 Payung, it was found that the teacher's role was still very dominant in classroom learning.

Student analysis

Student analysis is a study of student characteristics in accordance with the learning test development design. These characteristics include mathematical abilities and attitudes towards learning materials. Student analysis activities are focused on class IXA students as trial subjects because the material has been studied in class VIII. The average number of students in each class is 25 students. Based on observations and interviews from mathematics teachers, it can be seen that the mathematics knowledge of grade IX students of SMP Negeri 3 Payung varies. There are those with low, medium and high abilities. This shows that there is a factor of interest that each student has in different mathematics lessons.

Material analysis

Material analysis is a study to select and apply, detail and systematically arrange the relevant material for use in the test. Based on the curriculum analysis activities, it was found that the material to be used in developing the test instrument was in accordance with the material in the 2013 Curriculum for junior high school students' mathematics subjects. The material is a two-variable system of linear equations.

Furthermore, the material is seen from the Basic Competence. Based on this, indicators can be developed for each question, namely: (1) Presented a problem related to SPLDV, students can draw conclusions from the problem at hand, (2) Presented a problem related to SPLDV, students can determine the form of Two Variable Linear Equations, (3) Presented a problem related to the Two Variable Linear Equation System, students can solve these problems with the methods and methods that have been studied previously, (4) Presented a shopping receipt, students can determine the problem solving related to the Two Variable Linear Equation system, (5) Presented a conversation about daily life problems, students can determine solutions related to the Two Variable Linear Equation System, (6) Presented a picture of the problems of everyday life, students can determine solutions related to the Two Variable Linear Equation System, (7) A picture of the problems of everyday life is presented, as well as presented with the answers, students are asked to find the truth and provide arguments.

Design

This stage aims to design an identified assessment test instrument based on the results of the analysis stage. The test instruments to be designed consist of grid questions on the test instrument, test instrument questions on the ability of the higher order thinking skill and answer keys to the test instrument. The initial stage carried out by the researcher was designing a grid of high-order thinking skills test questions. After that the researcher also made test instrument questions, answer keys, validation sheets for the validator to check the validity of the high-order thinking skills questions. The test grid is designed to refer to the achievement indicators and the cognitive domain of each question. Designing test questions and answer keys is the most difficult stage for researchers. This is because in this activity the researcher has to design problems and possible student responses based on indicators of thinking ability of each question. This was done to obtain a test instrument product that could measure or determine students' high-level thinking skills.

Development

The development stage in this research is in the form of: (1) Developing test instruments in the form of test question grids, test instrument questions, and test instrument answers. (2) The assessment of the validation of the test instrument consisted of 2 validators, both of whom were material experts. The two validators are lecturers from Ahmad Dahlan University. After assessing the two validators, the test instrument is suitable for use. (3) The initial revision of the test instrument was carried out after the test instrument was assessed for its validity by the material expert.

Implementation

After the test instrument is declared feasible to be tested with revisions by material experts, the test instrument can be implemented in mathematics learning in schools. The implementation in this research is the process of testing the instructional test in learning activities. The trial was conducted in 2 meetings on Monday 27 & 28 July 2020 during the 1-2 lesson hours for 25 students of class IXA at SMP Negeri 3 Payung.

Evaluation

The purpose of this stage is to consider the quality of the design of the test instruments that have been developed. In the form of analysis of the validity of the contents of the questions, the reliability test, the level of difficulty and the differentiation of the developed instruments.

The results of the data analysis of the feasibility of the question instruments developed obtained the results of the validity, difficulty index, discrepancy, deceptive function, and reliability as follows.

Validity

Material expert validation data can be obtained from the results of filling out the validation sheet to two material experts.

The validation of material experts was carried out by the validator 1 on June 28, 2020. The instrument for validating this material consisted of 22 specific multiple-choice questions and 14 specific description questions. Comments and suggestions obtained on the validation of material experts are used as the basis for making revisions before the questions developed are tested on students.

Validation of material experts was carried out by validator 2 on July 16, 2020. The instrument for validating this material consists of 22 specific multiple-choice questions and 14 specific description questions. Comments and suggestions obtained on the validation of material experts are used as the basis for making revisions before the questions developed are tested on students.

Difficulty index

The test items can be said to be good if the test items have a difficulty level in the interval 0.16-0.85, this shows that the items are not too difficult and not too easy. The level of difficulty of the tests developed was also obtained from the data on the results of student work in the field test. Following are the results of the analysis of the level of difficulty on the higher order thinking skills test.

Table 1. Multiple-choice problem difficulty index

No.	Index of Difficulty	Criteria
1.	0.84	Easy
2.	0.36	Moderate
3.	0.64	Moderate
4.	0.4	Moderate
5.	0.36	Moderate
6.	0.36	Moderate
7.	0.52	Moderate
8.	0.48	Moderate
9.	0.56	Moderate
10.	0.4	Moderate

Based on Table 1 it is known that question number 1 has an easy level of difficulty, meaning that many students answered correctly on these questions. While questions number 2, 3, 4, 5, 6, 7, 8, 9, 10 have a difficulty level with the "medium" category, meaning that students who answer correctly and answer wrongly are balanced. According to the test instrument quality criteria in chapter 3, all test questions are good.

