Journal of Educational Science and Technology
Volume 8 Number 1 April 2022 page 74-86
p-ISSN:2460-1497 and e-ISSN: 2477-3840
DOI: https://doi.org/10.26858/est.v8i1.25910

Analysis of Instrument Development Recount Text Writing Test HOTS for Vocational High School using Exploratory Factor Analysis (EFA)

Muhammad Haris Setiawan[1], Supriyadi[2], Dwi Anggani L. Bharati[3]
[1] Educational Research and Evaluation, Universitas Negeri Semarang, Indonesia
Email: mharissetiawan@students.unnes.ac.id
[2] Educational Research and Evaluation, Universitas Negeri Semarang, Indonesia
Email: supriyadi@mail.unnes.ac.id
[3] English Education, Universitas Negeri Semarang, Indonesia
Email: dwianggani@yahoo.com

**(*Received*: 01-12-2021; *Reviewed:* 01-2-2022; *Accepted:* 26-03-2022;
Available online: 28-04-2022; *Published:* 29-04-2022)**

Abstract. This study aims to determine the construct validity and reliability of the instrument recount text writing test HOTS in English subjects for vocational high school students. This method used a development research design from Djemari Mardapi which has been modified with the three steps. The analysis of the construct validity instrument is based on how many factors are needed to explain the relationship between a set of indicators and is used the SPSS Statistic v24 application and reliability using Cronbach's Alpha. The sample in this study was 30 students with a quota sampling approach and was taken randomly. The results of the analysis construct validity with EFA got KMO values > 0.5, Anti Image Correlation > 0.5, Eigenvalue 1 and Factor Loading 0.3 and instrument reliability is 0.771. The results showed that overall, the items of instrument recount text writing test HOTS in English subjects for vocational high school students had good construct validity scores, all items were able to measure every cognitive ability and showed that the construct validity of each test item developed was good because of construct validity. Teachers need to use the development of this standard instrument to reduce the level of the subjectivity of the assessment. Teachers should start using learning models and include HOTS questions to improve students' thinking skills.

Keywords: EFA, Writing Test, HOTS, Vocational High School.

INTRODUCTION

An instrument is a measuring tool used to obtain quantitative and qualitative information about the variation in the characteristics of research variables objectively. The instrument functions to reveal facts into data, so that if the quality of the instrument used is good, then the data obtained is by the facts (Rahmawan, Sumaryanto, and Supriyadi, 2016). In the field of education, the instrument is used to measure student learning achievement, factors that are thought to have a relationship or influence on learning outcomes, the development of student learning outcomes, the success of the teacher's teaching and learning process, and the success of achieving a particular program. Kartoningsih, et al (2018: 475) in their research revealed that the world of education is developing dynamically, especially in creating more interactive and comprehensive media, methods, and materials.

Assessment in the field of education has an important role to improve the quality of the teaching process. Teachers are also responsible for efforts to improve the quality of education. The results of this assessment are used as material in evaluation activities which are then used to make decisions regarding learning outcomes (Tucker et al in Gareis, 2007:20). Quality improvement must begin by paying attention to the current situation obtained through the process of learning outcomes assessment activities (Fuadi, Sumaryanto, and Lestari, 2014: 54).

The implementation of the assessment must be carried out intensively and regularly to foster good study habits for students. Educators in schools generally carry out assessments as they are and the assessments carried out do not pay attention to the process aspect (Arifin, 2009: 39). One of the reasons that the assessment objectives have not been met following the standards that have been set is the assessment process that does not use standardized instruments (Hamid 2010: 28). Permendikbud No.66 of 2013 describes assessment techniques and instruments under the 2013 Curriculum, but the reality is that educators are still experiencing difficulties in implementing assessment techniques and instruments standardized by the government.

The 2013 curriculum has been launched by the government since the 2013/2014 academic year and is implemented gradually at all levels of schools throughout Indonesia. The 2013 curriculum is prepared for generations to face globalization. Suraji, Sumaryanto, and Khumaedi (2019) in their research explained that the government developed the 2013 curriculum by increasing students' potential for competitiveness in the 21st century. Skills in the 21st century consist of three components, namely learning skills, reading skills, and life skills. Four skills are often found in learning skills known as the 4Cs (communication, collaboration, critical thinking, creativity, and innovation).

Vocational education is secondary education that prepares students especially to be ready to work in certain fields. Certain fields are fields that are chosen and studied as long as students are in vocational education institutions. Kartoningsih, et al (2018: 475) in their research revealed that the world of education is developing dynamically, especially in creating more interactive and comprehensive media, methods, and materials. Vocational education is an education subsystem that specifically assists students in preparing themselves to enter the workforce (Law No. 20 of 2003).

