Dinamika Pendidikan 16 (2) (2021) 194-203

Dinamika Pendidikan

http://journal.unnes.ac.id/nju/index.php/dp

Development of Higher Order Thinking Skills Based Assessment Instruments Using Live Worksheet Application

Ratih Octaviana Puspitasari✉, Siti Sri Wulandari

DOI: 10.15294/dp.v16i2.33433

Department of Economics Education, Faculty of Economics, Universitas Negeri Surabaya, Indonesia

History Article

Submitted 2021-10-22
Revised 2021-11-25
Accepted 2021-12-26

Keywords

Assessment Instruments; HOTS; Live Worksheet Application; Sugiyono; Development; Multiple Choice Test

Abstract

The research aimed to develop higher-order thinking skill (HOTS) based assessment instruments based on Live Worksheet applications on Creative Products and Entrepreneurship (PKWU) subjects, especially on Basic Competencies 3.2 Analyzing business opportunities and Basic Competencies 3.3 Applying business administration documents. This research used Sugiyono's development model with seven development steps. The research subject consisted of 34 students of class XI Department of Office Governance Automation SMKS Ketintang Surabaya. The development of this assessment instrument resulted in 20 questions that were declared valid with the calculation of r calculated > r table (0.3291). The instrument developed was said to be reliable because the reliability value of the problem of 0.840 > 0.60 The results of product validation development from three expert validators obtained an average validation of 95.6% with "very strong" criteria. The average student responded to the use of the Live Worksheet application "very positively by 88%. As per results of this study trial, the average ability to think high-level students fell into the category of "less". Therefore, the learning process to improve high thinking skills in students must be further improved with a learning model or method that can encourage the improvement of students higher-order thinking skills.

How to Cite

Puspitasari, R. O., & Wulandari, S. S. (2021). Development of Higher Order Thinking Skills Based Assessment Instruments Using Live Worksheet Application. *Dinamika Pendidikan,* 16(2), 194-203.

✉ Correspondence Author:
Jl. Ketintang No. 2, Ketintang, Kec. Gayungan, Surabaya
E-mail: ratih.18015@mhs.unesa.ac.id

p-ISSN 1907-3720
e-ISSN 2502-5074

INTRODUCTION

Education is a component of the dynamic living conditions of individuals and societies, constantly changing over time according to the demands of personal competence (Suratman et al., 2019). In the 21st century, the demands for quality improvement in the field of education are very complex and challenging (Retnawati et al., 2018). According to Kurniawan & Lestari (2019) said that the demands of the 21st century have to do with increasing the ability to think clearly. These demands bring about changes in the implementation of education, which requires students to be active, think critically, structured and creative during the learning process (Jannah & Pahlevi, 2020). Changes in the implementation of education lead to high-level thinking skills called Higher Order Thinking Skills (HOTS) which are part of the educational goals of the 21st century (Kemdikbud, 2018).

Higher Order Thinking Skills (HOTS) are skills that not only require memory and memorization skills but need other skills (Wardany & Ramli, 2017). One of the problems that is currently happening is that the high-level thinking ability of Indonesian students is still low, at the international level, based on the results of the Programme for International Student Assessment (PISA) examination conducted by the Organisation for Economic Co-operation and Development (OECD), it is mentioned that Indonesian students are ranked 64 out of 70 countries (Suwarna & Fatimah, 2018). Therefore, the quality of education in Indonesia must be improved (Suratman et al., 2020).

Efforts to improve the quality of education in Indonesia were implemented in the 2013 curriculum as an update to the previous curriculum (Septianingrum & Indartono, 2019). Curriculum 2013 as an update of the Education Unit Level Curriculum (KTSP) has the goal of preparing students to be able to be productive, creative, innovative, and effective through the learning stages where students are initially given information turned into students looking for information themselves (Fauzi, 2015). Therefore, the pattern of habituation of high-level thinking is very important to realize the 2013 curriculum goals (Arifin & Retnawati, 2017). According to Suparno (2018) he stated that by thinking critically students will try to find solutions, find benefits, and be able to provide ideas to solve everyday problems.

One of the important learning planning in the 2013 curriculum is the evaluation of learners' learning outcomes (Nelly Rhosyida, Mahmudah Titi Muanifah, Trisniawati, 2021). According to Desilva et al. (2020) they stated that learners' learning outcomes can be evaluated using an assessment instrument. In the opinion of Asrul et al. (2014), there are 2 forms of assessment instruments, namely the form of the test and the non-test form. Test form assessment instrument, according to Trisnamansyah (2014)especially within recurrent neural network (RNN there are four types: 1) the form of the problem is right or wrong, 2) the form of the problem is multiple choice, 3) the form of the problem is matching, 4) the form of the short answer question (stuffing). Trisnamansyah (2014)especially within recurrent neural network (RNN also stated that there are several types of non-test instruments, including observation, questionnaires, interviews, checklists, and rating scales that can be used to correct deficiencies in test form assessment instrument techniques. Therefore, it can be concluded that assessment instruments are an important part of learning activities to know the quality of students' learning skills against the material that has been studied.

