Jurnal Inovasi Pembelajaran Fisika
(INPAFI)
Available online http://jurnal.unimed.ac.id/2012/index.php/inpafi
e-issn: 2549 – 8258, p-issn 2337 – 4624

DEVELOPMENT OF OBJECTIVE TEST PHYSICS HIGHER ORDER THINKING (HOT) FOR THE SUBJECT KINEMATICS IN SENIOR HIGH SCHOOL

Daulatta Ras Surbakti dan Sahyar
Department Of Physics Education Faculty Of Mathematic and Science State University Of Medan
daulattaras@gmail.com, sahyarpasca@gmail.com
Diterima: Maret 2022. Disetujui: April 2022. Dipublikasikan: Mei 2022

ABSTRACT

This research aims to develop objective test physics of higher order thinking (HOT) in kinematics topics in senior high school that has a good standard of qualification test in terms of validity, realibility, difficulty level, discriminantion power and effectiveness distractor. This research adopts the development model of ADDIE which is five stages of development namely, analyze, design, development, implementation and evaluation. This reseacrh was conducted on 65 students of Sains Eleven at Senior High School 2 Kisaran. This objective test produces 50 multiple choice test and has been tested twice. The second test is selected from the first test where there the questions were valid.This research shows the result in the first test 31 questions (62%) are valid, 19 questions (38%) are invalid and the second trials 26 questions (83,8%) valid, 5 questions (19,2%) are invalid. The reliablity value is 0.893 and 0.740. This shows that the objective test is reliable. While the test for the quantity of items consists level difficulty of the question, discriminantion power and effectiveness distractor that has the quality of the question, HOT-based questions. Based on the results of the second trials it was found that 26 questions can be received and stored in the Objective Test bank Physics Higher Order Thinking (HOT). So, the efforts to support the learning process continue to be needed which is useful for improving higher order thinking in students.

Keywords: *objective test HOT, Kinematics, validity, reliability*

INTRODUCTION

Education is the main subject in the development of human resorces and society as future assets that are essential for the suitainability of human civilization in the world, one of which is science education. Science education aims to prepare students for real life. One important element in teaching science, is the actual implications of scientific inquiry in the classroom to ensure practice in real life (Alabdulkareem, 2017).

Science education is an important part of education system in various countries including Indonesia, but based on the result of the evaluation of the quality of science education throughout the world in 2015 by the Trend in International Mathematics and Science Study (TIMSS) shows that Indonesian elementary and secondary level students are on the order of 36 out of 49 participating countries (Mullis et al., 2016)

According to the program for International Students Assessment (PISA),

Nadlir (2018) explain that Indonesian in 2015 was ranked 64th in the Organization for EconomicCooperationand Development(OECD) which has 72 countries. In this regard, the goverment continues to make evaluations in the field of education in Indonesia, namely by improving existing deficency. The deficency in the implementation of the curiculum, the goverment continues to review the curriculum taht is suitable to be applied in Indonesia. For this reason, the 2013 curriculum is currently being implemented which aioms to improving the quality of education in Indonesia. In the implementation of the 2013 curriculum, Widiyanto (2016) said that the Ministry of Education and Culture continues to make revisons to the curriculum and the last revision was carried out in 2018 which produced instruments to faciliatate eduvators in assessing student performance. In the 2013 curriculum, the Ministry of Education and Culture also made improvements to content standards and assessment standards, both of them focused on higher order thinking skills (HOTS) (Astutik, 2016).

Based on the TIMSS results, teachers are currently expected to be able to arrange higher order thingking (HOT) questions, which are questions that are able to express higher cognitive levels. However, to compile HOT questions that are still many teachers who do not understand and master them, both the characteristics of HOT questions and how to turn ordinary questions into HOT questions. Even the teacher competency test which has recently been carried out leads to questions that are included in the HOT questions category.

High school Physics Teachers are important to be trained to develop questions of HOT. This is in line with the recommendations contained in PP No.19/225 Article 19 Paragraph 1 which states that the learning process in educational units in carried our inspiratively, interactively, fun, challenging and motivates students to participate actively. Provision of HOT questions aims to be able to create learning that makes students challenged to think and reasoning.