Table 2. Difficulty index questions in description

No.	Level of Difficulty	Criteria
1.	0.58	Moderate
2.	0.57	Moderate
3.	0.25	Hard
4.	0.28	Hard
5.	0.18	Hard

Based on Table 2 it is known that questions number 1 and 2 have a level of difficulty with the "medium" category, meaning that students who answer correctly and answer incorrectly are balanced. While questions 3, 4, and 5 have a difficult level of difficulty, meaning that few students are able to answer these questions. According to the test instrument quality criteria in chapter 3, all test questions are good.

Discriminant index

The items on the test instrument can be said to be good if the test items have the smallest distinguishing power of 0.2. This indicates that the items have sufficiently minimal distinguishing power. The following are the results of the analysis of the distinguishing power of the mathematics test instrument.

Table 3. Discriminant index of multiple-choice questions

No.	Discernment	Criteria
1.	0.33	Enough
2.	0.53	Well
3.	0.11	Ugly
4.	0.29	Enough
5.	0.53	Well
6.	0.37	Enough
7.	0.52	Well
8.	0.44	Well
9.	0.60	Well
10.	0.45	Well

Based on Table 3, it is known that questions number 2, 5, 7, 8, 9, 10 have distinguishing power with the "good" category, meaning that the question is good for distinguishing high-skilled test takers from low-ability test participants. Questions number 1, 4, and 6 have distinguishing power with the "sufficient" category, meaning that the questions are good enough to distinguish high-ability test takers from low-ability test takers. Meanwhile, question number 3 has "bad" distinguishing power, meaning that it is unable to distinguish high-ability test takers from low-ability test takers.

Table 4. Discriminant index of essay test

No.	Discernment	Criteria
1.	0.51	Well
2.	0.71	Well
3.	0.09	Ugly
4.	0.25	Enough
5.	0.08	Ugly

Based on Table 4, it is known that questions number 1 and 2 have distinguishing power with the "good" category, meaning that these questions are good for distinguishing high-ability test participants from low-ability test participants. Question number 4 has a distinguishing power with the "sufficient" category, meaning that the question is good enough to differentiate between high-ability test takers and low-ability test takers. Meanwhile, questions 3, 5 have "bad" distinguishing power, meaning that the questions cannot differentiate between participants. a high-performing test with a low-ability test taker.

Effectiveness of distractors

Iskandar & Rizal (2018) argues that in multiple choice questions there are alternative answers which are distractors. Good item, the trickster will be chosen evenly by students who answer wrong.

Table 5. Effectiveness of distractors

Item Question Number	Effectiveness of Distractors
1, 2, 3, 5, 6, 7, 8, 9, 10	Effective
4	Not effective

Based on Table 5, it can be seen that number 4 the function of the distractor does not work, meaning that the deceiver is not chosen by the students, indicating that the deceiver is bad. Questions number 1, 2, 3, 5, 6, 7, 8, 9, and 10 function very well, meaning that the deceiver has a great appeal so that the test taker chooses the cheater's answer.

Reliability

Based on the results of field trials (students) involving students in class IXA SMP Negeri 3 Payung with the number of students in that class is 25. Based on the data analysis, the reliability of the test obtained was 0.73 in multiple choice questions and 0.72 in essay questions with interpretation. high reliability.

The development of test instruments to measure higher order thinking skills is very important to instill students' higher order thinking skills. Higher order thinking skills can be viewed from the point of view of two major fields of science, philosophy and psychology. How sources are defined and how they are found. Psychology with science is philosophy with humanity. Both contribute to higher order thinking skills (Lewis & Smith, 1993). According to the National Research Council, National Science Teachers Association, and Standard for Professional Development in Schools, science teachers are expected to apply higher order thinking to their students (Barak, 2007; Dori & Belcher, 2005; Tobin et al., 1990). Testing students' high-order thinking skills can improve higher-order thinking skills.

Learning higher order thinking is generally accepted as one of the goals of general education (Ivie, 1998). In learning activities, high-level learning is described as a function of the interaction between cognitive strategies, metacognition, and certain domains during problem problems (Young, 1997). High-level thinking is a non-algorithmic and complex type of thinking that often solves problems with more than one solution (Barak & Dori, 2009). The concept of higher order thinking is cared for by Bloom's taxonomy of educational goals (Scully, 2017). Testing students' high-order thinking skills can improve higher-order thinking skills (Barak & Dori, 2009).