Given the importance of assessment instruments in learning activities, the assessment instrument that can answer the objectives of the 2013 curriculum is an assessment instrument in the form of HOTS questions (Desilva et al., 2020). The process of preparing HOTS-based assessment instruments requires educators to develop skills in understanding the dimensions of bloom taxonomic thinking in the cognitive realm that lead to high-level thinking skills, namely analyzing (C4), evaluating (C5), and creating (C6) (Trisnawati et al., 2020). In addition, according to Widana (2017) to compile a HOTS-based assessment instrument educators must have the ability to master teaching materials, be able to construct problems, and be creative in innovating stimulus in accordance with the surrounding problems. There are three aspects that must be considered in the preparation of HOTS based assessment instruments, namely: 1) aspects of language, 2) aspects of content/ material, and 3) aspects of construction/ evaluation.

Efforts that can be used to implement the 2013 curriculum objectives related to students high-level thinking skills one of the steps that can be taken is to pay attention to the quality of student learning related to how students think and provide solutions to a problem. Because according to Hamidah & Wulandari (2021) they stated that in this way students will indirectly use their reasoning skills to find solutions critically.

The educational institution that has led to 21st-century education is SMKS Ketintang by implementing the 2013 curriculum.

The results of preliminary studies that have been carried out by researchers, it was found that teachers had not applied HOTS-based assessment instruments on Creative Products and Entrepreneurship (PKWU) subjects on basic competencies 3.2 Examining business opportunities and fundamental competencies 3.3 Making use of business administration documents. The Basic Competence material explains how to plan business products, identify goods/services, and plan business document because the mastery of materials on basic competencies is very important for students to provide skills in the business/industry later. So, for these competencies it is necessary to develop HOTS-based assessment instruments.

It is also found that the type of problem that is often used by teachers is a matter of description. In addition, assessment activities that are often used by teachers are also still conventional, namely students download assignments and then work on written problems, and then upload the answer sheet again. Hamidah & Wulandari (2021) said that in conventional assessment there are several weaknesses, one of which is less effective and efficient in the time of anorexia. In the 21st century as it is today, assessment is not only done conventionally but there have been widely developed applications that can be used for assessment activities, called media Information and Communication Technology (ICT).

By utilizing ICT, the media is able to minimize deficiencies in conventional assessment because it is more interesting and interactive (Wardoyo, 2016). One of the ICT media is the live worksheet. The live worksheet is a web-based online platform that can provide alternative material delivery in the form of worksheets that students can access and respond to online without having to be downloaded first. The live worksheet can be accessed via mobile phone or computer for free using an internet connection and is fairly light and does not spend internet quota (Nelly Rhosyida, Mahmudah Titi Muanifah, Trisniawati, 2021).

In addition to attaching importance to learning outcomes and assessment instruments provided, students also need to feedback as an evaluation of the learning process provided by educators. The student's response in this study was used for evaluation of the use of live worksheets in assessment instruments. According to Jannah & Pahlevi (2020) student response is a response or feedback from students as a result of the stimulus provided by educators during the learning process. The learning process is called successful if it has gone through the process of assessing learning outcomes.

The previous research, namely the development of HOTS-made assessment instrument developed by Jannah & Pahlevi (2020) stated that the assessment instrument that had been developed was valid and could measure students high thinking levels with the acquisition of 92% of the average validation from three expert validators, thus being included in the category of "very strong". Furthermore, the results of research conducted by Nelly Rhosyida, Mahmudah Titi Muanifah, Trisniawati (2021) showed that live worksheets application were worth using as a practical, fun assessment tool and increased students learning motivation.

The novelty in this study is that the development of assessment instruments is presented by using an online live worksheet platform where the use of live worksheets can minimize conventional assessment activities to facilitate educators. Based on the background above, it is necessary to develop higher-order assessment instruments based on thinking skills assisted by live worksheets on PKWU subjects for OTKP majors in SMKS Ketintang that can be used to train students' high-level thinking skills. The purposes of the study are: 1) developing higher-order thinking skills based on assessment instruments using live worksheets on PKWU subjects, 2) testing the feasibility of higher-order thinking skills-based assessment instruments developed, 3) analyzing students' high-level thinking skills, and 4) analyzing students' responses to the use of live worksheets. So it can be concluded this development research is expected to be used as input and contribution to the world of education, especially regarding the development of assessment instruments in the form of higher-order thinking skills based questions for PKWU subjects.

