P-ISSN xxxx-xxxx || E-ISSN xxxx-xxxx
Indonesian Physics Education Research
Volume 5, No. 1, June 2024, pp. 24-38

DEVELOPMENT OF HOTS-BASED TEST INSTRUMENTS TO MEASURE STUDENTS' CRITICAL THINKING SKILLS MAJOR MATERIALS AND UNIT

Frisma Winda Afifah Panuqih[1] and Sabani[2]
[1] Universitas Negeri Medan, Medan 20113, Indonesia
[2]Universitas Negeri Medan, Medan 20113, Indonesia
Emai: frisma@mhs.unimed.ac.id

Abstract

This study aims to develop a HOTS-based test instrument in measuring students' critical thinking skills on the material of quantities and units at SMAN 2 Rambah Hilir which has met the criteria for testing the instrument's feasibility including content validity, item validity, reliability testing, level of difficulty, discriminating power, and student response test. This research is a research development or Research and Development (R & D) using the ADDIE (*Analysis, Design, Development, Implementation, and Evaluation) model.* This research was conducted at SMA Negeri 2 Rambah Hilir with the research subjects of class X MIPA SMA Negeri 2 Rambah Hilir. Based on the data obtained from content validity with the Lawshe technique, 10 questions were obtained that could be used. Meanwhile, from testing the validity of the items, there are 8 valid questions and 2 invalid questions. The reliability, test obtained a value of 0.543390 with moderate correlation criteria. For the level of difficulty, scores ranged from 0.297 to 0.70 with the criteria that 1 item was said to be easy, 1 item was said to be difficult, and 8 items were said to be moderate. While different power derives a value of -1,8- 2.5 with a category 2 a matter of saying ugly and the two problems said to be good at all. Meanwhile, to test students' response was obtained by 78,1 % in both categories. The conclusion from the overall test results is that the HOTS-based test instrument is feasible to use to measure critical thinking skills.

Keywords: Development of Test Instruments, Critical Thinking Skills, Quantities and Units

Learning in the 2013 curriculum requires a student-centered learning process. In the 2013 curriculum, students are required to be able to observe, ask questions, reason, communicate and apply what they have learned at school. Students can use what they have learned, but the premise is that they must be guided to solve problems, find problems and draw inspiration from them. In this case, students' advanced thinking skills play an important role in the success of achieving these educational goals. As written in the journal (Soemanto, 2012; Budiningsih, 2012), children aged 11 years and over already have mature abstract thinking, so the thinking skills needed are *high-order thinking skills* (HOTS).

Thinking skills are divided into lower-order thinking skills (LOT) including memory, understanding, and application, and advanced thinking skills (HOT) including analytical (analysis), evaluation (assessment), and creative (create) skills. Meanwhile, higher thinking skills, including the ability to analyze, evaluate, and be creative are very much needed for the progress of learning, especially in physics class.

When individuals face unknown problems, challenging problems, or uncertainties, HOTS will develop. According to Sani (2019), if a person stores information and acquires new information, then connects, compiles, and develops that information to achieve

goals or obtain possible solutions to chaotic situations, forward-thinking occurs. HOTS is very important in the process of learning science, one of which is physics. This is in line with the views of Pratama and Istiyono (2015) which state that the realization of scientific learning objectives is not only determined by the concept but rather on the influence of mentoring learning, one of which is HOTS.

Learning physics is a natural science that studies the basic elements that make up the universe, natural phenomena, and all the interactions that occur in it. According to Pratama and Istiyono (2015), physics is a science that studies natural phenomena. The rapid development of science and technology as well as discoveries in the field of science and technology make physics one of the important subjects mastered by students. More advanced thinking skills are needed to study physics material to understand the material in-depth and be able to solve problems in physics.

Higher-order thinking skills in students can be measured by conducting an assessment. This is confirmed by Hanifah (2019), which states that one way to find out whether students already have higher-order thinking skills is by conducting an assessment. Assessment in the form of tests can be used to hone students' thinking skills, and affect determining students' thinking skills. (Abdul Malik 2018; Law Rosidin 2018; Chandra Ertikanto) Higher order thinking skills have become one of the key points in learning physics. Knowledgeability requires students to understand, apply and analyze facts, concepts, processes, and metacognitive knowledge based on their curiosity about science. Likewise, with core competencies, students are expected to be able to develop skills, reason, and perform in specific and abstract fields related to the development of independent learning in schools, be able to show their creativity effectively, and be able to use science-based methods.

