Development of HOTS Assessment Instruments in Vibration and Wave Courses

Nurhayati Nurhayati, Wahyudi Wahyudi*, Syarif Lukman Hakim

IKIP PGRI Pontianak, Indonesia
*e-mail: wahyudi.kakap@gmail.com

Abstract: This study aims to 1) produce a HOTS assessment instrument; 2) knowing the quality of the test instrument in terms of the feasibility of construction, material feasibility, and language feasibility according to the expert; and 3) knowing the quality of the test items in terms of validity, reliability, difficulty level and distinguishing power based on the test results. Research and Development ware used as a research method with 4D procedural development model consists of four stages, namely: the define stage, the design stage, the development stage, and the dissemination stage. The questionnaire was used for expert judgment validation. The characteristics measurement of the HOTS items instrument including the validity, reliability, difficulty level and distinguishing power of the questions. The HOTS assessment instrument developed was in the form of multiple-choice options with a reason based on HOTS in aspects of analyzing, evaluating and creating. The results of expert validation show that the average item with criteria is very good in terms of content, construct and language aspects. Instruments that have been validated and revised were tested on students who had studied vibration and wave material and the test results showed that 77% of the questions developed were of good quality with valid criteria, good distinguishing criteria, level of difficulty at moderate and easy levels and very strong reliability so that feasible and ready to be used to measure students' higher order thinking skills in vibrations and waves material.

INTRODUCTION

The industrial revolution 4.0 sparked a new vision of the industrial revolution that was born in Germany in 2011 and has been inspiring a lively (Pfeiffer, 2017). According to experts (Lasi et al., 2014; Neugebauer et al., 2016; Rüßmann et al., 2015; Schmidt et al., 2015) the industrial revolution 4.0 is predicted to have great potential benefits including improved production flexibility speed, improved service to customers and increased revenue. If the potential benefits are realized, it will certainly have a positive impact on the economy of a country (Prasetyo & Sutopo, 2018). But to realize these benefits, in general there are five challenges that will be faced (Zhou et al., 2015) namely aspects of knowledge, technology, economy, social and political. In the next 10 years, 75 percent of all types of jobs will be lost (Harahap & Rafika, 2020). This is because the work played by humans gradually replaced with digitalization technology programs. As a result, the production process becomes faster to work on and more easily distributed massively with minimal human involvement.

To face these challenges, there needs to be innovative human resources, creative, and able to create new things, namely being able to create solutions although not necessarily original. One of those who are able to print the generation (human resources) is universities. Universities are expected to prepare human resources that are able to compete in the era of revolution 4.0 which is a generation that is able to solve every problem faced with its ability to think critically, logically and analytically.

In order to achieve these goals, of course learning must lead to the improvement of students' ability to think high level or known as High Order Thinking Skills (HOTS), because high-level thinking is one of the stages of thinking that cannot be released from daily life and every student is directed to have a high level of thinking patterns because of the ability high-level thinking allows one to think critically. In addition to the use of learning approaches that encourage students to be able to think critically, logically and analytically, there is also a need for a test instrument that can measure students' high-level thinking abilities. The assessment instrument consists of HOTS questions that measure students' abilities including (1) the ability to transfer one concept to another, (2) the ability to process and apply information, (3) the ability to find links from different information, (4) the ability to use information to solve problems, and (5) the ability to study ideas and information critically.

Based on the observations of several midterm exams and final semester exams for several courses and interview results conducted with several lecturers studying physics education at one of the campuses in West Kalimantan, information was obtained test instruments used to measure students' ability to understand the lecture materials have not yet can be classified into HOTS problem instruments because the instrument has not been able to train students to develop their skills in analyzing, evaluating and creating. The problem instruments used are mostly still measuring students' abilities in the stages of remembering, understanding and applying. This is contrary to what (Kartowagiran, 2012) stated, that the assessment used by educators not only asks students to repeat the concepts they have learned in class but is expected to improve the reasoning of learners so that they can hone the skills of learners in solving problems, thinking critical and creative.