METHODS

The type of research used in this research was the development research method (R&D). According to Sugiyono (2015), the development research method (R&D) is a research method to produce products and test the effectiveness of products to know the effectiveness and feasibility of the product. The development model used in this research was the Sugiyono development model, which consists of 10 stages, but in this study only used 7 stages of development, namely analysis of potential problems, collecting data, pro-

duct design, design validation, design revision, product trials, and product revisions.

This research only used 7 stages of development because this research was only conducted at SMKS Ketintang Surabaya and only developed on creative products and entrepreneurship (PKWU) subjects in KD 3.2 and KD 3.3 so that it was not widely disseminated. The test subject in this study was a class XI OTKP student consisting of 34 students. The development stage of HOTS-based assessment instruments on this study is presented in the following Figure 1.

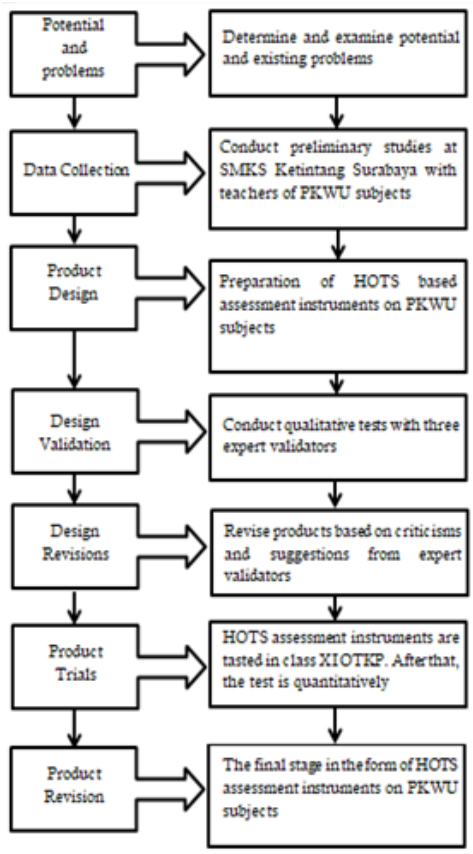

Figure 1. The flow of the research and development model adopted from Sugiyono

The explanation of each stage of development in figure 1 is in the first process of determining potential and problems, namely researchers determine the place of research and examine the potential and problems that exist. The second process is to collect data through a preliminary study to the source, namely, teachers of PKWU subjects to collect information about assessment instruments used in schools. The third is that the product design process is carried out by formulating a question grid and preparing a HOTS-based question item instrument in the form of multiple-choice tests. The fourth stage is design validation by expert validators, namely material experts, linguists, and evaluation experts, to test product feasibility. Furthermore, the fifth process revises the design following the suggestions and criticisms of each expert validator so that it is ready for trial. The sixth process is to conduct product trials on students of class XI OTKP, but previously HOTS-based assessment instruments have first been inputted on the live worksheet. For the last stage, the seventh stage is to revise the product to produce HOTS-based assessment products on PKWU subjects for KD 3.2 and KD 3.3.

This study used two types of data, namely qualitative data and quantitative data. Qualitative data is obtained from the results of validation sheets of material experts, linguists, and evaluation experts as well as criticisms and suggestions given by each validator. Furthermore, quantitative data was obtained from the results of the analysis of problem points include validity, reliability, difficulty level of the problem, differentiating power of the problem, problem distractor, and analysis of students' high-level thinking skills. This study used research instruments in the form of interviews, expert validation sheets, and student response sheets. This type of interview was conducted with teachers of Creative Products and Entrepreneurship (PKWU) subjects at SMKS Ketintang Surabaya.

The technical data analysis, the methods were divided into qualitative data analysis and quantitative data analysis. Qualitative data was obtained from validator assessments which were then analyzed with the Guttman scale with the answer "Yes" will get a value of one, while for the answer "No" will get a value of zero (Riduwan, 2016). For the results of the calculation of validation values experts expressed the criteria "Very Weak" when obtaining results of between 0%-20%, the criterion "Weak" when obtaining the result 21%-40%, the criterion "Sufficient" when obtaining the result 41%-60%, the criterion "Strong" when it is at the result 61%-80%, and for the condition "Very Strong" when obtaining the result of 81%-100%. Kinanti & Subagio (2020) the feasibility of assessment instruments by validators is at least on the criteria of "Sufficient" which is to obtain results between 41%-60%.

As for quantitative data it was obtained from the results of analysis of 30 points of questions that had been developed. Analysis of the first problem item is validity, the problem item can be declared to meet the valid aspect if the r calculation is greater than the value of the r table (0.3291) (Kurniawan & Lestari, 2019; Maulidia & Pahlevi, 2020; Aniq Rif'atun Najihah, Vina

Serevina, 2018).