Sasongko's research (2010) is in line with Fave's (1966) statement that essay tests can be investigated in depth without being unreasonable because they allow test takers to defend their answer choices by showing properties that encourage giving their answers. Essay tests can be standardized in several ways (based on evidence from experimental psychology) to establish objective scoring criteria. Rusilowati (2017: 19) states that the test is divided into three, namely objective tests, description tests, and performance tests. The form of the essay test provides an opportunity for test-takers to organize ideas and or things that have been learned using their own words.

Widoyoko (2016: 86) mentions several advantages of the essay form test. The advantages of the description test are: (1) it can be used to measure complex learning outcomes, such as the ability to apply principles, the ability to interpret relationships, and the ability to formulate valid conclusions, (2) Increase the motivation of test-takers to learn (3) Easy to prepare and organize, (4) There are not many opportunities for speculation or luck, (5) Encouraging test takers to dare to express their opinions, (6) Providing opportunities for test-takers to express their meaning in their language and way.

The essay test showed effectiveness in improving the achievement of the essay test, there was no significant difference between the treatment groups in students who had a disability (disability) in reading and writing skills (Therrien et al, 2009). Writing skills in essay tests can correlate scores on the level of writing ability in admissions rather than student placement tests (Goodwin, 2016). Day et al (1990) test essays can be constructed within the scope of the clinical assessment domain. The results in the form of essay tests can be linked to other clinical competencies so that they can provide information. Essay tests are arranged through the stages of planning, writing, reviewing, administering tests, scoring, analyzing and interpreting, and revising (Suyata, 2006). Good preparation will be able to produce an essay test that can provide information. Essay tests can show evidence of awareness in the metacognitive understanding of the function of

feedback and formative test assessment procedures (Ellery, 2008).

Preparation of HOTS using Problem Based Learning needs to consider educational goals and teachers must design problems to meet the specified goals (Weiss, 2003). Project-based learning has been proven to help students become collaborators, develop thinking skills, ask questions, share ideas and discuss ideas, find and analyze information from various sources and make multimedia presentations (Susanawati et al, 2014) (Faizah et al, 2015) (Nuswowati et al, 2017). Barnett and Francis (2012) conducted a study to test whether quizzes containing higher-order thinking questions were associated with critical thinking and performance tests when used together. Students to be able to face changing circumstances or challenges in the learning process need to have higher-order thinking skills (HOTS) in the form of critical thinking (Arafah et al, 2012:48). Higher Order Thinking Skills (HOTS) is a very important aspect to be developed in learning (Susanto & Retnawati, 2016).

The HOTS-oriented Problem-Solving learning model has a positive effect on learning outcomes. Problem-solving learning can stimulate higher-order thinking skills such as critical thinking and creative thinking (Handayani and Priatmoko, 2012) (Jailani et al, 2017) (Anazifa and Djukri, 2017). The material is given to students who can provide a stimulus to think creatively to solve the given problem (Handayani et al, 2015). The critical and creative thinking patterns possessed by students can identify, and formulate problems, identify evidence and phenomena, draw conclusions, and communicate conclusions (Noviani et al, 2017). The constructed instrument can distinguish critical thinking skills between students who receive a learning process that trains critical thinking skills (Ritdamaya and Suhandi, 2016). Problem-solving learning uses different ways of using knowledge to solve problems and involves metacognition (Zuhaida et al, 2014: 2).

Higher-Order Thinking Skill is a student's thinking pattern by relying on the ability to analyze, create, and evaluate all aspects and problems. According to Zaini (2015), higher-order thinking is a thinking skill that combines critical thinking and creative thinking. Schraw in Kusuma et al (2017:26-32) classifies thinking skills based on Bloom's taxonomy into two categories, namely Lower Order Thinking Skill (LOTS) which consists of knowledge, understanding, and application. Higher-Order Thinking Skill (HOTS) consists of Analyze, Synthetic, and evaluation. Krathwohl (2002:215) states that Bloom's revised taxonomy consists of (1) Remember, (2) Understand, (3) Apply, (4) Analyze, (5) Evaluate, and (6) Create. Higher-order thinking skills are Analyze (C4), Evaluate (C5), and Create (C6) (Trisnawati et al, 2017). Juhanda (2016) mentions the average percentage of questions that develop higher-order thinking skills starting from questions C4, C5 and C6 have a low average percentage. The question of the revised Bloom's cognitive level that develops HOTS still needs to be improved.

Based on the explanation above, it encourages researchers to choose the topic of analysis of the instrument development of recount text writing test HOTS for vocational high school. With the existence of an instrument whose validity has been developed in a constructive manner using exploratory factor analysis (EFA) and is reliable, it is hoped that in assessing the test it can facilitate and assist the assessors in assessing higher-order thinking skills, especially in the description test on the competence of writing recount material.