SMAN 2 Rambah Hilir is one of the many schools that have implemented the 2013 curriculum. Based on the results of interviews, in compiling items the teacher has not fully implemented HOTS-based questions. But it is more likely to measure lower-order thinking (LOT) and higher-order thinking (HOT). Such as practice questions and daily test questions for students. Cognitive assessment instruments are in the form of questions so that the questions used by students are still included in LOT and HOT questions which only remember, understand and apply, while at this time students are required to think critically and creatively. In accordance with the results of observations on the 2019 UNBK questions, 30% were based on critical thinking indicators. Based on these problems, the researcher believes that HOTS questions need to be developed in several materials so that students can often be trained to work on HOTS questions and can improve students' skills in learning Physics, especially critical thinking skills. This is in accordance with the view of Hanifah (2019) which states that the application and development of HOTS in learning is very important. If students have advanced thinking skills, they will be able to use problem solving methods correctly, appropriately and confidently in their learning. Learning activities that are focused on the objectives of developing HOTS have a major impact on improving the quality of education. However, due to the lack of test kits that can be used to train

and improve students' advanced thinking skills or HOTS in physics class, it is necessary to develop HOTS-based test kits in physics class to train critical thinking skills.

Application of ratings ability to think critically is often done by using a test essay Mukti and Istiyono, (2018). Critical thinking skills are one aspect of HOTS. According to Fanani (2018), the steps for writing HOTS questions are: a) analyzing the basic skills that can be achieved by HOTS questions, b) writing a grid of questions, c) choosing interesting and contextual stimuli, and d) writing interesting questions. According to the question grid, e) Develop assessment criteria (standard) or answer keys.

Based on the description of the background above, the formulation of the problem in this study is how the process of developing a HOTS-based test instrument to measure students' critical thinking skills on the material of quantities and units. This aims to see the feasibility of the developed test instrument.

RESEARCH METHODS

The research was conducted at SMAN 2 Rambah Hilir with the research subjects of class X students. This research is development research which is better known as *Research and Development* (R&D). This research design uses the ADDIE model developed by Dick and Carry (1996) to design a learning system. This model is arranged sequentially and programmed so that the sequence of activities is systematic. This model consists of five stages, namely *Analysis, Design, Development, Implementation, and Evaluation.*

Data collection techniques aim to collect data needed in research. Data collection is carried out during the process of preparing the assessment instrument, including through:

1. Testing the feasibility of the assessment instrument developed using expert judgments that will prove the validity of the content of the instrument.
2. Conducting trials of the assessment instrument developed to students of SMA Negeri Rambah Hilir class X MIPA and seeing the resulting score to determine the feasibility of the developed question instrument.
3. Documentation in the form of value data from the work carried out by students.

The data analysis technique in this study went through three stages, namely:

a. Analysis of validation results by experts

Validation data by experts is obtained from the validation of each item on the validation sheet. Expert validation questionnaires are arranged with an interval scale of 1 to 4 for each item. Conclusion section will obtain a decent instrument to use, fit for use with revisions, or unfit for use. Validator experts analyze the problem and then the data will be calculated using a CVR.

b. Item analysis

After that, the questions were tested on students with data obtained from students and then processed using:

1. Item validity

$$rxy = \frac{N\sum XY - \sum X \sum Y}{\sqrt{\left\{N\sum X^2 - \left(\sum X\right)^2\right\}\left\{N\sum Y^2 - \left(\sum Y\right)^2\right\}}} \quad (1)$$

Information:

rxy = Correlation coefficient

N = Number of question items

X = Value for each item

Y = Total value for each item

2. Reliability

$$r_{11} = \left(\frac{n}{n-1}\right)\left(1 - \frac{\Sigma\sigma_i^2}{\sigma_t^2}\right) \quad (2)$$

where:

r_{11} = Reliability sought

n = number of questions

$\sum \sigma_i^2$ = Total score variance of each item

σ_t^2 = total variance

3. Difficulty Level

$$P = \frac{\sum X}{S_m N} \quad (3)$$

Information:

P = difficulty level

x = number of test takers who answered correctly

S_m = maximum score

N = number of test takers

4. Different Power

$$DP = \frac{WL - WH}{n} \quad (4)$$

DP = the amount of discriminating power you want to find

WL = number of students who answered incorrectly from the lower group kelompok

WH = number of students who answered incorrectly from the upper group kelompok

n = sample size from one of the groups.

c. Analysis of student responses to questions

$$\%respon = \frac{scoreobtained}{\max imumscore} \times 100\% \quad (5)$$

The diagram or research chart can be seen as follows:

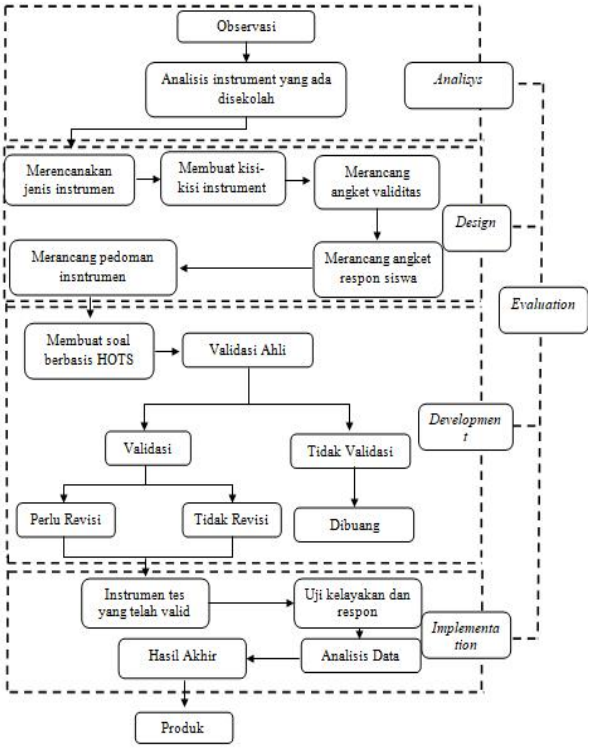

Figure 1. Research flow

Results and Discussion

The results of the research conducted in the form of a HOTS-based test instrument that can be used to measure students' critical thinking skills on the material of quantities and units. Instruments developed can be declared eligible by the validity by experts, forecast validity, reliability, level of difficulty, the power difference, and the student questionnaire responses. The development of this test instrument uses the ADDIE development model, which includes the *Analysis, Design, Development, Implementation,* and *Evaluation* stages. The description of the development stages can be seen in the following description:

1. Stage of Analysis (Analysis)

At this early stage, observations were made by analyzing the problems and needs of students at SMA N 2 Rambah Hilir. The main activity carried out at this stage is to analyze and see the test instruments used by teachers in schools. And conduct interviews with teachers related to the test instruments commonly used by teachers to evaluate student learning outcomes. Based on the results of observations made, the researchers found that the application of the HOTS-based test instrument to measure critical thinking skills had not been implemented as much as possible. This is not in accordance with the target needs of the revised k13 curriculum. An example of an instrument used in schools is as shown in the table below:

Table 1. Instruments used in schools

No	Question	Answer
1.	The attractive force between two objects whose masses are m_1 and m_2 and separated by a distance *r is*	The physical concept related to this problem is the force of gravity.