The second problem analysis is reliability, while reliability obtains results 0.80-1.00 is on the criteria "Very High", the result 0.60-0.79 entered the "High" criteria, for the result 0.40-0.59 the criterion "Sufficient", and for the result of the value between 0.20-0.39 is on the low criteria, while the criterion "Very Low" when obtaining a value of 0.00-0.19 (Arifin, 2012).

The third problem analysis is the difficulty level of the problem. As for problems that have a value of more than 0.90, they are considered very easy so that the problem is rejected, then the problem that has a value of 0.71-0.89 is included in the easy criteria so that the problem is considered less good and needs improvement, the problem that has a value between 0.31-0.70 is on the medium or sufficient criteria and requires improvement, while the problem that gets the result of 0.21-0.30 is on the difficult criteria, which is less good and needs improvement, then for problems whose value is less than 0.20, they are included in the criteria that are very difficult so that the problem is rejected (Sabina Ndiung, 2020).

Next is the differentiating power of the matter. Problems that have a differentiating power result of more than 0.40 are on the criteria very well so that the problem is accepted, for problems with differentiating power results get a value between 0.30-0.39 enter the criteria well so that the problem is accepted with improvement, the problem that gets a differentiating power value of 0.20-0.29 enters the criteria enough and requires improvement, while the problem with a differentiating power value of less than 0.19 is on the criteria not good so that the problem is discarded (Sabina Ndiung, 2020).

Analysis of the next question item is a question distractor that can be seen from the answer pattern that has been selected by students. Arifin & Retnawati (2017) stated that problem distractor can be said to be effective if select at least 5% of the number of test-takers.

The next quantitative data is the analysis of students' high thinking skills according to Purbaningrum (2017) can be known after students work on HOTS-based problems. Then for students who score high in thinking skills between 81-100 included in the criteria "Very Good", students who score between 61-80 are on the criteria "Good", students whose grades are 41-60 included in the criteria "Enough", students whose grades are 21-40 are on the criteria "Less", and students whose grades are less than 20 falls into the criteria "Very Less" (Purbaningrum, 2017).

Then after the end of the work on the problem students conduct the student response analysis stage when using the live worksheet in assessment activities. Students who respond more than 85% are included in the "Very Positive" response criteria, for students who respond 70%-85% are on the "Positive" response criteria, students who give a response of 50%-70% are included in the "Less Positive" response criteria and students who give responses less than 50% are included in the "Not Positive" response criteria (Aisyah et al., 2015).

RESULTS AND DISCUSSION

HOTS-Based Instrument Development Process

Higher Order Thinking Skills (HOTS) based assessment instruments using live worksheet applications in PKWU KD 3.2 and 3.3 subjects were developed by using Sugiyono's development model with ten stages of development, but limited to 7 stages of development only. The stages of development are: *First*, potential and problems. At this stage, researchers determined the location of the study and analyzed potential and problems in SMKS Ketintang Surabaya. Based on the results of interviews conducted by researchers it obtained results that in PKWU subjects in SMKS Ketintang Surabaya had not implemented HOTS-based assessment instruments. In addition, learners were not used to arguing and solving factual problems, because educators gave the form of problems limited to the ability to understand and remember. Then at SMKS Ketintang Surabaya already had an ICT -based learning facility (Information and Communication Technology) called learning SMK Ketintang.

Second, data collection. At this stage, the researcher conducted an unstructured interview with PKWU teachers as sources. The results of the interview showed that the assessment instrument used by the teacher was still in the form of a description and only racing on theory so that students were not used to working on forms of factual problem-solving problems, as well as assessment activities carried out by teachers, were still conventional.

Third, product design. At this stage, researchers compiled problems based on Higher Order Thinking Skills (HOTS) in the form of multiple-choice questions. According to Widiyawati et al., (2019) multiple-choice problems could be used to measure students' high-level thinking skills and measure more complex student learning outcomes. In addition, multiple-choice assessment instruments were often used because they were easier, objective, and easier in data proces-

sing (Hartini & Sukardjo, 2015). The stages of preparing the product design were analyzing the KD to be made about HOTS, creating a grid of HOTS problems, compiling questions that fit the grid of questions using interesting and contextual stimulus, and finally compiling answer keys. The assessment instrument that the researchers developed was 30 questions in the cognitive realm of C4, C5, and C6 which were then taken 20 of the best questions.

Fourth, design validation. The validation stage of the assessment instruction design was carried out by three expert validators, namely material validators who are PKWU subject teachers at SMKS Ketintang Surabaya, language validators who are teachers Indonesian at SMKS Ketintang Surabaya, and evaluation validators who are lecturers of assessment experts at Surabaya State University. Following the results of the assessment, it obtained a value of 97.5% with the criteria "very strong" from material experts, then it obtained a value of 92.5% with the criteria "very strong" from linguists, and it obtained a value of 96.7% with the criteria "very strong" from the evaluation expert. So that the average value of expert validation results was 95.6% with "very strong" criteria. In addition, criticism and advice were also given by experts on HOTS-based assessment instruments to be improved so that it became a product that was worth testing.