Dewey (1993) was the first educator who differentiated levels of thinking. He desdcribed thinking as a sequenced chaining of events that moves from reflection to inquiry and into critical though process, and lead to conclusion that can be substantiated by more than personal beliefs and images.

Newmann (1998) reminds us that hugher-order thinking implies a challenge and expended use of the mind, while lower order thinking takes place during routine, mechanistic applications and constraints on the mind. He continues stating that that the challenging and expanded use of the the mind is achived through interpeting, analyzing and manipulating information to solve a problem, because a problems cannot be resolved through mechanistic application of previously learned knowledge.

It's pertinent to add the importance of the ability to recognize the context of the situation in the ability to recognize, and apply and elaborate knowledge for that is situation (Kinget et al., 1998). Transfer is the higher order thinking skill of being able to apply acquired knowledge to new situations, or new context, including across knowledge domains (Barak et al., 2007).

Newmann's statments cited by Abosalem (2016) higher order thinking skills can be defined as a challenge for students to interpret, analyze, or manipulate information. While (King, Goodson, & Rohani , 2011) say higher order thinking skills are the ability to think that activated when students face unfamiliar problems, dilemmas over the answers to a questions, in these situations students do not only use memorizing solutions but required to think critically and creatively in solving problems.

Students have difficulty in working on higher order thinking skills questions because they are not used to working on higher order thinking skills based questions. While Anderson, et al. (2001) say that the cognitve domains of remember, understand and apply are categorized as Lower Order Thinking Skills (LOTS) and the cognitive domains are analyze, evaluate and create categorized as Higher Order Thinking Skills (HOTS).

HOTS (High Order Thinking Skill) Lerning Development. States that competence

of thinking can be classified according to Bloom's Taxonomy, as in Table 1 below.

Table 1. Classification of thinking competencies according to Bloom's Taxonomy

Bloom's Taxonomy	Levels of Thinking	Review
Knowledge (C1)	Lower-order	Remembering
Comprehension (C2)	Lower-order	Understand
Application (C3)	Lower-order	Apply
Analysis (C4)	Higher-order	Analyze
Synthesis (C5)	Higher-order	Create
Evaluation (C6)	Higher-order	Evaluate

Schraw et al. (2011) classifies bloom's thinking skill into two categories that is Lower Order Thinkinh Skill which consists of knowledge, understanding and applicatio. Higher Order Thinking Skills which consist of analysis, synthetic and evaluation.

Sani (2019) classifies the HOTS betwen HOT into different categories that is HOTS (high order thinking skill) which consists of critical thinking, creativie thinking, problem solving and make decision. High order thinking (HOT) whicj consists of analyze, evaluation and created. Description of differences can be seen in table 2.

Table 2. Difference of HOT between HOTS

HOT	HOTS
Analyze	Critical Thinking
Evaluation	Creatif Thinking
Created	Problem Solving
	Make Desicion

Therefore, based on these problems researchers want to develop objective test in the form of higher order thinking based questions in order to hepl education improve their higher order thinking in students and help students understand and solve various problems with higher order thinking based assessment on physics subjects using the development model ADDIE, to determine the feasibility of higher order thinking skills based test objective on physics. This development research is expected to be used as input and contribution to the world of education, especially regarding the development of objective test in the form of higher order thinking based questions.

RESEARCH METHODS

This research is a development research. The developed product is instrument assessment to train student's higher order thinking (HOT). Development type is adapted from ADDIE type which consists of 4 development steps. The steps of ADDIE type, which consist of : 1) analyze, 2) design, 3)development, 4) implementation and 5) evaluation.