	expressed by $F = G\frac{m_1 m_2}{r^2}$ where G is a constant, then the dimensions of G are...	G = $\frac{F \times r^2}{m_1 m_2}$ G = $\frac{(kgm/s^2)(m^2)}{kg \times kg} = \frac{m^2}{s^2 kg}$ Then the dimensions of this unit are $M^{-1} L^2 H^{-2}$
2.	The results of measuring the length and width of a rectangular plot of land are 13, 35 m and 12.5 m. The area of land according to the rule of significant figures is…. m^2	Area = Length × Width Area = 15.35 × 12.5 = 192 m^2 In the multiplication operation, the number of significant figures must be equal to the least significant number, the length has 4 significant figures and the width has 3 significant figures so that it follows the width which has 3 significant figures.
3.	If a sign states a maximum speed limit of 55 miles per hour (mph), what is the speed: a. In meters per second (m/s) In km per hour (km/h)	a. Can be written 1 mile as 1 mile = (5280 ft) $\left(12\frac{in}{ft}\right)\left(2{,}54\frac{cm}{in}\right)\left(\frac{1m}{100cm}\right) = 1609m.$ we also know 1 hour consists of 3600 s, so $55\frac{mil}{jam} = \left(55\frac{mil}{jam}\right)\left(1609\frac{m}{mil}\right)\left(\frac{1jam}{3600s}\right) = 25\frac{m}{s}$ where we round the result to both significant figures. b. 1 mile = 1609 m = 1,609 km $55\frac{mil}{jam} = \left(55\frac{mil}{jam}\right)\left(1{,}609\frac{km}{mil}\right) = 88\frac{km}{jam}$

While the core competencies in the tenth-grade physics subject syllabus, students are required to:

KI 3: Understanding, applying, recognizing factual, conceptual, procedural knowledge based on their curiosity about science, technology, art, culture, and humanities with insight into humanity, nationality, state, and civilization related to the causes of phenomena and events, as well as applying procedural knowledge in a specific field of study according to their talents and interests to solve problems.

KI 4: Processing, reasoning, and presenting in the concrete and abstract realms related to the development of what they learn at school independently, and being able to use methods according to scientific rules.

After analyzing the curriculum and questions for 2019 and 2020, the researchers concluded that the HOTS-based test instrument to measure critical thinking skills really needs to be developed at SMAN 2 Rambah Hilir.

2. Product Design Phase (Design)

At the product design stage (*design*) is a continuation of the analysis stage that has been carried out. The results obtained at this stage are the design of the test instrument to be developed, scoring guidelines, validity questionnaires, and student response questionnaires. The design of the test instrument is in the form of an instrument grid. In this study, the researchers designed a HOTS-based test instrument to measure students' critical thinking skills on 10 items of Quantity and Units. The test instrument made by the researcher was in the form of essay questions with reference to critical thinking indicators according to Ennis which included (1) providing simple explanations (*elementary clarification*), (2) setting strategies and tactics (*strategy and tactics)* , (3) building basic skills (*basic support*), (4) conclude (*inference*), and (5) provide further explanation (*advance clarification*). Each indicator consists of 2 development items .

However, when the questions were shown to one of the expert validators, the questions were still not said to be HOTS questions to measure critical thinking skills. Because one of the HOTS criteria for measuring critical thinking skills is that the question must have a stimulus, then the question directs students to think critically about the answer. After reading and finding out from several sources about HOTS to measure thinking skills, the researchers remodeled all the instruments that were previously made into new instruments according to the criteria for critical thinking .

In addition to the test instrument grid, the results obtained in the product design stage in the development process carried out there are also scoring guidelines that are used as a guide in checking the results of student answers, validity questionnaires used to validate test instruments by a team of experts, and student response questionnaires. used to see students' opinions on the test instrument.

3. Product Development and Testing Phase (Development)

Furthermore, the stages of making and testing the product are carried out, while the results obtained at this stage are in the form of details of the test instruments developed in accordance with the grid that has been made. The instrument used to measure content validity in this study used an instrument developed by Yulisa (2019) . The content validation of this test instrument was validated by five experts or validators, then the results of the validation were processed by finding the value of Lawshe's *content validity ratio* (CVR). (1975) to determine the validity of each item. Furthermore, the test instrument that has been revised

according to the suggestions and comments by the expert team will be directly tested on the students.

The data obtained from the validator is then calculated using content validity to determine which questions are eligible to be tested on students and which are not. Based on the results of CVR Lawshe obtained, all questions are declared valid.

