Jurnal Pembelajaran Dan Biologi Nukleus p-ISSN: 2442-9481
Vol 10 (3): 1083-1100, November 2024 e-ISSN: 2685-7332

Development of Instrument Based on The Higher Order Thinking Skills (HOTS) to Practice of Critical Thinking Ability of High School Students on The Motions Systems Material

Eka Wahyu Putri M. Ritonga(*), Ummi Nur Afinni Dwi Jayanti

Biology Education of Faculty of Tarbiyah and Teacher Training,
Universitas Islam Negeri Sumatera Utara
Jl. William Iskandar Pasar V, Medan Estate, Deli Serdang Regency,
North of Sumatera 20221, Indonesia

***Corresponding author:** ekawahyuputriritonga@uinsu.ac.id

Submitted August 07Th 2024 and Accepted October 30Th 2024

Abstract

This research aims to develop a valid and practical Higher Order Thinking Skills (HOTS)-based instrument that is used to measure or train students' critical thinking abilities. The research uses the Bord and Gall model research and development (R&D) method. The research instruments are validation sheets (validation sheets from instrument experts and evaluation experts), questionnaires (teacher responses), and tests (effectiveness). The Second grade students (Class XI MIPA) of senior high school SMA Gajah Mada as subjects. Testing was carried out on student and biology teachers as respondents for the practicality of the module using student and biology teacher response questionnaire sheets. Based on the results of material expert validation, it was found that the expert validator's evaluation demonstrated the validity of the instrument by classifying it as appropriate for use based on an analysis of the material, construction, and linguistic aspects. This instrument also meets the reliability criteria. Multiple choice questions have a medium level of difficulty, good discriminating power

Keywords: *HOTS; Instruments; Motion Systems*

https://doi.org/10.36987/jpbn.v10i3.6226

INTRODUCTION

High-level thinking skills are an important aspect of education (Karim & Marzita, 2019). This ability is very important in the educational process because it can influence students' abilities (Ramdiah et al., 2019). HOTS does not only think about remembering and applying, but also high-level thinking skills including aspects of critical, logical, reflective, metacognitive and creative thinking abilities (Amin et al., 2018). In line with Hamdi et al., (2018) HOTS includes critical,

creative thinking and problem solving. HOTS also requires students to reach possible answers in new situations (Arafah et al., 2017). There are several skills that individuals must have in the current era. One of these skills is the ability to think critically in problem solving. The 21st century is characterized by rapidly developing science and technology. Facione (2011) believes that a person's critical thinking ability is very influential for their future life. So that when someone is able to make decisions using critical thinking, they can make wise decisions in their life. Therefore, in the 21st century, every individual is able to think critically, solve problems and use technology as well and wisely as possible.

Based on research results Putriningtyas et al., (2022) which stated that Indonesian students' critical thinking abilities in 2015 submitted level 61st out of 63 command on value of 403 points (Saputra et al., 2019). Rofi'uddin & Zuhdi (2000) further stated that the critical thinking skills of students graduating from primary to tertiary education are still relatively low because thinking education has not been handled well. Therefore, it is very important to integrate critical thinking skills into every subject, including biology learning.

Critical thinking abilities consist of two components, namely cognitive skills and critical thinking tendencies. Critical thinking will be easier to achieve if someone has the tendencies and abilities that are considered characteristics of critical thinkers (Safitri et al., 2020). The tendency to think critically can be the difference between someone who has competence in critical thinking and someone who only has the cognitive skills to think critically (Facione, 2011). The lack of analytical skills is based on several things, such as students often memorizing topics and formulas so they don't understand the topic. This case is the same as the findings of Sianturi et al., (2018) explaining that the lack of student response and the tendency to memorize rather than understand concepts as a result, students' analytical skills are not yet capable. The active role of students is still lacking, shown by the small number of students questions and giving opinions. Proven on a number of students and forward input. students have this teacher without analyzing, criticizing, evaluating what the teacher conveys.

Students in their first year tend not to have the confidence to answer questions and do not have the courage to ask when experiencing difficulties. Therefore, students' critical thinking tendencies in the first year are lower (Agustanti, 2012). Syahfitri et al., (2019) in their research also revealed that it is possible that students in the first semester of their studies still tend to be afraid to ask questions, are awkward about exploring the knowledge they have, and are not sure about the right decisions to make. This can also be related to the large number of references, readings and students' knowledge in applying them to solve a problem. So the higher a person's strata is in line with the amount of knowledge information obtained, so that with his ability to think correctly based on the knowledge he has, it will be processed appropriately and creatively to produce logical conclusions (Syahfitri et al., 2019). To support critical thinking skills, teachers must design and choose learning models that can increase students' thinking activities, make students accustomed to solving problems and can develop their critical thinking abilities (Heong et al., 2012). This means that educators carry out essential obligations during

instruction. Educators are required to be creative, innovative, and not flashy during instruction. In reality, instructional educators display conventional learning. Apart from that, teaching staff must be able to see the problems that students have. Not only creative, reliable in teaching, but student skills are really needed to create a sense of self-confidence in students, therefore educators can have an impact on learning. Educators can create active classes, not only focused on educators but on students (Wafi & Arif, 2020). However, it is still not optimal because they have not thought much about evaluation, which is an important component in learning, which is supported by data that teachers are generally not used to using learning evaluation tools that can train students to think critically.

The second grade (Class XI biology subject matter) which requires a minimum KD in analyzing is KD 3.5 and 4.5. The minimum demands for analysis are listed in the human movement system material. In this material, students are able to analyze and solve problems because they relate to the mechanisms of life. The movement system is related to the mechanisms in the bodies of living things that produce movement, one of which is to move places, making this system easy for students to observe if problems or functional disorders occur. Its contextual nature/everyday problems can shape students' thinking patterns. Movement system material that is directly related to students makes this material capable of developing knowledge, so it is suitable to be developed to familiarize students with higher level thinking (Nirwana et al., 2020).