To overcome these limitations, a high level of thinking ability-based assessment instrument (HOTS) is required. The development of this assessment instrument aims to measure and observe the level of understanding that students have in analyzing, evaluating and creating. Vibrations and waves were chosen as the object of research material because students still often misunderstand and interpret various concepts of vibration and waves (Ekawati et al., 2017). Problems related to the concepts of vibration and waves are often found in everyday life, because the world is full of waves which are abstract concepts (Serway & Jewett, 2018). For this reason, reasoning skills are needed to solve various problems related to the concept of vibration and waves. According to Pusat Penilaian Pendidikan Kemdikbud, the cognitive level of reasoning is a higher-order thinking ability that includes the ability to analyze, evaluate, create, think critically, think logically, think creatively, solve problems in new contexts (Kemdikbud, 2019). To that end, this study will develop HOTS assessment instrument devices on vibration and wave courses. It is expected that through this HOTS problem device, students will get used to solving problems with a high level of thinking skills so as to develop students' thinking skills.

METHOD

This research uses research and development (R&D) methods (Haryati, 2012; Hasnunidah & others, 2017) to developed and validation a HOTS assessment instrument in vibration and wave courses. The development model used is the 4D model (four *D* model) (Sugiyono, 2019; Thiagarajan & others, 1974). The 4D development model consists of 4 main stages, namely: (1) Define, (2) Design, (3) Develop, and (4) Disseminate. The stage of test development is presented in Figure 1.

Subject of Research

The subject of this research is the subject of product development and trial subject. The product development subjects consisted of material experts, educational evaluation experts and linguists as validators. The test subjects were one class physics education students (21 students) who had taken vibration and wave courses.

Data Collection and Instrument

Indirect communication technique using closed questionnaire was used to measure the feasibility of the developed vibration and wave HOTS test instrument. The aspects of the feasibility of the questionnaire that were measured included aspects of content, construct, language, time and instructions. The validator validation results that have been obtained are used to revise the developed questions items. The instruments used in this study have been validated by experts and declared eligible for use in research data retrieval. Measurement technique was used when testing the HOTS vibration and wave test instrument to measure the level of validity and reliability of the test.
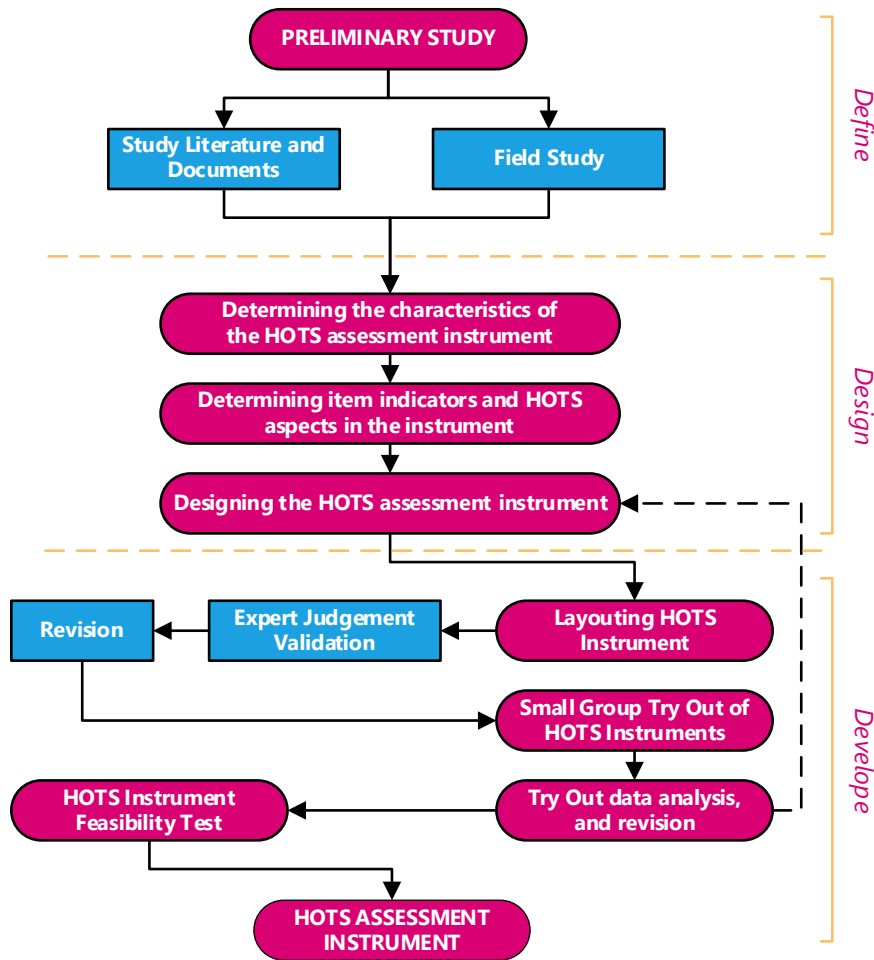
Figure 1. HOTS Assessment Instrument Development Research Design Chart