This research has succeeded in developing a test instrument to measure the high-level thinking ability of two-variable linear equations system material. The test developed consisted of 10 multiple choice items and 5 descriptions that were declared valid by the expert's report. The trial results showed that from the analysis of the difficulty index of multiple choice questions, 1 item was obtained in the easy category, 9 items in the medium category, the analysis of the difficulty index for the questions in description obtained 2 questions in the medium category, 3 questions in the difficult category; In the analysis of the difference in power of multiple choice questions, there were 1 question in the bad category, 3 questions in the sufficient category, and 6 questions in the good category, in the analysis of the difference in power of the questions in the description, there were 2 questions in the bad category, 1 question with enough categories, 2 questions with good category; based on the analysis of the function of the distractor, it was found that 1 question was not functioning properly, 9 questions were functioning properly; and from the reliability analysis obtained

0.73 for multiple choice questions and 0.716 for essay questions with high reliability interpretation.

CONCLUSION

The analysis of the feasibility of the test instrument to measure the high-level thinking ability of two-variable linear equation system material at SMP Negeri 3 Payung shows that of the 15 questions, (1) The validity is feasible in terms of the results of the material expert's assessment; (2) The test instrument reliability coefficient was 0.73 for multiple-choice questions and 0.72 for essay questions; (3) There are 1 easy and 9 medium items for multiple-choice questions. For the essay questions, there are 2 moderate items and 3 difficult items; (4) There are 9 items having effective and 1 item having effective distractors; (5) There are 1 item having bad discriminant, 4 items having good discriminant, and 3 items having medium discriminant for multiple-choice questions. Meanwhile, for the essay, there are 2 items having good discriminant, 2 item having bad discriminant, and 1 item having moderate discriminant.

In this study, the test instrument conducted in this researcher focused on mathematics subjects, the two-variable linear equation system material. This instrument can be used as a reference for teachers in developing HOTS questions for material on two-variable linear equation systems. It can also be used as items bank for teachers.

ACKNOWLEDGEMENT

The author would like to thank Universitas Ahmad Dahlan who has funded this research and SEAMEO QITEP in Mathematics who has been the colleague of the research.

REFERENCES

- Barak, M. (2007). Transitions from traditional to ICT-enhanced learning environments in undergraduate chemistry courses. *Computers & Education*, 48, 30-43.
- Barak, M., & Dori, Y. J. (2009). Enhancing higher order thinking skills among interface science teachers via embedded assessment. *Journal of Science Teacher Education*, 20(5), 459-474.
- Branch, R. M. (2009). *Instructional Design: The ADDIE Approach*. New York: Springer.
- Dori, Y. J., & Belcher, J. W. (2005). How does technology-enabled active learning affect students' understanding of scientific concepts? *The Journal of the Learning Sciences*, 14, 243-279.
- Iskandar, A., & Rizal, M. (2018). Analisis kualitas soal di perguruan tinggi berbasis aplikasi TAP. *Jurnal Penelitian dan Evaluasi Pendidikan*, 22(1), 12-23.
- Ivie, S. D. (1998). Ausubel's learning theory: An approach to teaching higher order thinking skills. *The High School Journal*, 82(1), 35-42.
- Lewis, A., & Smith, D. (1993). Defining Higher Order Thinking. *Theory Into Practice*, 32(3), 131-137.
- Martina. (2017). Pengembangan instrumen tes higher order thinking skill (HOTS) pokok bahasan sistem persamaan linear dua variabel dan teorema Pythagoras kelas VIII SMP Citra Samata Kab. Gowa. *Skripsi*. Makassar: Universitas Islam Negeri Alauddin Makassar.
- MOEC. (2016a). *Permendikbud Nomor 58 Tahun 2016 Tentang Rincian Tugas Unit Kerja Di Lingkungan Lembaga Pengembangan Dan Pemberdayaan*

- Pendidik Dan Tenaga Kependidikan Bidang Kelautan, Perikanan, Teknologi Informasi Dan Komunikasi*. Jakarta: Ministry of Education and Culture.
- MOEC. (2016b). *Peraturan Menteri Pendidikan dan Kebudayaan RI Nomor 21 Tahun 2016 tentang Standar Isi Jenjang Pendidikan Dasar dan Menengah*. Jakarta: Ministry of Education and Culture.
- Saputra, H. (2016). *Pengembangan Mutu Pendidikan Menuju Era Global: Penguatan Mutu Pembelajaran dengan Penerapan HOTS (High Order Thinking Skills)*. Bandung: SMILE's Publishing.
- Scully, D. (2017). Constructing multiple-choice items to measure higher-order thinking. *Practical Assessment, Research & Evaluation*, 22(4), 1-13.
- Tobin, K., Kahle, J., & Fraser, B. (1990). *Windows into science classrooms: Problems associated with higher-level cognitive learning*. London: Falmer.
- Young, A. (1997). Higher-order learning and thinking: What is it and how is it taught? *Educational Technology*, 37(4), 38-41.