1. Analyze
   The analyze step in this research is as follow : 1)conduct initial identification in the form of material, 2) determine the research problems, 3) determine the products to be develop to overcome research problems, 4) review the literature (pre-existing products)
2. Design
   Design of this research is as follow : 1) arrange the indicator of instrument assessment towards HOT, 2) determine the instrument validity through the helping of physicist's test to validate the instrument made, 3) do instrument revision based on validastor's suggestion, 4) do limited try out (small class), 5) determine distinguishing power, difficulty level, and reliability of question items, 6) do product revision based on the result small class test, 7) do field try out (big class)
3. Development
   a. Determine the purpose of instrument assessment that is to train high school student;s high order thinking (HOT)
   b. Arrangement of instrument assessment form. Instrument assessment which developed is questions drill HOT, Multiple choice , and the answer key based on the indicatorof KD and HOT
   c. Validationn of HOT question items. Valid or deserve questions used will be measured based on lecturer's assessment. Data collecting instrument used by questionnaire validation

d. Revision of question items and arrangement of questions. Result of validation by validator which has been got is used for revising developed question items
e. Do the limited try out (small class)
f. Revision of question items which has been revised is limited try out that have been known reliability, distingushing power and difficulty level.

4. Implementation
Then, After doing field try out, is doing HOT arrangement questions which can examine stedent's HOT indeed. The arrangement questions are tried out again to see if it is really effective in measuring student's HOT.
5. Evaluation

a. Conduct evaluation according to the results of data from field trials and after that product has been produced.

RESULT AND DISCUSSION

a. Research Results

The product of the developmental study is a valid and reliable HOT test instrument, consisting of multiple-choice test items for senior school physics. The istrument development passes two assesment phases. The first phase is to assess the validity of the instrument, conducted by three experts of physics education. The seconds involves a small class try-out with 25 testees and a big class try-out with 40 testees.

Validation by experts is to look at the contents of the initial product and obtain feedbacks for revising the first draft. In the process, the experts are given the taable of the specification of the test, the test items, and the evaluation sheets. Data of the experts' evaluation are subjected to the Aiken's V formula to find the content validity coefficient. The result can be seen in Table 3.

Tabel 3. Results of Experts' validation

Item Number	Total	Criteria
2,3,4,5,7,9,10,11,12, 13,14,15,16,17,19, 20,21,22,23,24,25,27, 28,29,30,31,32,33, 34,35,36,37,38,39, 40,41,42,43,44,45,46, 47,48,49,50.	45	Good to be used
1,6,8,18,26.	5	Need revision/ delection

In table , it can be seen that, out of 50 test items, 45 are feasible for use and 5 needs revision or deletion.

After that the realibility in this research test in small class was calculated using the Kuder-Richardson 20 formula (KR-20) and it is known that the realibility of the questions develope is 0,89 (high category).

Table 4. Difficulty levels of small class test

Category	Item Number	Total
0.71< P ≤1.00 (Easy)	9,10,13,19,24.	5
0.31≤P ≤0.70 (Medium)	1,2,3,4,5,6,7,11, 12,15,16,17,18, 20,21,23,25,27, 28,29,30,31,32, 33,35,36,37,38, 39,40,41,42,43, 44,45,46,47,48, 49,50.	40
P < 0.30 (Difficult)	8,14,22,26,24.	5

Table 4 shows that 40 items (80%) have dificulty level in the medium category. Table 5 shows that five items (10%) have a discriminating power of the medium category.

Table 5. Discrimination power of small class test

Category	Item Number	Total
0.71< P ≤1.00	9,10,13,19,24.	5

(Easy)		
0.31≤P ≤0.70 (Medium)	1,2,3,4,5,6,7,11, 12,15,16,17,18, 20,21,23,25,27, 28,29,30,31,32, 33,35,36,37,38, 39,40,41,42,43, 44,45,46,47,48, 49,50.	40
P < 0.30 (Difficult)	8,14,22,26,24.	5

Efectiveness distractor it is clear that distractor distribution of all of the test items is functioning, it means that all the distractors are chosen by 5% of the testees. Based on the results of the analyses of the item characteristics above, the number of items that are accepted and rejected can be seen in Table 6.

In Table 6, a total of 31 items (62%) are accepted and 19 (38%) are rejected. The accepted are than reformatted to become the implementation class and test to big class again.