Data Analysis Methode

Analysis of research data in the form of expert validation results analyzed using the percentage of HOTS test feasibility score. The percentage of score obtained is interpreted based on the eligibility criteria (Sudijono, 2017), presented in Table 1.

Table 1 Assessment Criteria by Experts

Percentage Range	Criterion
$81.25\% < \text{score} \leq 100\%$	Excellent
$62.50\% < \text{score} \leq 81.25\%$	good
$43.75\% < \text{score} \leq 62.50\%$	Good Enough
$25.00\% < \text{score} \leq 43.75\%$	bad

Hots assessment instruments developed are considered feasible to use if the rating score $> 62.5\%$. If the score of the assessment is still ≤ 62.5 then the product needs to be revised again. The test result data was analyzed in the form of validity test, difficulty level test, differentiation index test and reliability test. The validity coefficient is obtained using the Product Moment Correlation Equation (Arikunto, 2010). The validity of HOTS items of instruments is obtained by comparing the coefficient of validity

(r_{hitung}) with apostates. Hots question item is said to be valid if $r_{hitung} > r_{tabel}$ and otherwise (Sugiyono, 2019). The difficulty level of the problem item is indicated by the difficulty level index (Arikunto, 2010). The test question items is said to be good if the problem difficulty index is in the range of 0.25-0.75 (Koyan, 2012). The differentiation of the item is used to know the level of each question item to distinguish between the upper group students and the lower group. The differentiation item of instrument is shown from the distinguishing index criteria (Arifin, 2012). According to Fernandes (Koyan, 2012), the test question item is said to be good if the problem differentiation index is in the range of 0.15 – 1.00. The question of the test is known from the reliability index of the question. The reliability index of the problem is obtained using the Alpha Cronbach equation (Surapranata, 2004), to describe the internal consistency of the instrument.

RESULT AND DISCUSSION

Test Development Results

This development research resulted in the product of high-level thinking ability assessment instrument (HOTS) for vibration and wave courses. The 13 items of the HOTS test instrument were developed in the form of multiple choice with reasons at the level of thinking, analyzing, evaluating, and creating. (Table 2).

Table 2. Item Spread High Level Thinking Ability Test Vibration and Wave Material

No	Realm of Knowledge	Problem Number	Number of Questions
1	Analyzing (C4)	2, 4, 6, 13	4
2	Evaluating (C5)	1, 3, 5, 9, 11	5
3	Create (C6)	7, 8, 10, 12	4

Based on Table 2, it appears that the problem spread for each HOTS level is relatively the same. The feasibility of HOTS questions by experts is seen from the aspects of constructs, materials and language. Table 3 shows that HOTS assessment instruments developed have feasibility in terms of content, constructs and languages with excellent criteria for content and construct while criteria are good for languages. In addition, the allocation of time given very well criteria means that the time provided in the work of the question is adequate and the instructions presented on the question have good criteria meaning the question order is easy to understand by students. However, there are several revisions (improvements) based on validator suggestions including (1) Redaction improvements to questions 1, 2 and 3 and operational verbs on Basic Competencies.