Apart from that, there are no HOTS questions that measure indicators, there are still a few questions that have been tested as valid and reliable in the movement system material, making this research necessary to carry out. Thus, innovation is needed to create instruments based on HOTS to train high school students' critical thinking skills on movement systems material. Learning characteristics are able to require students to reason and analyze (Saptono et al., 2013).

Several analysis has been carried out that discuss the development of HOTS instruments and assessment of students' critical thinking components, including research by Afrita & Darussyamsu (2020) which discusses high-level thinking skills involving high cognitive levels from Bloom's taxonomy, namely C4-C6. In order to improve students' high-level thinking abilities, an assessment instrument in the form of multiple choices was developed. Valid criteria are obtained from descriptive statistical analysis to analyze data from the results of validation questionnaires filled in by validators (Marwiyah et al., 2018). This shows that the instrument for assessing high-level thinking abilities in this study was validly developed in terms of aspects of material, construction, language and cognitive level. The assessment instrument is at the level of high-level thinking abilities, which identifies this tool can be used to test students' high-level reasoning skills. Then further research discussing the advance of the HOTS instrument and assessment of students' critical thinking components has been carried out, including research by Marada et al., (2021) as for the topic of host evaluation.

The validity of the assessment and feasibility by expert lecturers and fellow teachers shows that the questions developed are worthy of being tested. Testing of the questions is carried out by students. The literature study stage comes from

development model theory, HOTS, critical thinking, Bloom's taxonomy, all of which are included in theoretical studies. From the initial acquisition at the location, the reason why students cannot achieve the standard score for passing the national exam is because most teachers do not implement HOTS questions during evaluation, resulting in minimal understanding of students about HOTS-based questions. Teachers often give evaluation questions at knowledge levels C1, C2, C3. This problem is the benchmark for researchers in developing Higher Order Thinking Skills (HOTS) based questions. No previous research had discussed movement systems, thus encouraging researchers to discuss the development of instruments based on Higher Order Thinking Skills (HOTS) to train high school students' critical thinking skills on movement system material.

Based on the background explanation presented, the problem formulation that can be studied in this research is how the HOTS instrument trains students' critical thinking skills in motion systems material. Therefore, the aim of this research is to develop a valid and practical (HOTS) based instrument that is used to measure or train students' critical thinking abilities. Furthermore, the expected benefit of this research for learners is that it can provide an enjoyable learning experience so that it can train learner critical thinking skills using instruments Higher Order Thinking Skills (HOTS).

METHOD

Research & Development is the methodology employed in this study. The media created is an instrument based on HOST, HOST scoring media such as test questions in the form of multiple choices and essays, in this research it was changed by adapting Borg & Gall (1983) development model into five development steps namely: 1) requires preliminary information gathering and analysis, 2) Organizing and creating the evaluation tools (Instruments of assessment), 3) early expert testing of the product, 4) evaluation and 5) implementation. Research and information gathering are conducted using pertinent theoretical research as a basis. Validation of HOTS-based instruments is carried out to evaluate the validity of HOTS-based instruments. Validation is carried out at the initial product development stage by two validations, namely instrument expert validation and evaluation expert. To test the feasibility of the HOTS assessment instrument in biology learning on movement system material, it was first validated by an instrument expert, then revised in stage one. The revised product is then validated by the teacher, then the second stage of revision is carried out.

Sample or Participant

This research was conducted at Gajah Mada High School especially on second grade (class XI MIPA) which consists of 3 classes. The product was tested on 20 student in class XI MIPA 1 and 30 student in class XI MIPA 2 and 20 XI MIPA 3 so that the total sample for testing was 50 student.

Instrument

The research instrument used is an instrument that has passed the theoretical validation stage with the research supervisor. The research instruments are validation

attachments (assessment by instrument experts and evaluation experts. questionnaire (educator responses) and analysis tests, reliability. The measurement scale for each indicator on the validation sheet is measured using a Likert scale with a score of 1-4.

Data collection

The data collection instrument consists of: 1) Test instrument, 15 questions in a multiple-choice test with five possible answers that refer to markers of advanced cognitive skills (level thinking abilities), 2) Validation sheet, conducted by two experts validator: an instrument expert for test instrument validation and an evaluation expert for HOTS assessment validation. Furthermore, The biology teacher conducted validation in order to ascertain the teacher's responsiveness and practicality as an educator in the classroom.

Procedure

Research and development sources following the steps of this research were developed in several stages, namely Potential and Problems, Literature Study and Collection, Design Validation, Revision I, Limited Trial, Revision II, and Extensive Trial. Based on several stages of research development, we can describe the steps of level 4 research and development (creating new, tested products) which have been modified, namely the first is potential and problems, this stage is to find out the potential and problems that exist in the field through observation. Both literature studies and information collection, information collection comes from the results of initial observations in the field which are supported by literature. Third, product design/design, at this stage, design the product by following the steps for preparing questions in the 2013 Curriculum guidelines. Fourth, validation, the next step for the product is to test whether the product being developed is valid or not. The purpose of validation is to determine the level of appropriateness of Higher Order Thinking Skills (HOTS) based questions.

The five design/design revisions are based on suggestions, opinions, comments from the results of expert assessments of the product design, then the product is improved so that it can be used. Sixth, namely product creation, at this stage the researcher creates a product according to the steps for Higher Order Thinking Skills (HOTS) questions, namely creating a question grid, creating questions and creating a rubric or assessment score. The seventh limited trial, limited field testing is carried out by using the product design in real conditions. The test plan is carried out once. And the last one is product revision, at this stage the results of limited testing do not meet the expected specifications, so there needs to be a revision of the product. The results of the revision are used to improve the Higher Order Thinking Skills (HOTS) question item product.