Fifth, design revision. At this stage, improvements were made to the assessment instruments developed by researchers following the criticism and advice given by validators to become a product that was worth testing. Before the trial, the HOTS assessment instrument was first created in the live worksheet application on the "create" menu. Here's a look at the HOTS-based scoring instrument in the live worksheet application.

Figure 2. HOTS-based Assessment Instruments in Live Worksheet Application

Sixth, product testing. At this stage, the improved instruments were tested on 34 OTKP class XI students. The assessment instrument was tested with the help of a live worksheet application and was implemented in the classroom. The problem was displayed on the smartphone screen of each student. After the student finished answering the question, the results of the work could be seen by students, and then the results of the student's work would be directly sent to the teacher's live worksheet account. Then the researcher conducted a quantitative hots-based assessment instrument quality test by testing the validity, reliability, difficulty level of the problem, the differentiating power of the problem, and the distractor of the problem using the item application. The quantitative test aimed to produce valid and decent problems.

Seventh, product revision 1. At this stage, 20 valid questions were obtained and 10 invalid questions were based on the results of quantitative data analysis. Therefore, the 20 best questions were taken as the final product of the HOTS-based assessment instrument developed by the researcher. As for the problem that was said to be feasible there was at number 4, 5, 6, 7, 8, 9, 10, 13, 14, 15, 17, 18, 19, 20, 21, 22, 23, 24, 26, 27, 29, while the unworthy question must be discarded at number 1, 2, 3, 11, 12, 16, 25, 28, 30, 14. In line with the opinions of Hamidah & Wulandari (2021), Desilva et al. (2020), Jannah & Pahlevi (2020) and Kurniawan & Lestari (2019) which also developed HOTS-based assessment instruments with the form of multiple-choice questions establishing and taking the best questions after being declared valid and feasible.

Feasibility of HOTS Instruments Assisted by Live Worksheet Application on Basic Competencies 3.2 and 3.3

The feasibility of higher-order thinking skills (HOTS) based assessment instrument products were derived from qualitative and quantitative test results. As for qualitative test results, it obtained from the results of expert validators, namely material experts, linguists, and evaluation experts. This was in line with the opinions of Hamidah & Wulandari (2021), Desilva et al. (2020), Jannah & Pahlevi (2020), Kurniawan & Lestari (2019), and Widana (2017) which stated that aspects of validation of HOTS-based assessment instruments can be seen based on material aspects, language aspects, and evaluation aspects. The results of validator assessments are presented in the Table 1.

Table 1. Expert Validation Results

Validator	Assessment Validator	Criteria
Material Validator	97,5%	Very Strong
Language Validator	92,5%	Very Strong
Evaluation Validator	96,7%	Very Strong
Everage	95,6%	Very Strong

Source: Primary Data Processed (2021)

The Results of validator assessment, it was known that from the material expert got a score of 97.5% with the criteria "very strong", from linguists got a value of 92.5% with the criteria "very strong", and the assessment of the evaluation expert got a value of 96.7% with the criteria "very strong", therefore it can be concluded that the assessment results of validators obtained an average score of 95.6% with criteria "very strong".

The results of feasibility test of HOTS-based assessment instruments quantitatively which included tests of validity, reliability, difficulty level of problems, problem differentiating power, and problem distractors conducted using *iteman* applications. What is the validity of the matter according to Arifin (2012) this is one of the criteria in determining a good measuring instrument. The validity of the problem in the *iteman* application can be seen in the Point Biser column. The problem can be valid if the r count is greater than the table r.

Based on the product trial, 21 points of the problem r count was greater than the r *table* (0.3291), and 9 invalid problems were number 1 with a value of 0.166 < 0.3291, problem number 2 with a value of -0.077 < 0.3291, problem number 3 with a value of 0.232< 0.3291, problem number 11 with a value of 0.306 < 0.3291, problem number 12 with a value of 0.263 < 0.329, number 16 with a value of 0.216 < 0.3291, number 25 with a value of 0.280 < 0.3291, number 28 with a value of 0.261 < 0.3291, and number 30 with a value of 0.2291 0.182 < 0.3291.

The reliability of the problem can be known in the Alpha table contained in the item application. A problem can be said to be reliable if the alpha value is greater than 0.60. According to the results of the trial that had been done the reliability value of the problem was 0.840. The value was greater than 0.60 so that the assessment instrument developed by the researcher was declared reliable with "very high" criteria.