Table 3. Recapitulation of Test Results by Experts Of Each Item

No	Aspects of Study	Study Results						P Total	criterion
		Validator 1		Validator 2		Validator 3			
		P	criterion	P	criterion	P	criterion		
1	Content	80.00	good	80.00	good	53.33	Good Enough	86.67	Excellent
2	Construct	86.67	Excellent	100.00	Excellent	86.67	Excellent	91.67	Excellent
3	Language	93.33	good	95.00	Excellent	86.67	Excellent	75.56	good
4	Time	80.00	good	100.00	Excellent	80.00	good	86.67	Excellent
5	Instructions	40.00	bad	80.00	good	80.00	good	73.33	good

The results of the feasibility analysis of HOTS assessment instruments in terms of content are considered eligible to be used with very good feasibility criteria in terms of content, construct and language. This explains that the test instruments made are able to measure the high level of skills of students as evidenced by expert judgment (Istiyono et al., 2014).

Test Results

The test instrument developed is an instrument used to measure students' high level of thinking skills at the level of analyzing, evaluating and creating. According to Pusat Penilaian Pendidikan (Kemdikbud, 2019), there are 3 principles in assessing high thinking skills, namely (1) The existence of a stimulus presentation in the form of problems (cases), scenarios, discourses and introduction to text that can make it easier for student to think; (2) Using contextual and new problems for students; and (3) Distinguishing between difficulty levels (easy, moderate or difficult) and cognitive levels (low-level thinking and high-level thinking). Therefore, the HOTS developed using contextual and interesting stimulus that is sourced from global issues related to the material in question and allows students to be able to solve problems by evaluating some problem solving strategies so that students can find new solutions models different ways (Thomas & Thorne, 2009).

The results of the trial analysis showed that out of the 13 HOTS question items developed there were 9 valid criteria question items and 4 invalid criteria question items (Table 4). Valid critical questions are considered worthy to be used in measuring students' high-level thinking skills. It is as stated by (Barnett & Francis, 2012) that test questions containing high levels of valid criteria can encourage students to think deeply about learning materials. The valid test can describe the indicators of the learning material to be tested. The HOTS assessment instrument that was developed effectively in training students to have good HOTS can be used as a learning assessment for students (Merta Dhewa et al., 2017). A valid HOTS test instrument can present questions correctly to measure analytical skills, evaluation abilities and the ability to be creative in the concept of vibration and waves. so that with the right instrument it can develop students' higher order thinking skills (Rahmi & Alberida, 2017).

Table 4. HOTS Assessment Instrument Validity Test Results

Problem Number	r_{count}	r_{table}	criterion
1	0.878	0.811	Valid
2	0.878	0.811	Valid
3	0.168	0.811	Invalid
4	0.619	0.811	Invalid
5	0.814	0.811	Valid
6	0.925	0.811	Valid
7	0.989	0.811	Valid
8	0.878	0.811	Valid
9	0.878	0.811	Valid
10	-0.04	0.811	Invalid
11	0.901	0.811	Valid
12	0.835	0.811	Valid
13	0.988	0.811	Valid

The results of the analysis of the differentiation index of HOTS problem items obtained that 50% of the problem items are good category, 15% are quite good category and 26% are less good category. Differentiation analysis is needed to know the intensity of a problem in difficulty, namely the ability of the question item to distinguish between students who master the material tested and students who have not mastered the material tested (Fatimah & Alfath, 2019). The question item that has bad differentiation criteria means that the question item can be answered correctly by students who are clever or less clever so it is not good to be used in measuring students' high-level thinking ability.

Furthermore, the difficulty level of HOTS problem items developed shows that the problem is dominated by questions with a moderate category of 77% of medium category questions, 15% of easy category questions and 8% of difficult category questions. A test should not be too easy, nor should it be too difficult. An instrument has good criteria if the problem in test instrument is not too easy and not too difficult because the problem item is too easy will make all students can answer correctly and the problem item is too difficult to answer by students (Fatimah & Alfath, 2019). This indicates that the HOTS problem items developed fall into both categories in the difficulty level of the problem. There are still students who find it difficult to answer HOTS questions because students are still not familiar with the type of HOTS questions in the form of multiple choice with a reason (Agustihana & Suparno, 2019). There are still students who have difficulty answering HOTS questions because students are still not familiar with the types of HOTS questions in the form of multiple choice and the reasons. Another reason is that Higher Order Thinking Skills (HOTS) have not been properly trained and accommodated in the classroom. Students are still accustomed to using Lower Order Thinking Skills (LOTS) is the logical reason students still have difficulty answering HOTS questions, in addition to psychological and intellectual characteristics, class environment, (Budsankom et al., 2015) and inactive learning (Santi et al., 2018). Teachers should use learning models or media that can stimulate students' higher-order thinking skills (Aulia et al., 2019).