Data analysis

Data analysis was carried out from two methods of analysis, namely qualitative data on evaluation instrument values and quantitative data analysis of HOTS test questions. From the qualitative data analysis, The evaluation questionnaire's usefulness is determined by the findings from three different

perspectives: linguistic, construction, and material. Valid test questions are used from expert assessments, namely 1 expert lecturer and 1 Biology teacher. There are several examples of HOST questions that have been validated by the validator, namely, (1) Jurisprudence had an accident so that the bones were cracked or broken. However, the position of the fracture has not shifted from its initial position and does not injure the surrounding muscles. The disorder experienced by this person is categorized as... a) Simple fracture b) Partial fracture c) Complex fracture d) Multiple fracture e) Complete fracture; (2) Osteosarcoma is a type of bone cancer. This type of cancer attacks many teenagers and adults of productive age. Bone cancer treatment is generally carried out by surgery to remove the tumor completely. One of the technologies used in treating bone cancer is Limb ablation. What is the mechanism for treating osteosarcoma with limbablation? a) Connecting the bones of patients affected by tumors with bones taken from patients who have died. b) Removal of bone cancer cells with radiation. c) Amputating bones that have been affected by cancer and cannot be saved anymore. d) Dissect bone tissue and remove cancer cells. e) Curing cancer through chemotherapy; (3) Below are the compounds that play a role in the muscle contraction mechanism: O: acetylcholine, P: excitatory, Q: actin and myosin, R: actomyosin, S: energy from ATP. The correct sequence of muscle contraction mechanisms is... a)P→ Q → S → O → R; b)P → O → Q → S → R c); P → Q → O → S → R; d)R → S → O → Q → P; e) S → O → P → Q → R.

From this example question, the value given to each validation value item is: 1 as "not correct" category, 2 as "less legitimate" category, 3 as "fairly correct" category, 4 as "correct" category, and 5 "very correct" category. The Aiken's V calculation formula is used to analyze HOTS test questions.

When evaluating HOST test questions quantitatively, information gleaned from student answers is examined using Microsoft Excel. The evaluation is carried out to see the characteristics of the HOST question items including 4 characteristics, The first step is to use the point-by-point correlation formula to assess the test's validity, as long as the computation results are compared with computations at a 5% significance level. The item is correct if the calculation is greater than or equal to the calculation; if the calculation is less than the calculation, the item is incorrect. The second is a reliability test that use the KR-20 formula due to the binary of question scores (0 and 1). To ascertain whether the coefficient interval value is ≥ 0.7 in order to determine the reliability criteria. Third is the difficulty level test, It is necessary to analyze the test questions' degree of difficulty in order to determine which ones belong into the easy, medium, and difficult categories. The formula used to calculate the level of difficulty of the questions is: P=NP/N, utilizing the following criteria: 0.71-1.00 as "too easy" criteria, 0.31-0.70 as "medium" criteria, and 0.00-0.30 as "too difficult" criteria (Kuswana, 2011).

Lastly, distinguishing power, differentiating power is the test's ability to differentiate between students who have high abilities and students who have low abilities. The strength index of different items uses the formula DP=BA/JA-BB/JB. If the discrimination index is known, then the figure is interpreted according to the criteria: value of 0.00-0.20 as bad criteria, 0.21-0.40 as fair criteria, 0.41-0.70 as good criteria and 0.71-1.00 as very good criteria (Kuswana, 2011).

Table 1. Likert Scale Assessment Guidelines (Sugiyono, 2015)

Assessment criteria	Score
Strongly agree	4
Agree	3
Don't agree	2
Strongly disagree	1

Based on research by Ikhwani & Kuntjoro (2021), explains the formula 1 used for the percentage of product success as follows,

$$Presentase\ (\%) = \frac{score\ obtained}{Maximum\ score} \times 100\ \%$$

Analysis of the question instrument's usefulness can be demonstrated in the teacher's feedback about how the learning phases are implemented (see in Table 2). Practicality evaluation is carried out by processing data obtained from student response questionnaires. HOST media is stated to have a sufficient level of practicality, namely above 61% (Sugiyono, 2015).

Table 2. Practicality Criteria (Sugiyono, 2015)

Score ranking	Criteria
81%-100%	Very practical
61%-80%	Practical
41%-60%	Quite practical
21%-40%	Very impractical

RESULT AND DISCUSSION

Development result

From the results of research that has been carried out regarding the development of instruments based on Higher Order Thinking Skills (HOTS) to train high school students' critical thinking skills on valid and reliable movement system material. This is also in accordance with Saputra (2020) that The development of critical thinking abilities is crucial for students' academic and social success, and this can be achieved through a teaching methodology that fosters critical thinking.

The HOTS-based instrument developed has gone went through two scoring processes. First scoring is used to see validity the HOTS-based instrument which was carried out by a biology education expert. The second stage of assessment was a field trial involving 50 students from 2 classes, namely XI MIPA 2 and XI MIPA 3, short scoring on special of the HOTS test items. The process carried out in this development includes preparing HOTS test question products. The designed HOTS test questions are assessed by expert validators, revised and then an initial HOTS test question product is obtained which is ready to be used as limited trial material. According to Saraswati & Agustika (2020); Yuliana et al., (2024) The HOST test has 4 indicators such as the ability to solve problems, the ability to make decisions, the

ability to think analytically, and the ability to think creatively In line with Salirawati et al., (2017) states that the Higher Order Thinking Skills (HOTS) questions that are developed must have the characteristics of Higher Order Thinking Skills (HOTS). The quality of the questions is improved one by one so that the questions are high level thinking and the obscens must also be at the Higher Order Thinking Skills (HOTS) level.