4. Product Implementation Phase (Implementation)

At the implementation stage of this product, the researcher tested the revised test instrument. The instrument was given to students for class X MIPA at SMAN 2 Rambah Hilir as many as 30 students who had studied the material for Quantities and Units to determine the feasibility of the test instrument. The test was carried out by giving questions to the teacher and then the teacher gave the questions to the students through the class WhatsApp group listed in Appendix 18. The results of the students' answers obtained were then processed to determine the validity of the predictions, reliability, level of difficulty, and distinguishing power of the test instrument. Judging from the results of the validation of the test instrument predictions by experts, there were 10 questions, so the test instruments tested in the field also amounted to 10 questions.

5. Product Evaluation Phase (Evaluation)

The evaluation stage is used at the end of each stage to make improvements or revisions to the product. So that it can reduce errors in conducting research. The evaluation is in the form of input such as suggestions, revisions, and comments at each stage. The following is a formative evaluation conducted by researchers:

a. Analysis stage

At this stage the researchers only observed interviews with teachers at SMAN 2 Rambah Hilir school. So it was found that the teacher had used the HOTS-based question instrument but the use of the instrument was still not fully implemented. So that the supervisor provides suggestions for analyzing the curriculum and finding out the demands of the curriculum used at this time.

b. Design stage

At this stage the researcher has designed a hots-based test instrument to measure critical thinking skills. But when discussing with one of the validators, the instrument designed did not meet the criteria for critical thinking. Because critical thinking questions should be given a long enough stimulus, then the questions are directed so that students think about the information provided critically. So in accordance with the suggestions that have been given, the previously designed instrument must be changed according to the requested criteria.

c. Development stage

At this development stage, the researcher made improvements to the test instrument through a validation process in accordance with the validator's direction.

d. Implementation stage

At this evaluation stage, the researcher conducts a summative evaluation, namely making decisions about the sustainability of the test instrument. In the summative evaluation , the prediction validity and reliability, level of difficulty, distinguishing power of the developed test instrument, and student responses were carried out.

1) Item Validity

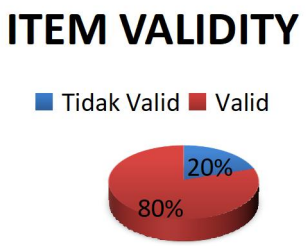

Figure 2. Item Validity Image

Item validity was carried out during the implementation of student trials. The results obtained are 8 questions have valid criteria for having $r_{\text{count validity per item about}} > r_{\text{table}}$ with a significance level of 5%, and the matter is not valid.

2) *Reliability*

The reliability used by researchers in this research uses the Alpha formula. The reliability results obtained have moderate correlation criteria with a value of 0.543390. Then $r_{\text{count the whole question}} > r_{\text{table}}$ that is 0.543390 > 0.361, then the test is declared reliable.

3) Difficulty Level

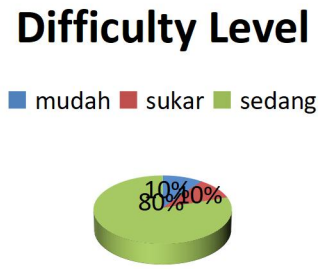

Figure 3. Difficulty Level

There are three levels of difficulty in this study, namely easy, medium, and difficult.

4) Differing Power

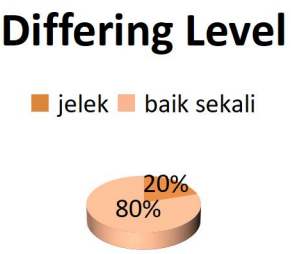

Figure 4. Different Power

There are three differentiating factors in this study, namely bad, good, very good. The results of the differentiating power in this study are shown in Fig

5) Student Response

At the time of implementation to students, not only instruments were distributed but also student response questionnaires. Student response questionnaires were distributed using a google form so that it could make it easier for students to fill out the questionnaire. The student response questionnaire used a student response questionnaire developed by Yulisa (2019) with slight changes. The results obtained were then calculated by finding the percent (%) average value obtained with a value of 78 , 1 % in both categories.

Discussion

This research was conducted at SMAN 2 Rambah Hilir in the even semester of TP 2020/2021 in class X MIA 1. This study was conducted to determine the feasibility of developing a HOTS-based test instrument to measure students' critical thinking skills on the material of Quantities and Units. This study uses a *Research and Development (R&D)* research model, with the ADDIE method, namely *analysis, design, development, implementation, evaluation*. In testing the feasibility of the HOTS-based test instrument to measure critical thinking skills, validity, reliability, level of difficulty, and differentiability were carried out, as well as student response tests. This is by the requirements of a good test according to Arikunto (2010) including validity, reliability, objectivity, practicality, and economy.