Next was the difficulty level of the problem can be known in the Prop Correct table. Based on the results of trials that had been conducted, the average difficulty level of the problem was on the "moderate" criteria. There were 4 "difficult" questions, 21 "moderate" questions, 2 simple questions and 3 "very difficult" questions. The results were in line with Arifin (2012) which stated that good problems are problems that are not very difficult and not very easy so that the problem with the criteria "very easy" and "very difficult" will be discarded.

Furthermore, the differentiating power of the problem can be known in the Biser table. As for the trials that had been done, it showed that not all differentiating power was in the criteria of "very good". There were 21 questions included in the "very good" criteria, 5 questions included in the "good" criteria, 3 questions included in the "enough" criteria, 1 question included in the criteria of "less" problem number 2 with a value of -0.098 so that the problem with the criteria "less" would be discarded. Arifin (2012) argued that the higher the proportion of differentiating power, the better the problem.

The next was the affectivity of distractor that can be known on the Prop Endorsing table in the *iteman* application. Following the opinion of Arifin & Retnawati (2017) question, distractor can be said to be effective if selected 5% of the number of students. Based on the results of trials that had been conducted by question distractor, it was said to have worked well because overall the choice of answers had exceeded 5%, but there were 2 questions with the complainer did not work, namely question number 1 because for the choice of answer D was not selected 5% of all students and problem number 3 was in the choice of answer A and D was not selected 5% of all students, Therefore, the problem must be discarded because it was an invalid matter.

Analysis of Students' High-Level Thinking Skills Judging from the Results of HOTS-Based Problem Work

Students' high-level thinking skills according to Purbaningrum (2017) is the ability of students to analyze, evaluate and create. Students' high-level leveraged abilities can be measured after students have finished working on HOTS-based assessment instruments. Based on the results of the trials that had been conducted, the high-level thinking skills of students of class XI OTKP at SMKS Ketintang Surabaya can be described in the Figure 3.

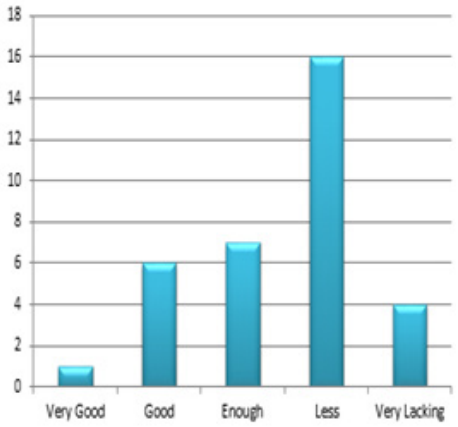

Figure 3. Diagram of Students' High-Level Thinking Ability Assessment Results

Based on the results of student answers can be known the results of the assessment of students' high-level thinking skills according to the following Table 2.

Table 2. Student Grade Results

Number of Students	Interpret Student Score	Criteria
1	81-100	Very Good
6	61-80	Good
7	41-60	Enough
16	21-40	Less
4	<20	Very Lacking

Source: Primary Data Processed (2021)

The results of the grades obtained by students showed that there were students who got grades in the categories "less" and "very less". This can be caused by several factors. The factors referred to according to Widiyawati et al., (2019) that is; 1) limitations of conceptual understanding, 2) limitations of thinking, 3) limitations of reasoning, and 4) limitations of skills in solving problems. In line with opinions Arifin & Retnawati (2017) another factor that causes students to think less critically is because learners are less accustomed to working on HOTS-based problems, so it needed teacher skills in providing a more interactive and innovative learning model in compiling HOTS-based questions so that students are accustomed to critical thinking. The results of the research that had been conducted by researchers were in line with the results of research conducted by Kurniawan & Lestari (2019) which also used HOTS-based assessment instructions to find out students' high-level thinking skills showed results that students' high thinking level abilities were diverse.

Student Response to the Use of Live Worksheet Application

Based on the results of the student response test to the use of live worksheet applications for the work of higher-order thinking skills (HOTS) based assessment instruments, the results showed that the average student gave a "very positive" response. Here's a table of student response results.

Table 3. Student Response Results

Indicators		
Responses	Reactions	
Format and Relevance	Attention and Satisfactiion	Confident
88%	87%	91%
Average	88%	

Source: Primary Data Processed (2021)

Table 3 showed that for the indicator of format and relevance responses obtained a value of 88% with the criteria "very positive", meaning that the problem displayed in the live worksheet application was interesting because there was a stimulus of images and videos. Furthermore, on the reaction indicator of attention and satisfaction criteria obtained a score of 87% with the criteria "very positive", meaning that the live worksheet application was an application that was interesting and easy to use so that students would be interested in working on the problem through the live worksheet application. On the criteria of confidence got a percentage of 91% with the criteria "very positive". As per the table of student response results above, it can be concluded that the use of live worksheet applications in implementing higher-order thinking skills (HOTS) based assessment instruments received a "very positive" response with an average response assessment of 88%.