Limited test results are considered as evaluation material in order to create good HOTS-based test media for use in field tests. After obtaining estimates of the reliability coefficient, difficulty level criteria, distinguishing power, and distracting alternatives on the field test scores, the media finally got into the HOST question questions is obtained which is ready to be used. Aisah & Pahlevi (2020); Desilva et al., (2020); Rizal & Wulandari (2020), Saputro & Supahar (2018) in accordance with the research they have conducted who also developed HOTS-based assessment instruments in the form of multiple choice questions and using the best questions, as well as research conducted by Jannah & Pahlevi (2020); Sari et al., (2018) who developed a HOTS-based assessment instrument using IT applications.

Product Trial Results

Evaluation from experts is used to see the first media. This evaluation aims to obtain input and improve the media, as well as initial media assessment before it is carried out in a limited trial. According to Sugiyono (2009), this stage aims to determine the validity of the product being developed from both media and material aspects.

Table 1. Results of Expert Validation Analysis of Multiple Choice HOTS Test Questions

Question Item Number	Aiken's V coefficient	Criteria
1	1.00	Suitable To Use
2	0.83	Suitable To Use
3	0.83	Suitable To Use
4	1.00	Suitable To Use
5	0.83	Suitable To Use
6	1.00	Suitable To Use
7	1.00	Suitable To Use
8	0.83	Suitable To Use
9	1.00	Suitable To Use
10	0.83	Suitable To Use
11	0.83	Suitable To Use
12	0.83	Suitable To Use
13	1.00	Suitable To Use
14	0.83	Suitable To Use
15	0.83	Suitable To Use

Validation activities are carried out by providing an initial product script, namely in the form of a question grid and HOTS test questions Sugiyono (2009) view, this phase aims to ensure the validity of the media being promoted both in terms of media and material. Evaluation activities are carried out by displaying the

initial product script, namely in the form of a grid of questions and HOST test questions. As well as a validation sheet to the expert validator (Syahfitri, 2019). Next, an analysis of the assessment of the HOST test questions are carried out with evaluation assessment criteria using the Aiken's V formula to calculate content. validity coefficient (Purwadhi, 2019). Analysis of expert validation data acquisition can be seen in Table 1.

Based on the results of the analysis using the Aiken's V formula, the HOTS test questions consisting of 15 multiple choice questions were all declared suitable for use. However, there were several questions that were improved according to input and suggestions from the validator, namely regarding improvements to the movement system material, including sentence formulation and indicators that did not match the question items. From the results of the investigation using the Aiken's V formula, in the HOST 5 essay questions all of the questions were declared correct.

Table 2. Results of Expert Validation Analysis of HOTS Essay Test Questions

Question Item Number	Aiken's V coefficient	Kriteria
1	0.83	Suitable To Use
2	0.83	Suitable To Use
3	0.83	Suitable To Use
4	0.83	Suitable To Use
5	1.00	Suitable To Use

Limited Product Trial Results

Item validity

The results of the trial the resource person found that the time needed to work on the HOST test questions, on multiple choice questions and essays, was approximately 120 hours needed to work on the questions. Apart from that, through interpreting the question evaluation, the level of questions can be determined from the characteristics of the questions consisting of validity, reliability, level of difficulty and distinguishing power. Based on table 3, the results show that the test is limited, the results show that the 12 multiple choice HOTS test questions are included in the "Valid" category. Based on table 4, the results show that the limited test results show that the multiple choice HOTS test questions are included in the "Valid" category.

Table 3. Summary results of the validity of multiple choice HOTS test questions

Question number	Criteria
1, 2, 4, 6, 7, 9, 10, 11, 12, 13, 15	Valid

Table 4. Summary results of the validity of the HOTS Essay test questions

Question number	Criteria
1, 2, 4, 5	Valid

Item reliability

Testing the reliability of the HOST questions using the KR-20 formula with the Microsoft Excel program, the results proved that the reliability of the questions was 0.850 and the questions were in the very high reliability category. Obtain the reliability of the questions in table 5 and table 6. From table 5, it is obtained that the

limited test shows that the multiple choice HOST test questions meet the "Reliable" criteria. Based on table 6, the results show that the limited test results show that the multiple choice HOTS test questions are included in the "Reliable" category.

Table 5. Summary results of the reliability of multiple choice HOTS test questions

Reliability Testing Criteria		
Reference Value	**Cronbach's Alpha value**	**Conclusion**
0.7	0.850	Reliable

Table 6. Summary results of the reliability of the HOTS Essay test questions

Reliability Testing Criteria		
Reference Value	**Cronbach's Alpha value**	**Conclusion**
0.7	0.780	Reliable

Difficulty

The level of difficulty of the HOST questions is calculated using the formula P=NP/N with the Microsoft Excel program. Prolehan classified 3 questions as difficult, 12 questions as medium.in the multiple choice questions while 2 in the essay questions are classified as difficult, 3 are classified as medium. The results of the difficulty level of the questions are as in Table 7 and Table 8. Referring to Sudjana (2017) opinion, there are several basic considerations in determining the proportion of the number of questions in the easy, medium and difficult categories. The first consideration is that there is balance, namely the same number of questions for the three categories.

3 including medium. The results of the level of difficulty of the questions can be seen in tables 7 and 8. In line with the opinion of Sudjana (2017), several studies have been conducted by determining the proportion of the number of questions classified as light, medium and difficult. The initial assessment is that there is balance, that is, the number of questions is the same for all three criteria. Based on Table 7, it can be seen that the level of difficulty ranges from the medium category for 12 questions (76.67 %). Based on Table 8, it can be seen that the level of difficulty ranges from the medium category with 3 questions (86.60 %).