This study uses a feasibility test of content validity and item validity. For content validity, calculate the results of the data provided by the validator using CVR *Lawsh* , where the value of the CVR obtained is 1 with 10 questions declared valid. This means that the developed test instrument has met the criteria for assessing critical thinking indicators that have been validated by 5 validators to assess their validity. In line with previous research by Yul Ifda Tanjung and Yulisa Ardiani Dwiana (2019) where it was said that the results of the validity of the Lawshe CVR conducted, 13 items out of 15 items were declared valid, with a percentage of 87%

valid questions and 13% invalid questions. Then a field test was conducted, namely, the validity of the items with the acquisition of valid questions as many as 8 questions and 2 invalid questions. Because on the validity of the items that are said to be feasible or valid as many as 8 questions, then these 8 questions can be used. It is said to be valid because the question is considered to have a validity coefficient past 0.361 where 0.361 is the r table in this study. For the reliability feasibility test, the results obtained were 0.543390 with moderate correlation criteria. These results indicate that the r count > r table is 0.543390 > 0.361, which means that if this instrument is measured repeatedly on the same subject, it will give relatively the same results. The results of the research for the reliability aspect are in line with the research conducted by UU Rosidin, Agus Suyatna, and Abdurrrahman Abdurahman (2019) where in this study the multiple-choice test reliability score was 0.65 which was included in the high category and the score on the essay was 0.71 which also included in the high category with an r-table of 0.345. This is obtained by comparing the r-table with the r-count value obtained so that the result of r-count > r-table.

Judging from the aspect of the level of difficulty, there are 3 criteria, namely easy, medium and difficult. In the results of this study, question number 1 is in the easy category, and question number 3 is difficult, and for the other questions, the category is moderate. This means that the test instrument developed is good because the questions are said to be good if the questions are not too easy and not too difficult. If the question is a too easy, not stimulating way of student thinking and if too difficult, then these students become desperate and do not have the spirit of the work on the problems. The results of this study are in line with research conducted by Abdul Malik, UU Rosidin, Chandra Ertikanto (2018), which is that the difficulty level is obtained with a proportional number of relatively balanced easy and difficult questions. The shape of the problem is multiple-choice with an easy difficulty level (10%), moderate (73, 33 %), difficult (16.67%) and the essay obtained easily (10%), moderate (60%), and hard (10 %).

Judging from the aspect of distinguishing power, this study was divided into two groups, namely the upper 50% of the upper group and the lower 50% of the lower group. From the test results obtained 2 questions with an index of -1.8 to -1.9 so that it is categorized as having bad criteria, and 8 questions with an index of 1.2 - 2.5 with a very good category. Based on these results it can be said to be on a different power either category as a matter of value under different power index is 0, 00 then the question is not good. The meaning of the power difference is when students are clever and not clever to answer the questions correctly then this issue does not have the power difference, whereas if the student clever and not clever is equally unable to answer the question, then this question does not have a different power. In line with research conducted by Husnawati Hall, Hartono Hartono, Masturi Masturi (2019) where the discriminatory test of questions is calculated after obtaining questions after revision with a value of 90% very good, 10% good. With the first results about the distinguishing features are excellent, 3 has the power difference is quite good, and one question has different powerless well.

From the results of testing student responses in this study, a score of 78.1 % was obtained, meaning that students showed a positive response to the HOTS-based test instrument as a result of development. This is in line with research conducted by Lucy Asri Purwasari and Nur Fitriyana

(2020) where the student response results obtained were 3.82 which had a good category, so the HOTS-based worksheets developed were ready to be used as teaching materials.