CONCLUSION

The results of research that had been conducted by researchers on the development of Higher Order Thinking Skills (HOTS) based assessment instruments on PKWU subjects using live worksheets can be concluded that: 1) Assessment instruments developed in KD 3.2 Analyze Business Opportunities and KD 3.3 Apply Business Administration Documents with the form of

multiple-choice questions in the cognitive realm C4, C5 and, C6 number 30 questions which were then taken 20 best questions with assisted live worksheet application, 2) qualitative instrument feasibility test that was the assessment of three validators obtained an average of 95.6% with "very strong" criteria, 3) The high thinking ability of students in class XI majoring in OTKP at SMKS Ketintang Surabaya averages in the criteria of "less", 4) the use of live worksheet applications got a percentage of 88% with the criteria of "very positive" response, 5) live worksheet had the advantage of easy access because students could do problems online without having to download first, could be accessed via mobile phone or computer for free and the use of the internet was fairly light. Spend internet quotas.

The research results showed two theoretical and practical implications. For theoretical implications shown from the results of assessment instruments that had been made had fulfilled several aspects of the feasibility test of the problem in terms of language, material, and problem construction, so that the assessment instruments that have been made can increase the ability of high thinking levels of students. Then the use of live worksheet applications can be a form of assessment innovation that can attract learners to work on the given problem. While the practical implications are shown from the results of research that has been done can be an input for educators to be more innovative in the preparation of problem stimulus, improve the quality of problems and better familiarize the use of ICT media in the assessment process, and for prospective educators, the results of this research can be used as study materials to compile appropriate HOTS assessment instruments.

Limitations in this study was development models applied by using the sugiyono model which is limited to only seven stages of development only. So based on this research, advice that can be given to the next researcher is, 1) the development of HOTS assessment instruments using the Sugiyono model can be complete ten steps, 2) can test assessment instruments with a wider subject, 3) can use ICT media, in addition, to live worksheets to provide variations in assessment activities.

REFERENCES

Aisyah, Panjaitan, R. G. P., & Marlina, R. (2015). Respon Siswa terhadap Media *E-Comic Bilingual* Sub Materi Bagian-Bagian Darah. *Jurnal Pendidikan dan Pembelajaran*, *5*(3), 1–12.

Aniq Rif'atun Najihah, Vina Serevina, M. D. (2018). The Development of High Order Thinking Skills (HOTS) Assessment Instrument for Temperature and Heat Learning. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, *4*(1), 19–26. https://doi.org/10.21009/1.04103

Arifin, & Retnawati. (2017). Pengembangan Instrumen Pengukur *Higher Order Thinking Skills* Matematika Siswa SMA Kelas X. *PYTHAGORAS: Jurnal Pendidikan Matematika*, *12*(1), 98. https://doi.org/10.21831/pg.v12i1.14058

Arifin, Z. (2012). *Evaluasi Pembelajaran* 2 (Edisi R).

Asrul, Ananda, R., & Rosinta. (2014). Evaluasi Pembajalaran. In *Ciptapustaka Media*.

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

Fauzi, A. (2015). Pengembangan Media Pembelajaran *E-Book* pada Materi Jenis dan Karakteristik Bahan Baja Kelas X TGB SMK Negeri 1 Mojokerto. *Pendidikan Teknik Bangunan*, *1*(1), 148–151.

Hamidah, M. H., & Wulandari, S. S. (2021). Pengembangan Instrumen Penilaian Berbasis HOTS Menggunakan Aplikasi "Quizizz." *Efisiensi : Kajian Ilmu Administrasi*, *18*(1), 105–124. https://doi.org/10.21831/efisiensi.v18i1.36997

Hartini, H., & Sukardjo, S. (2015). Pengembangan *Higher Order Thinking Multiple Choice Test* untuk Mengukur Keterampilan Berpikir Kritis IPA Kelas VII Smp/Mts. *Jurnal Inovasi Pendidikan IPA*, *1*(1), 86. https://doi.org/10.21831/jipi.v1i1.4535

Kemdikbud, P. (2018). *Pembelajaran*.