The reliability of the test category is very strong which is 0.95 which indicates that the test has high stability so that the test can be used to make decisions about students' high-level thinking ability. The reliability coefficient of the test has a large of at least 0.90 can be used to make decisions about individuals (Suryabrata, 2000). To make instrument materials according to the need for higher order thinking, that must feature novelty and be supported by strong theoretical and empirical foundation so that have consistency between components (Agustihana & Suparno, 2019). The stability of the developed HOTS instrument can be seen from its high reliability index (Mooi et al., 2018). The development of the HOTS test for vibrations and waves has met the criteria including content and construction consistently describing the stability of the HOTS instrument.

CONCLUSION

The HOTS instrument was developed in the form multiple-choice options with a reason to measure students' high-level thinking skills at the level of analyzing, evaluating and creating. The quality of the question points is obtained from the analysis of expert studies and try out analysis consisting of validity, reliability, difficulty and differentiation of the items, it can be drawn conclusions (1) Reviewed from the study of quality experts HOTS problems developed good criteria with an average percentage of content aspects of 86.67% with excellent criteria, construction aspects of 91.67% with excellent criteria and language aspects of 75.56% with good criteria; (2) In terms of validity, on HOTS question instruments as many as ten points of valid category questions and three items of invalid category questions; (3) In terms of reliability with an index of 0.95 has qualified reliability; (4) Judging from the differentiation index, there are seven points of question with good differentiation, two points of question with sufficient differentiation and four points of question with poor differentiation; (5) Judging from the difficulty level, there are 77% of medium category question items, 15% of easy category question items and 0.08% of difficult category question items. In general, 77% of HOTS items instrument are in good criteria.

Based on the results of the analysis, it is recommended: (1) Lecturers can assess high-level thinking skills in courses by applying HOTS test; (2) The existence of training on the preparation of high-level thinking skills assessment instruments; and (3) The development of HOTS-based questions for other physics courses.

REFERENCES

Agustihana, S., & Suparno, S. (2019). Development of HOTS Oriented Cognitive Problems in Thermodynamics for Senior High Schools. *Jurnal Penelitian Fisika Dan Aplikasinya (JPFA)*, *9*(1), 44. https://doi.org/10.26740/jpfa.v9n1.p44-54

Arifin, Z. (2012). Evaluasi pembelajaran: Prinsip Teknik Prosedur. In *Bandung: Remaja Rosdakarya*. Bandung: PT Remaja Rosdakarya.

Arikunto, S. (2010). *Prosedur Penelitian: Suatu Pendekatan Praktik* (6th ed.). Rineka Cipta, Jakarta.

Aulia, F., Suyatna, A., & Viyanti. (2019). STEM-Based Multimedia Design for Stimulating HOTS in Theory Bioenergy and Wind Energy: Perception Physics Teacher. *Jurnal Pembelajaran Fisika*, *7*(2), 77–85.

Barnett, J. E., & Francis, A. L. (2012). Using higher order thinking questions to foster critical thinking: A classroom study. *Educational Psychology*, *32*(2), 201–211.

Budsankom, P., Sawangboon, T., Damrongpanit, S., & Chuensirimongkol, J. (2015). Factors affecting higher order thinking skills of students: A meta-analytic structural equation modeling study. *Educational Research and Reviews*, *10*(19), 2639–2652.

Ekawati, F., Handhika, J., & Huriawati, F. (2017). Profil Kesalahan Siswa dalam Menyelesaikan Soal Getaran dan Gelombang Berbasis HOTS (Higher Order Thinking Skill). *Prosiding SNFA (Seminar Nasional Fisika Dan Aplikasinya)*, *2*, 293–298.

Fatimah, L. U., & Alfath, K. (2019). Analisis Kesukaran Soal, Daya Pembeda dan Fungsi Distraktor. *Al-Manar*, *8*(2), 37–64.