The results of the instrument recapitulation based on item analysis can be seen in the table below:

Table 2. item analysis results analysis

Number	Validity	Reliability	Difficulty Level	Differing Level
1	invalid	0,543390	Easy	Bad
2	Valid		Currently	Very well
3	Valid		Hard	Very well
4	Valid		Currently	Very well
5	Valid		Currently	Very well
6	Valid		Currently	Very well
7	Valid		Currently	Very well
8	Valid		Currently	Very well
9	Tidak valid		Currently	Bad
10	Valid		Currently	Very well

Based on the relevant research, the product of this research is feasible to use by the feasibility test criteria. Some relevant studies that become the reference for this research are the research conducted by Yul Ifda Tanjung and Yulisa Ardiani Dwiana (2019) with Straight Motion material for *critical thinking* test instruments which obtained research results that the instrument developed can be said to be feasible if it has carried out expert validity, reliability, the level of difficulty, and differentiating power, as well as the instrument developed also received a very good response with a percentage of 75%. Meanwhile, research conducted by Nirwana, Syaiful Rochman, Alif Yanuar Zukmaidi (2018) has research results that are based on estimated parameters, their instrument product is declared valid and reliable, and the instrument has high stability and can be used to measure high-order thinking skills of school students. medium. The results of research conducted by Nurris Septa Pratama and Edi Istiyono (2015) show that the design of physics learning is based on the moderately implemented category (TS). According to the results of research conducted by Edi Istiyono (2018), teachers must give extra effort in applying HOTS-based learning assessments. And in the research of Merta Dhewa Kusuma, Invite Rosidin, Abdurrahman, Agus Suyatna (2017) where the use of the HOTS instrument is an alternative for teachers to train and determine students' HOTS levels. By working on HOTS questions, students can provide answers that match their thinking abilities.

In conducting this research, it is undeniable that there are obstacles experienced by researchers. One of them is because, during this pandemic, schools are still conducting online learning. Where the physics learning process is only carried out for one hour, and there are several sub-materials of Quantities and Units that are not taught. In addition to the constraints of the pandemic, the material of Quantity and Unit has also been studied for a long time. So that students have difficulty in answering questions. In addition, because the questions are distributed through the class's WhatsApp group and given a week to work. It is undeniable that not all students participate and contribute to answering questions according to their abilities. As a result, there are several students' answers that are the same so that it affects the differentiating power and the level of difficulty of the HOTS-based test instrument to measure critical thinking skills.

CONCLUSION

Based on the results of the study and the data obtained, the conclusions of this study by taking into account the research objectives that have been formulated are as follows:

1. The development of a HOTS-based test instrument to measure students' critical thinking skills on the Materials of Quantities and Units through the development process using the ADDIE model. Where this model consists of five stages, namely *analysis, design, development, implementation,* and *evaluation.*
2. The HOTS-based Physics test instrument to measure students' critical thinking skills on the Materials of Quantities and Units is through several feasibility tests such as content validity, item validity, reliability, level of difficulty, and discriminating power. With content validity, 100% of the questions are bound to be valid, the validity of the items has 80% of the questions said to be valid and 20% are said to be invalid. To be reliable, it has a moderate correlation. The difficulty level has 10% easy, 10% difficult, and 80% moderate. And for the difference, 80% of the questions have a very good difference and 20% of the difference is bad. And the student's response was declared good with a value of 78.1%

ACKNOWLEDGMENTS

I am very grateful to all those who have supported and helped during my study. And also thank all those who helped me in My research. I would like to thank all the supervisors and examiners for the guidance, direction, and time given to me. The biggest thanks to my supervisor, Mr. Sabani, S. Pd., M. Pd. who has guided me, provided input and support during the preparation of this assignment. I also thank Mr. Drs. Abu Bakar, M. Pd., Dr. Dewi Wulandari, M. Si., Yul Ifda Tanjung, S. Pd., M. Pd. who have helped me and provided very supportive criticism and suggestion during the research process. And a big thank you to my beloved family who has prayed for me, supported, and helped me so that I could complete my study.

REFERENCES

Arikunto. (2010). *Prosedur Penelitian: Suatu Pendekatan Praktik*. Rineka Cipta.