Table 6. Interpretation of Small Class Test

Category	Item Number	Total	Percentage
Accepted	1,2,3,4,5,6,8, 11,12,14,15, 16,18,20,21, 22,26,28,31, 32,33,35,36, 37,39,41,44, 46,48,48,50.	31	62%
Rejected	7,9,10,13,17,1 9,23,24,25,27, 29,30,34,38,40 ,42,43,45,47.	19	38%

From the analysis obtained not all the istruments that have good criteria so that need a revision for 31 question that are valid. In general, the thirty one questions need revision beecause : (1) difficulty level is easy and hard, (2) discriminantion power is bad and (3) efectiveness of distractor is not good.

Implementation

The implementation of this study is big class test it concluded 31 items and involving of 40 testess.

Realibility in this research test in big class was calculated using the Kuder-Richardson 20 formula (KR-20) and it is known that the realibility of the questions develope is 0,74 (high category).

Table 7. Difficulty levels of big class test

Category	Item Number	Total
0.71< P ≤1.00 (Easy)	1,2,20,23,26,31.	6
0.31≤P ≤0.70 (Medium)	3,4,6,7,8,9,10,12, 13,14,15,16,17,18, 19,21,22,24,25,28, 29,30.	22
P < 0.30 (Difficult)	5,11,17.	3

Table 7 shows that 22 items (70,9%) have the difficulty level in the medium category. Meanwhile, table shows that 8 items (22,58%) have a discriminating power of the medium category.

Table 8. Discrimination power of small class test

Category	Item Number	Total
DP<0.19 (poor)	2,6,7,20,23,27.	6
0.20≤DP<0.40 (Medium)	8,13,11,15,18, 28,31.	7
0.40≤DP<0.70 (Good)	1,3,5,10,12,14, 17,21,29,30.	10
0.70≤DP≤1.00 (Very Good)	4,9,16,19,22,24, 25,26.	8

Efectiveness distractor it is clear that distractor distribution of all of the test items is functioning, it means that all the distractors are

chosen by 5% of the testees. Based on the results of the analyses of the item characteristics above, the number of items that are accepted and rejected can be seen in Table 9.

In Table 9, a total of 26 items (83.8%) are accepted and 5 (19.2%) are rejected. The accepted are than reformatted to become the final product test instrument of HOT in terms of the test validity.

Tabel 9. Interpretation of small class test

Category	Item Number	Total	Percentage
Accepted	1,3,4,5,6,8,9, 10,11,12,13, 14,16,17,18, 19,21,22,24, 25,26,27,28, 29,30,31.	26	83.80%
Rejected	2,7,15,20,23.	5	19.20%

The final product revision is conducted to obatain a test instrument that is valid and reliable. Revision is done by looking at the results of evaluation in the analysis of results.

b. Final Product Discussion

The product of the study is valid and reliable test using high order thinking instrument. It is a fact that, up to the present time, no effort has been done for evidance of test validity and reliability. The development of the instrument begins with the review HOTS which, according to Harahap and Sahyar (2019), consist of the validity and reliablility is feasible.

It is followed by formulating the itenms indicators and writing of the test items. Then, the test items are subjected to content validation through expert judgement. Before being administered in the field try-out, the items are subjected to a limited-scale-try-out for realibility. The field try-out involves 40 students. Finally, item analysis and reliability estimation are conducted. The test instrument development has been conducted following the standard procedure and found that the test is valid and reliable.

The test items developed in the study are those of multiple-choices type. The development of the istrument begins with the review HOTS which, according to Ramadhani and Sahyar (2019) the istrument is feasible for use for measuring the HOTS because it was reliable. According to opinions and research result from experts, a multiple-choice test can be used to measure HOTS (Budiman & Jailali, 2014). It suggested that the format of the HOT test items consist of an introduction followed by response option.

CONCLUSIONS AND SUGGESTIONS

The conclussions of this research stated that there had been developed an instrument assessment of HOT with HOT indicators in kinematics based on cognitive dimension process in form of analyzing ability (C4), evaluating (C5), and creating (C6). The validity of the test is indicated by the experts' judgement showing that the test is good to be used in the aspects of format, content and language. The result shows the reliability coefficicient of 0.74 (good category), with average score of 0.59 for difficulty level (medium category), discriminatio power of 0.28 (medium category), and funtioning distractors.