Table 7. Initial Product Difficulty Level Results for Multiple Choice HOTS Test Questions

Category	Question Item Number	Amount	%
TK < 0.25 (difficult)	4, 11, 15	3	23.33
0,25 ≤ TK ≤ 0.80 (Medium)	1, 2, 3,5, 6, 7, 8, 9, 10, 12, 13, 14	12	76.67
TK > 0.80 (Easy)	-	0	0

Table 8. Initial Product Difficulty Level Results for HOTS Essay Test Questions

Category	Question Item Number	Amount	%
TK < 0.25 (difficult)	2,4	2	13.23
0,25 ≤ TK ≤ 0.80 (Medium)	1, 3,5	3	86.65
TK > 0.80 (Easy)	-	0	0

Differentiating Power

The differentiating power test regarding HOTS was calculated using the Microsoft Excel program. Shown in table 9 and table 10. Based on Table 9, it can be seen that the differentiating power ranges from the accepted without revision category to 13 questions (83.33%). Items in the good category and accepted without revision are immediately used in the main product. Category questions were accepted with revisions of 2 questions (16.67%), repair it first before using it on the main media.The items in the good category are accepted without revision and those that have been revised made like the first product for testing.

Table 9. Initial Product Differentiation Power Results on Multiple Choice HOTS Test Questions

Category	Question Item Number	Amount	%
DP ≥ 0.40 (Good)	-	0	0
0.30 ≤ DP ≤ 0.39 (Accept without revision)	2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15	13	83.3
0.20 ≤ DP ≤ 0.29 (Accept with revision)	1, 5	2	16.67
DP ≤ 0.19 (Replaced/ /completely revised)	-	0	0

Table 10. Initial Product Differentiation Power Results on HOTS Essay Test Questions

Category	Question Item Number	Amount	%
DP ≥ 0.40 (Good)	-	0	0
0.30 ≤ DP ≤ 0.39 (Accept without revision)	2, 3, 4,	13	79.33
0.20 ≤ DP ≤ 0.29 (Accept with revision)	1, 5	2	20.67
DP ≤ 0.19 (Replaced/ /completely revised)	-	0	0

Based on Table 10, it can be seen that the differentiating power of the essay questions ranges from the accepted without revision category to 3 questions (79.33 %). Items in the good category and accepted without revision are immediately used in the main product. Category questions were accepted with revisions of 2 questions (20.67 %), revised first before being used in the main product. The items in the good category are accepted without revision and those that have been revised are reassembled into the main product of the test questions.

Field Product Test Results

Field trials were carried out to determine the quality of HOTS test questions based on the characteristics of multiple choice HOTS test items developed as a result of limited trials. Characteristics of Multiple Choice HOTS Test Items from Field Trial Results.

Table 11. Summary of the validity of multiple choice HOTS test questions

Question number	Criteria
1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15	Valid

Table 12. Summary of the validity of the HOTS Essay test questions

Question number	Criteria
1, 2, 3, 4, 5	Valid

Based on tables 11 and 12, the results show that the limited test shows that 15 multiple choice HOTS test questions and 5 HOTS essay test questions are included in the "Valid" category.

Table 13. Summary of the reliability of multiple choice HOTS test questions

Reliability Testing Criteria		
Reference Value	**Cronbach's Alpha value**	**Conclusion**
0.7	0.905	Reliable

Tabel 14. Summary of the reliability of the HOTS Essay test questions

Reliability Testing Criteria		
Reference Value	**Cronbach's Alpha value**	**Conclusion**
0.7	0.895	Reliable

Based on table 13 and table 14, the results show that the test is limited. The results show that the HOTS test questions are multiple choice and the Essay test questions are included in the "Reliable" category.

Table 15. Initial Product Difficulty Level Results for Multiple Choice HOTS Test Questions

Category	Question Item Number	Amount	%
TK < 0.25 (Difficult)	-	0	0
0.25 ≤ TK ≤ 0.80 (Medium)	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 14, 15	15	95
TK > 0.80 (Easy)	-	0	0

Table 16. Initial Product Difficulty Level Results for HOTS Essay Test Questions

Category	Question Item Number	Amount	%
TK < 0.25 (Difficult)	-	0	0
0.25 ≤ TK ≤ 0.80 (Medium)	1, 2, 3, 4, 5	5	98
TK > 0.80 (Easy)	-	0	0

Based on Table 15, it can be seen that the level of difficulty ranges from the medium category for 15 questions (95 %). And table 16 shows that the difficulty level of the HOTS Essay test questions is in the medium category with 5 items (98 %). The differentiating power of test items is known by looking at the point biser correlation coefficient (RPBIS). In general, the differentiating power of multiple choice questions can be seen in Tables 17 and 18.

Table 17. Initial Product Differentiation Power Results on Multiple Choice HOTS Test Questions

Category	Question Item Number	Amount	%
DP ≥ 0.40 (Good)	8, 15	2	16.67
0.30 ≤ DP ≤ 0.39 (Accept without revision)	2, 3, 4, 5, 6, 7, 9, 10, 14,	9	63.33
0.20 ≤ DP ≤ 0.29 (Accept with revision)	1, 12	2	6.67
DP ≤ 0.19 (Replaced/ /completely revised)	11, 13,	2	13.3 3

Based on Table 8, 11 questions were categorized as good and accepted without revision (80 %). Items in the good category and accepted without revision are used directly as the final product. Category questions were accepted with revision and replacement, as many as 2 questions (20 %) were not used. Items in the good category and accepted without revision are reassembled into a final product of multiple choice HOTS test questions that are ready to be used.

Table 18. Initial Product Differentiation Power Results on HOTS Essay Test Questions

Category	Question Item Number	Amount	%
DP ≥ 0.40 (Good)	1, 5	2	16.67
0.30 ≤ DP ≤ 0.39 (Accept without revision)	2, 3, 4	4	83.33
0.20 ≤ DP ≤ 0.29 (Accept with revision)	-		
DP ≤ 0.19 (Replaced/ /completely revised)	-		

Based on the table 18 found that 2 questions were in the good category and 3 questions were accepted without revision (83 %). Items in the good category and accepted without revision are used directly as the final product of the HOTS Essay test questions which are ready to be used.