Hanifah, N. (2019). Pengembangan instrumen penilaian Higher Order Thinking Skill (HOTS) di sekolah dasar. *Conference Series*, *1*(1), 1–8. http://ejournal.upi.edu/index.php/crecs/article/view/14286

Husnawati, A., Hartono, H., & Masturi, M. (2019). Pengembangan Soal Higher Order Thinking Skill (HOTS) Fisika Kelas VIII SMP Materi Gerak Pada Benda. *UPEJ Unnes Physics Education Journal*, *8*(2), 133–140. https://doi.org/10.15294/upej.v8i2.33320

Istiyono, E. (2018). IT-based HOTS assessment on physics learning as the 21st century demand at senior high schools: Expectation and reality. *AIP Conference Proceedings*, *2014*(September). https://doi.org/10.1063/1.5054418

Kusuma, M. D., Rosidin, U., Abdurrahman, A., & Suyatna, A. (2017). The Development of Higher Order Thinking Skill (Hots) Instrument Assessment In Physics Study. *IOSR Journal of Research & Method in Education (IOSRJRME)*, *07*(01), 26–32. https://doi.org/10.9790/7388-0701052632

Malik, A., Rosidin, U., & Ertikanto, C. (2018). Pengembangan Instrumen Asesmen Hots Fisika Sma. *Lentera Pendidikan Pusat Penelitian LPPM UM Metro*, *3*(1), 11– 25.

Mukti, T. S., & Istiyono, E. (2018). Instrumen Penilaian Kemampuan Berpikir Kritis Peserta Didik SMA Negeri Mata Pelajaran Biologi Kelas X. *BIOEDUKASI: Jurnal Pendidikan Biologi*, *11*(2), 107–112. https://jurnal.uns.ac.id/bioedukasi/article/download/21624/pdf

Nirwana, N., Rochman, S., & Zukmadini, A. Y. (2019). *An assessment of Higher Order Thinking Skills (HOTS) Based on Rasch Models of Student in Physics Learning*. *295*(ICETeP 2018), 162–165. https://doi.org/10.2991/icetep-18.2019.40

Pratama, N. S., & Istiyono, E. (2015). Studi Pelaksanaan Pembelajaran Fisika Berbasis Higher Order Thinking (Hots) Pada Kelas X Di Sma Negeri Kota Yogyakarta. *PROSIDING : Seminar Nasional Fisika Dan Pendidikan Fisika*, *6*(2), 104–112. http://jurnal.fkip.uns.ac.id/index.php/prosfis1/article/view/7711/5687

Purwasi, L. A., & Fitriyana, N. (2020). *Pengembangan lembar Kerja Peserta Didik (LKPD) Berbasis Higher Order Thingking Skill (HOTS) Pendidikan Matematika STKIP PGRI Lubuklinggau , Indonesia*. *9*(4), 894–908.

Rosidin, U., Suyatna, A., & Abdurrahman, A. (2019). A combined HOTS-based assessment/STEM learning model to improve secondary students' thinking skills: A development and evaluation study. *Journal for the Education of Gifted Young Scientists*, *7*(3), 435–448. https://doi.org/10.17478/jegys.518464

Sani, Ridwan.Abdullah. (2019). *Cara Membuat Soal HOTS*, Tangerang : Tira Smart

Siwardani, N. W., Dantes, N., & IGK Arya, S. (2015). Pengaruh Model Pembelajaran ADDIE Terhadap Pemahaman Konsep Fisika dan Keterampilan Berpikir Kritis Siswa Kelas X SMA Negeri 2 Mengwi Tahun Pelajaran 2014/2015. *Jurnal Administrasi Pendidikan*, *6*(1), 1–10.

Tanjung, Y. I., & Dwiana, Y. A. (2019). Pengembangan Instrumen Tes Berbasis Critical Thinking Skill Pada Materi Gerak Lurus. *Jurnal Inovasi Pembelajaran Fisika (INPAFI)*, *7*(4), 80–86. http://jurnal.unimed.ac.id/2012/index.php/inpafi

Yuliantaningrum, L., & Sunarti, T. (2020). Pengembangan Instrumen Soal Hots Untuk Mengukur Keterampilan Berpikir Kritis, Berpikir Kreatif, Dan Pemecahan Masalah Materi Gerak Lurus Pada Peserta Didik Sma. *Inovasi Pendidikan FIsika*, *09*(02), 76–82.