Based on the conclusion of the study, it suggested that futher research is conducted by analyzing using the more modern method. This will expected be able to calculate the items of each items.

REFERENCES

Abosalem, Y. (2016). Assessment Techniques and Students' Higher-Order Thinking Skills. International Journal of Secondary Education, 4(1), 1.

Alabdulkareem. S. A. (2017). Saudi Science's Teacher's Perpeptions of Implementing Inquiry in Science Class. Journal of Education and Training Studies, 5, (12), 67-68

Anderson, L. W., Kratwohl, D. R., Airasian, P.W., A, K., Cruikshank, Meyer, R. E., Wittrock, M. C. (2001). A Taxonomy for Learning Teaching and Assessing (a

Bridged). New York : Addison Wesley Logman.

Astutik, P. Pudji. (2016). Integrasi Penguatan Pendidikan Karakter (PPK) dan High Order Thinking Skills (HOTS) dalam pembelajaran Tematik SD. In Seminar Nasional Pendidikan (pp. 343-354). Malang : Fakultas Ilmu Pendidikan Universitas Negeri Malang

Barak, M., David, B. C., & Uri, Z. (2007). Purposely teaching for the promotion of higher-order thinking skills: A case of critical thinking. Research in Science Eduaction, 37 (4), 353-369.

Budiman, A., & Jailani. (2014). Pengembangan instrumen aesmen higher order thinking skill (HOTS) pada mata pelajaran matematika SMP kelas VIII semester 1. Jurnal Riset Pendidikan Matematika, 1(2), 139-150. https://doi.org/10/21831/jrpm.vli2.2671

Dewey , J. (1993). How to think : A restatement of the relation of reflective thinking to the educative process. Boston : D. C. Health and Company

Hararap, W. Y and Sahyar. (2019). Pengembangan Tes Objektif HOTS Materi Suhu dan Kalor di SMA/MA. Jurnal Ikatan Alumni Fisika Universitas Negeri Medan 5(3) : 7-14

King, F. J., Goodson., & Rohani, F. (1998). Higher order thinking skills : Definition, teaching strategies, assessment. Tallahassee, FL : Center for Advancement of Learning and Assessment. Retrieved from hhtp://www.cala.fsu.edu/files/higher order thinking skills.pdf

King,F., Goodson, L., & Rohani, F. (2011). Higher Order Thinking Skills : Definitions, strategies, assessment. Center for advancement of learning and assesment. Tallahasse : FL : Florida State University. Retried from www.cala.fsu.edu

Mullis, I. V.S., Martin, M. O., Foy, P., and Hooper, M,. (2016). TIMSS 2015 International Results in Science. Retrieved from Boston College, TIMSS & PIRLS International Study Center

Nadlir, M. (2018). Komnas HAM Catat 4 Kondisi Darurat Pendidikan Indonesia. Retried November 14, 2018, from https://nasional.kompas.com/read/2018/05/02/12581141/komnas-ham-catat-4-kondisi-darurat-pendidikan-indonesia

Newmann, F. M. (1998). Higher order thinking in the high school curriculum. NASSP Bulleting, 72 (508), 58-64.

Ramadhani, T and Sahyar. (2019). Pengembangan Tes ObjektifFisika HOTS Pada Pokok Bahasan Fluida Untuk SMA/MA. Jurnal Ikatan Alumni Fisika Universitas Negeri Medan. 5(4): 40-46

Sani, R. A. (2019). Pembelajaran Berbasis HOTS (Higher Order Thinking), revisi edition. Tanggerang : Tira Smart

Schraw, Gregory, Robinson, DH. (2011). Assessment of Higher Order Thinking Skills. America : Information Age Publishing.

Widiyanto, N. (2016). Revisi Kurikulum 2013, Guru Lebih Dimudahkan. Retrieved November 14, 2018, from https://www.kemdikbud.go.id/main/blog/2016/06/revisi-kurikulum-2013-guru-lebih-dimudahkan