Product Revision

Product revisions are carried out the final media results have met the valid and reliable criteria. Evaluation carried out based on the results of the HOTS-based instrument assessment for each level of media testing. Media revisions in this research consisted of: media revisions resulting from Establishment, product revisions resulting from limited trials, and product revisions resulting from field trials.

From Purchasing improvements after expert scoring, limited and field testing the HOTS-based instrument developed underwent several evaluation. Beginning, revise the question items from on validator input and suggestions. In general, the input and suggestions relate to improvements in sentence formulation and indicators that do not match the question items. Second, revise the question items from results of limited trials of the initial HOTS test question product. Category questions are accepted with revision, revised based on the results of the analysis of the

characteristics of the questions. Third, revise the question items based on the results of field trials of the main HOTS test questions. Question items that are good and accepted without revision are re-verified using the HOTS indicator to see that all indicators are represented. The verification results of the question items are reassembled into the final product of the HOTS test questions which are ready to be used (Afdila & Sartika, 2020).

Final Product Review

The final media of this development research is a HOTS-based instrument to train high school students' critical thinking skills in class XI movement systems material in the form of a set of HOTS test questions. Tests are questions that must be completed, the form of the questions can be in the form of description questions, objective questions, multiple choice questions, matching questions, and short answer test questions (Fatimah & Pahlevi, 2020).

From results of expert validation, limited trials, field trials, and improvements, as well as data analysis carried out, it can be seen that the HOTS test questions developed have met the criteria for validity and reliability, and the quality of the questions is good. The validity of HOTS-based instruments in the form of HOTS test questions is based on predetermined product validation criteria developed. Validation was carried out by an expert from an academic lecturer at Yogyakarta State University, namely Mr. Ahmad Kamal Sudrajat, on the HOTS test question product which met logical validity and was developed to meet the criteria for being suitable for use. The reliability of the assessment instrument in the form of HOTS test questions is based on the results of the analysis of the main product items on the HOTS test questions.

The reliability coefficient obtained from the results of the analysis of multiple choice HOTS test questions is 0.713, while the reliability coefficient for the description HOTS test questions is 0.939. This is in accordance with Nurhasanah (2019), an instrument can be said to be reliable if the alpha value is > 0.600 or in the high category. The quality of the instrument in the form of HOTS test questions is based on the results of analyzing the main product items of the HOTS test questions, namely analyzing all question items based on empirical data (Ariyana et al., 2018). The multiple choice HOTS test questions have an average level of difficulty of 0.406 (medium). According to the opinion of Arikunto (2015), questions in the medium category can be used. Questions in the very easy category were not used, so the difficulty level of the HOTS-based assessment instrument in this study was in the "good" category and the average discriminating power was 0.330 (good). According to Basuki & Hariyanto (2016), questions can be accepted if they have the power to distinguish questions from the "very good" category, then the HOTS-based assessment instrument is a good question.

CONCLUSION

Based on the results of the research and discussion, it was found that the final product in this research produced a HOTS-based instrument on motion system material. The HOTS-based instrument is a set of HOTS test questions consisting of 15 multiple choice questions and five essay questions. The validity of the instrument is proven by the results of expert assessments which show that the instrument is suitable for use with an average score of 0.79 Aiken's V on multiple choice questions while for essay questions it is 0.83 Aiken's V based on a review of material, construction and language aspects. The instrument also meets the reliability criteria with a score for multiple choice questions of 0.850 Cronbach's Alpha and essay questions of 0.780 Cronbach's Alpha for the limited test while for the broad test the reliability criteria for multiple choice questions are 0.905 Cronbach's Alpha and essay questions are 0.895 Cronbach's Alpha for the limited test while the instrument is based on HOTS has a medium level of difficulty, good discrimination. For further research, you can create a graph between the differences between limited tests and extensive tests on HOTS-based test question instruments.

REFERENCES

Afdila, N, & Sartika, S.B. (2020). A secondary School Student's Critical Thinking Ability Profile in Natural Science Subject Matters Based on Learning Style. *Insecta: Integrative Science Education and Teaching Activity Journal,* 1(2), 140-170. https://doi.org/10.21154/insecta.v1i2.2230

Afrita, M., & Darussyamsu, R. (2020). Validitas Instrumen Tes Berpikir Tingkat Tinggi (HOTS) pada Materi Sistem Respirasi di Kelas XI SMA. *Mangifera Edu*, 4(2), 129-142. https://doi.org/10.31943/mangiferaedu.v4i2.83.

Agustanti, T.H. (2012). Implementasi Metode Inquiry Untuk Meingkatkan Hasil Belajar Biologi. *JPI: Jurnal Pendidikan IPA Indonesia,* 1(1), 16-20. https://doi.org/10.15294/jpii.v1i1.2007

Aisah, S., & Pahlevi, T. 2020. Pengembangan Instrumen Penilaian Higher Order Thinking Skills (HOTS) Pada Mata Pelajaran Korespondensi Kelas X OTP Di Smk Negeri 1 Jombang. *Jurnal Pendidikan Administrasi Perkantoran*, 8(1), 146-156. https://doi.org/10.26740/jpap.v8n1.p146-156

Amin, D. I.., Sutrisno, S.,& Sigit, D. (2018). Instrument asesmen pemahanan KonseptualBerorientasi HOTS-Keterampilan Proses dan sikap Terhadap Sains padaBahan Kajian Hidrokarbon dan Minyak Bumi. *Jurnal Pendidikan: Teori, Penelitian dan Pengembangan*, *3*(9), 105-115. http://dx.doi.org/10.17977/jptpp.v3i9.13342

Arafah, K., Amin, B.D., Sari, S.S & Hakim, A. (2021). The Development of Higher Order Thinking Skill (HOTS) Instrument Assessment In Physics Study. *Phys.: Conf. Series 1899 012140.* Volume 1899, 2nd Workshop on Engineering, Education, Applied Sciences and Technology (WEAST) 2020 5 October

2020, Makassar, Indonesia. 8 page. https://doi.org/10.1088/1742-6596/1899/1/012140

Arikunto, S. (2015). *Dasar-Dasar Evaluasi Pendidikan*. Fourth edition. Jakarta: PT. Bumi Aksara.