Harahap, N. J., & Rafika, M. (2020). Industrial Revolution 4.0: and the Impact on Human Resources. *Ecobisma (Jurnal Ekonomi, Bisnis Dan Manajemen)*, *7*(1), 89–96. https://doi.org/10.36987/ecobi.v7i1.1545

Haryati, S. (2012). Research and Development (R\&D) sebagai salah satu model penelitian dalam bidang pendidikan. *Majalah Ilmiah Dinamika*, *37*(1), 15.

Hasnunidah, N., & others. (2017). Metodologi Penelitian Pendidikan. *Yogyakarta: Media Akademi*.

Istiyono, E., Mardapi, D., & Suparno, S. (2014). Pengembangan tes kemampuan berpikir tingkat tinggi fisika (pysthots) peserta didik SMA. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *18*(1), 1–12.

Kartowagiran, B. (2012). Penulisan butir soal. *Yogyakarta: Universitas Negeri Yogyakarta*, 1–33.

Kemdikbud. (2019). *Panduan Penulisan Soal HOTS (Higher Order Thinking Skills)* (Asrijanty; Deni Hadiana; (ed.); 2019th ed.). Pusat Penilaian Pendidikan Kemdikbud.

Koyan, W. (2012). Konstruksi Tes: Singaraja. *Universitas Pendidikan Ganesha Press*.

Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., & Hoffmann, M. (2014). Industry 4.0. *Business \& Information Systems Engineering*, *6*(4), 239–242.

Merta Dhewa, K., 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 (IOSR-JRME)*, *7*(1), 26–32.

Mooi, E., Sarstedt, M., & Mooi-Reci, I. (2018). *Market Research, The Process, Data, and Methode Using Stata*. Springer Texts in Business and Economics, Singapore.

Neugebauer, R., Hippmann, S., Leis, M., & Landherr, M. (2016). Industrie 4.0 - From the Perspective of Applied Research. *Procedia CIRP*, *57*, 2–7.

Pfeiffer, S. (2017). The Vision of "Industrie 4.0" in the Making—a Case of Future Told, Tamed, and Traded. *NanoEthics*, *11*(1), 107–121.

Prasetyo, H., & Sutopo, W. (2018). Industri 4.0: Telaah Klasifikasi aspek dan arah perkembangan riset. *J@ Ti Undip: Jurnal Teknik Industri*, *13*(1), 17–26.

Rahmi, Y. L., & Alberida, H. (2017). Improving Students' Higher Order Thinking Skills through Portfolio Assessment on Biology Curriculum and Textbook Analysis Course. *Bioeducation Journal*, *1*(1), 22–33.

Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., & Harnisch, M. (2015). Industry 4.0: The future of productivity and growth in manufacturing industries. *Boston Consulting Group*, *9*(1), 54–89.

Santi, N. W., Suyatna, A., & Suyanto, E. (2018). Pengembangan Buku Elektronik Inti Atom Sebagai Bahan Ajar Mandiri Untuk Menumbuhkan Kemampuan Berfikir Kritis Siswa. *Jurnal Pembelajaran Fisika*, *55*(1), 134–145.

Schmidt, R., Möhring, M., Härting, R.-C., Reichstein, C., Neumaier, P., & Jozinović, P. (2015). Industry 4.0-potentials for creating smart products: empirical research results. *International Conference on Business Information Systems*, 16–27.

Serway, R. A., & Jewett, J. W. (2018). *Physics for scientists and engineers*. Cengage learning.

Sudijono, A. (2017). *Pengantar Evaluasi Pendidikan*. Raja Grafindo Persada, Jakarta.

Sugiyono. (2019). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D* (28th ed.). CV. Alvabeta, Bandung.

Surapranata, S. (2004). *Analisis, Validitas, Reliabilitas, Dan Interpretasi Hasil Tes*. Remaja Rosda, Bandung.

Suryabrata, S. (2000). Pengembangan alat ukur psikologis. In *Yogyakarta: Penerbit Andi*.

Thiagarajan, S., & others. (1974). *Instructional development for training teachers of exceptional children: A sourcebook.*

Thomas, A., & Thorne, G. (2009). How to increase higher order thinking. *Metarie, LA: Center for Development and Learning*.

Zhou, K., Liu, T., & Zhou, L. (2015). Industry 4.0: Towards future industrial opportunities and challenges. *2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, 2147–2152.