Kinanti, M. D., & Subagio, F. M. (2020). Pengembangan LKPD Bahasa Inggris Berbantu Aplikasi Quizizz Kelas Iv Sekolah Dasar. *Jurnal PGSD*, *8*(3), 539–548. https://jurnalmahasiswa.unesa.ac.id/index.php/jurnal-penelitian-pgsd/article/view/35319

Kurniawan, R. Y., & Lestari, D. (2019). The Development Assessment Instruments of Higher Order Thinking Skills on Economic Subject. *Dinamika Pendidikan*, *14*(1), 102–115. https://doi.org/10.15294/dp.v14i1.19226

Maulidia, F., & Pahlevi, T. (2020). Pengembangan Instrumen Penilaian Tes Soal Pilihan Ganda Berbasis HOTS pada Mata Pelajaran Administrasi Umum Jurusan OTKP SMK Negeri 1 Lamongan. *Jurnal Pendidikan Administrasi Perkantoran (JPAP)*, *8*(1), 136–145. https://journal.unesa.ac.id/index.php/jpap/article/view/8114

Nelly Rhosyida, Mahmudah Titi Muanifah, Trisniawati, R. A. H. (2021). Mengoptimalkan Penilaian dengan *Liveworksheet* pada Flipped Classroom di SD. *Jurnal Taman Cendekia*, *05*(01), 568–578.

Purbaningrum, K. A. (2017). Kemampuan Berpikir Tingkat Tinggi Siswa SMP dalam Pemecahan

Masalah Matematika Ditinjau Dari Gaya Belajar. *Jurnal Penelitian Dan Pembelajaran Matematika*, *10*(2), 40–49. https://doi.org/10.30870/jppm.v10i2.2029

Retnawati, H., Djidu, H., Kartianom, Apino, E., & Anazifa, R. D. (2018). Teachers' Knowledge about Higher-Order Thinking Skills and its Learning Strategy. *Problems of Education in the 21st Century*, *76*(2), 215–230. https://doi.org/10.33225/pec/18.76.215

Riduwan. (2016). *Skala Pengukuran Variabel-variabel Penelitian*. CV Alfabeta.

Sabina Ndiung, M. J. (2020). Pengembangan Instrumen Tes Hasil Belajar Matematika Peserta Didik Sekolah Dasar Berorientasi pada Berpikir Tingkat Tinggi. 10 (June), 94–111. https://doi.org/10.25273/pe.v10i1.6274

Septianingrum, A. D., & Indartono, S. (2019). The Implementation of Archive Learning Based on the Kurikulum 2013. *Dinamika Pendidikan*, *14*(1), 87–101. https://doi.org/10.15294/dp.v14i1.18755

Sugiyono, P. D. (2015). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D* (22nd ed.).

Suparno. (2018). Development of E-Book Multimedia Model to Increase Critical Thinking of Senior High School Students. *Dinamika Pendidikan*, *12*(2), 196–206. https://doi.org/10.15294/dp.v12i2.13567

Suratman, B., Wulandari, S. S., & Nugraha, J. (2019). Relevance of Office Administrative Education and Vocational High School Curriculum to Improve Teacher Learning : Empirical Study from Indonesia. *International Journal of Applied Business and Economic Research*, *15*(April), 263–272. http: www.serialsjournal.com

Suratman, B., Wulandari, S. S., Nugraha, J., & Narmaditya, B. S. (2020). Does Teacher Certification Promote Work Motivation and Teacher Performance? a Lesson from Indonesia. *International Journal of Innovation, Creativity and Change*, *11*(10), 516–525.

Suwarna, I. P., & Fatimah. (2018). Implementation of Digital Assignments to Improve High Order Thinking Skills (Hots) Ability of Senior High School Students in The Concept of Newton's Law. *Edusains*, *10*(2), 335–340.

Tan, S. Y., & Halili, S. H. (2015). Effective Teaching of Higher-Order Thinking (HOT) in Education. *The Online Journal of Distance Education and E-Learning*, *3*(2), 41–47.

Trisnamansyah, S. (2014). *Evaluasi Pembelajaran* (Vol. 148).

Trisnawati, N., Suratman, B., Ranu, M. E., Pahlevi, T., & Wulandari, S. S. (2020). Higher-Order Thinking-Based Question Preparation Training for Office Administration Teacher's Group at Magetan. *Jurnal Pemberdayaan Masyarakat Madani*, *4* (2), 279–294.
https://doi.org/doi.org/10.21009/JPMM.004.2.8

Wardany, K., & Ramli, M. (2017). Pengembangan Penilaian untuk Mengukur Higher Order. *6*(2).

Wardoyo, C. (2016). Developing Learning Media Based on E-Learning on Accounting Subject for Senior High School Students. *Dinamika Pendidikan Unnes*, *11*(2), 84–93.
https://doi.org/10.15294/dp.v11i2.893

Widana, I. W. (2017). *Modul penyusunan soal HOTS*.

Widiyawati, Y., Nurwahidah, I., & Sari, D. S. (2019). Pengembangan Instrumen *Integrated Science Test Tipe* Pilihan Ganda Beralasan untuk Mengukur HOTS Peserta Didik. *Saintifika*, *21*(2), 1–14. issn: 1411 – 5433