Ariyana, Y., Pudjiastuti, A., Bestary, R., & Zamroni. (2018). *Buku Pegangan Pembelajaran Berorientasi pada Keterampilan Berpikir Tingkat Tinggi.* Direktorat Jenderal Guru dan Tenaga Kependidikan Kementerian Pendidikan dan Kebudayaan. 95 page.

Basuki, I., and Hariyanto, H. (2016). *Asesmen Pembelajaran*. Bandung: PT. Remaja Rosdakarya.

Borg and Gall, (1983). *Educational Research, An Introduction.* First edition. New York: Longman Inc. 418 page.

Desilva, D., Sakti, I., & Medriati, R. (2020). Pengembangan Instrumen Penilaian Hasil Belajar Fisika Berorientasi HOTS (Higher Order Thinking Skills) Pada Materi Elastisitas Dan Hukum Hooke. *Jurnal Kumparan Fisika*, *3*(1), 41–50. https://doi.org/10.33369/jkf.3.1.41-50

Facione, P.A. (2011). *Critical Thinking: What It Is and Why It Counts*. California: Measured Reason and The California Academic Press. 28 page.

Fatimah, S., & Pahlevi, T. (2020). Pengembangan Instrumen Penilaian Berbasis HOTS (Higher Order Thinking Skills) Pada Kompetensi Dasar Menerapkan Sistem Penyimpanan Arsip Sistem Abjad , Kronologis, Geografis, Nomor, Dan Subjek Di Jurusan OTKP SMKN 1 Bojonegoro. *Jurnal Pendidikan Administrasi Perkantoran (JPAP),* 8(2), 318-328. https://doi.org/10.26740/jpap.v8n2.p318-328

Hamdi, S., Suganda, I.A., and Hayati, N. (2018). Developing higher-order thinking skill (HOTS) test instrument using Lombok local cultures as contexts for junior secondary school mathematics. *Research and Evaluation in Education Journal, 4*(2), 126-135. https://doi.org/10.21831/reid.v4i2.22089

Heong, Y.M., Yunos, J.M., Othman, W., Hassan, R., Kiong, T.T., Mohamad, M.M. (2012). The Needs Analysis Of Learning Higher Order Thinking Skills For Generating Ideas. *Procedia Social and Behavioral Sciences,* 59. 197-203. https://doi.org/10.1016/j.sbspro.2012.09.265

Ikhwani, P.N., & Kuntjoro, S. (2021). Pengembangan Lembar Kegiatan Peserta Didik Elektronik (E-LKPD) Berbasis Guided Inquiry pada Materi Perubahan Lingkungan untuk Melatihkan Keterampilan Berpikir Kritis Siswa Kelas X SMA. *BioEdu: Berkala Ilmiah Pendidikan Biologi, 10*(1), 21-30. https://doi.org/10.26740/bioedu.v10n3.p597-604

Jannah, K., & Pahlevi, T. (2020). Pengembangan Instrumen Penilaian Berbasis Higher Order Thinking Skills Berbantuan Aplikasi 'Kahoot !' Pada Kompetensi Dasar Menerapkan Penanganan Surat Masuk Dan Surat Keluar

Jurusan OTKP Di SMK Negeri 2 Buduran. *Jurnal Pendidikan Administrasi Perkantoran,* 8(1), 108-121. https://doi.org/10.26740/jpap.v8n1.p108-121

Karim, F.A., & Marzita, P. (2019). The development of higher order thinking skills (HOTS) assesment instrument for word problems. *International Journal of Academic Research in Business and Social Sciences,*9(2), 243-271

Marada, R., Nusantari, E., & Dama, L. (2021). Pengembangan Instrumen Berbasis Higher Order Thinking Skiil (Hots) Untuk Melatih Kemampuan Berpikir Kritis Siswa Pada Mata Pelajaran Biologi. *Jurnal Normalita,* 9(2), 188-194.

Nirwana, N., Akhriana, A., Aisa, S., & Saputro, A.T. (2022). Perancangan Aplikasi Flashcard Digital Sebagai Media Latihan Identifikasi Objek Pada Anak Berkebutuhan Khusus Berbasis Android. *Zonasi: Jurnal Sistem Informasi*, *4*(1), 42 - 51. https://doi.org/10.31849/zn.v4i1.9325

Purwadhi, P., (2019). Pengembangan Kurikulum dalam Pembelajaran Abad XXI. *Mimbar Pendidikan, 4*(2), 103-112. https://doi.org/10.17509/mimbardik.v4i2.22201

Putriningtyas, A., Muhlis, M., & Bachtiar, I. (2022). Perkembangan Kecenderungan Berpikir Kritis Siswa Pada Materi Biologi di MAN 2 Mataram. *Jurnal Ilmiah Profesi Pendidikan*, *7*(3), 1534–1542. https://doi.org/10.29303/jipp.v7i3b.746

Ramdiah, S., Royani, M., Malang, U. M., & Kleij, D. (2019). Understanding, Planning, and Implementation of HOTS by Senior High School Biology Teachers in BanjarmasinIndonesia. *International Journal of Instruction, 12*(1), 98-130

Rofi'uddin, A., & Zuhdi, D. (2002). *Pendidikan Bahasa dan Sastra Indonesia di Kelas Tinggi.* Malang: Universitas Negeri Malang.

Rizal, A. N., and Wulandari, S.S. (2020). Pengembangan Instrumen Penilaian Higher Order Thinking Skills (HOTS) Mata Pelajaran Otomatisasi Tata Kelola Humas Dan Keprotokolan Di SMK Negeri Mojoagung. *Jurnal Pendidikan Administrasi Perkantoran (JPAP), 8*(1), 190-234

Safitri, A., Wulandari, D., & Herlambang, Y. T. (2022). Proyek Penguatan Profil Pelajar Pancasila: Sebuah Orientasi Baru Pendidikan Dalam Meningkatkan Karakter Siswa Indonesia. *Jurnal Basicedu, 6*(4), 7076–7086. Https://Doi.Org/10.31004/Basicedu.V6i4.3274

Salirawati, D., Permanasari, R.L., Purtadi, S., Nugraheni, A.R.E., & Dina, D. (2017). Pelatihan Pengembangan Soal HOTS Sebagai Peningkatan Kompetensi Pedagogik Guru. *INOTEKS: Jurnal Inovasi Ilmu Pengetahuan, Teknologi dan Seni*, *21*(1), 14-25

Saptono, S,. Rustaman, N.Y., Saefuddin, S., & Widodo, A. (2013). Model Integrasi Atribut Asesmen Formatif (Iaaf) Dalam Pembelajaran Biologi Sel Untuk Mengembangkan Kemampuan Penalaran Dan Berpikir Analitik Mahasiswa Calon Guru. *Jurnal Pendidikan IPA Indonesia: Indonesian Journal of Science Education, 2*(1), 890-954

Saputra, H. (2020). Kemampuan Berpikir Kritis Matematis. Perpustakaan IAI Agus Salim. https://doi.org/10.17605/OSF.IO/TJ76P

Saputra, K. A. K., Pradnyanitasari, P. D., Priliandani, N. M. I., & Putra, I. G. B. N. P. (2019). Praktek Akuntabilitas Dan Kompetensi Sumber Daya Manusia Untuk Pencegahan Fraud Dalam Pengelolaan Dana Desa. *Krisna: Kumpulan Riset Akuntansi*, *10*(2), 168-176. https://doi.org/10.22225/kr.10.2.2019.168-176

Saputro, B., & Supahar, S. (2018). Pengembangan Instrumen Penilaian Kemampuan Berpikir tingkat tinggi untuk mengukur pencapaian hasil belajar fisika peserta didik SMA kelas XI materi optika. *E-Journal Pendidikan Fisika, 7*(1), 130-139

Saraswati, P. M. S., & Agustika, G. N. S. (2020). Kemampuan Berpikir Tingkat Tinggi dalam Menyelesaikan Soal HOTS Mata Pelajaran Matematika. *Jurnal Ilmiah Sekolah Dasar*, *4*(2), 257–269. https://doi.org/10.23887/jisd.v4i2.25336

Sari, D. R. U., Wahyuni, S., & Bachtiar, R.W. (2018). Pengembangan Instrumen Tes Multiple Choice High Order Thinking Padapembelajaran Fisika Berbasis E-Learning Di SMA. *Jurnal Pembelajaran Fisika, 7*(1), 356-378

Sianturi, A., Sipayung, T.N., & Simorangkir, F.M.A. (2018). Pengaruh Model Problem Based Learning (PBL) Terhadap Kemampuan Berpikir Kritis Matematis Siswa SMPN 5 Sumbul. *Jurnal Pendidikan Matematika*, *6*(1), 29-41

Sudjana, N. (2017). *Penilaian Hasil Proses Belajar Mengajar*. Bandung; PT Remaja Rosdakarya. 56-67

Sugiyono, S. (2015). *Metode Penelitian dan Pengembangan*. Bandung: Alfabeta. Page 45-58

Sugiyono, S. (2009). *Metode Penelitian Kuantitatif Kualitatif dan R & D*. Bandung: Alfabeta. Page 30-42

Syahfitri, J., Firman, H., Redjeki, S., & Sriyati, S. (2019). Pengujian Validitas Dan Praktikalitas Tes Disposisi Berpikir Kritis Dalam Biologi (Tdbkb). *Jurnal Pendidikan Sains (JPS), 7*(1), 30-35. https://doi.org/10.26714/jps.7.1.2019.30-35

Wafi, U. H., & Arif, S. (2020). The Effect Of The Application Of The Guided Inquiry Model With The Problem Solving Approach On Students 'Observation Skills In Additive And Addictive Materials. *INSECTA: Integrative Science Education and Teaching Activity Journal*, *1*(2), 143–154. https://doi.org/10.21154/insecta.v1i2.2393

Yuliana, S., Sunanti, T., & Kintoko, K. (2024). Analisis Higher Order Thinking Skills (HOTS) Siswa Dalam Menyelesaikan Soal Cerita Matematika. *Riemann: Research of Mathematics and Mathematics Education, 6*(1), 118-126. https://doi.org/10.38114/6z3pm559

How To Cite This Article, with *APA style* :
Ritonga, E.W.P.M., & Jayanti, U.N.A.D. (2024). Development of Instrument Based on The Higher Order Thinking Skills (HOTS) to Practice of Critical Thinking Ability of High School Students on The Motions Systems Material. *Jurnal Pembelajaran dan Biologi Nukleus*, *10*(3), 1083-1100. https://doi.org/10.36987/jpbn.v10i3.6226

Conflict of interest : The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author contributions : All authors contributed to the study's conception and design. Material preparation, data collection and analysis were performed by all authors. The first draft of the manuscript was submited by [**Eka Wahyu Putri M. Ritonga**]. All authors contributed on previous version and revisions process of the manuscript. All authors read and approved the final manuscript.