METHOD

This study was conducted using the development of the instrument proposed by Djemari Mardapi (2008) which has been modified with the following steps: (1) determining the specifications of the test instruments, (2) writing test instruments, and (3) analyzing test instruments with validity and reliability. The research flow chart can be described as follows:

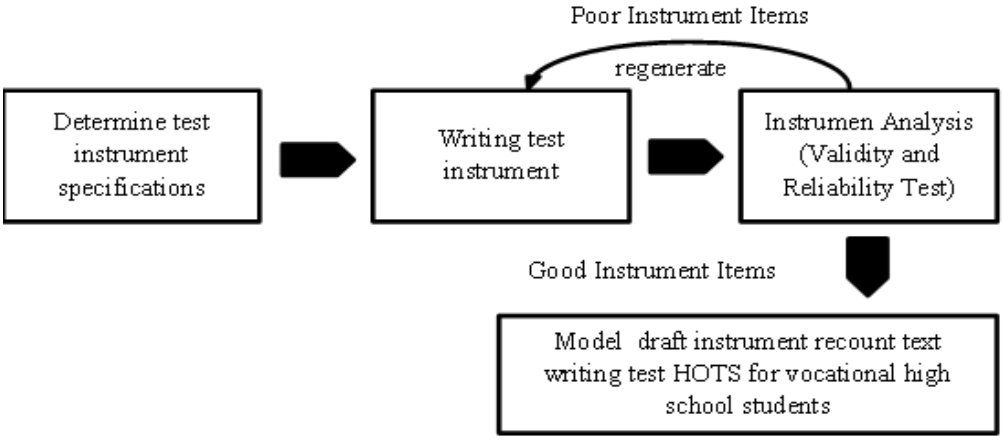

Figure 1. Research Flowchart

The first step in developing the test is to determine the specifications of the test instrument which contains a description that shows the overall characteristics that a test must-have. Researcher (1) determines the purpose of the test, namely the formative test. (2) Compile a grid based on the syllabus provided by the Ministry of Education and Culture in the 2013 curriculum, make a list of subjects to be tested in KI-3 for Class X English, determine indicators based on the dimensions of writing recount text material, and determine the number of questions for each subject. and sub-topics of the 5 item questions. (3) determine the form of the test is a description/free description. (4) determine the length of the test, which is 2 hours of lessons (45 minutes x 2 = 90 minutes).

Researchers develop test instruments in this case are questions and scoring guidelines (Rubric). The researcher arranges the description questions by describing the indicators based on the grid that has been made. Researchers in preparing questions will be guided by the signs that need to be considered in the preparation of the description test delivered by Djemari Mardapi, Saifuddin Azwar, S. Eko Putro Widoyoko, Abdul Majid, and Ani Rusilowati.

The researcher conducted an instrument analysis after obtaining empirical data from the test results of 23 students of class X SMK using a quota sampling approach and was taken randomly. Researchers in this step analyze each item of the question. The researcher tested the construct validity, the reliability of the test instrument. In the construct validity stage, the test instrument was analyzed using factor analysis to see the level of validity. The construct validity testing procedure departs from the computational results of the intercorrelation among various test results and is then followed by further analysis of the correlation matrix obtained, through various methods (Azwar, 2015, p.116). 1 of the 2 approaches used to test construct validity is the factor analysis approach. Construct validity will be analyzed by applying factor analysis. The factor analysis that will be carried out is Exploratory Factor Analysis (EFA) calculated using the help of the IBM SPSS Statistic v24 application program.

The exploratory factor analysis (EFA) procedure helps instrument developers identify and identify the various factors that make up a construct by finding the largest score variance with the least number of factors, expressed in the form of eigenvalue > 1.0. An exploratory approach is used to see how many factors are needed to explain the relationship between a set of indicators. The confirmatory approach is used to test whether the number of factors obtained empirically is following the number of factors that have been compiled theoretically. The construct validity test in this study used an exploratory analysis approach, and factor analysis procedures with the help of the IBM SPSS Statistic v24 application program. The construct validity test using factor analysis can be run if the KMO Value > 0.5, Anti Image Correlation > 0.5, Eigenvalue 1, and Factor Loading 0.3 (Nugroho et al., 2016, p.3).

The reliability test was carried out using the IBM SPSS Statistic v24 application by entering data and then clicking the Analyze, Scale, Reliability Analysis menu by selecting the Alpha model. The alpha value is based on the results of Cronbach's Alpha Reliability Statistics output. The reliability coefficient of 0.50 is sufficient to be accepted as good

reliability (Khumaedi, 2012, p.29). Meanwhile, according to Rusilowati (2014, p.29), the acceptance reliability criteria for the Cronbach Alpha coefficient are in Table 1.

Table 1. Acceptance of Cronbach's Alpha Coefficient

Alpha Cronbach's	Internal Consistency
$0 \leq \alpha < 0.2$	Very Low
$0.2 \leq \alpha < 0.4$	Low
$0.4 \leq \alpha < 0.6$	Average
$0.6 \leq \alpha < 0.8$	High
$0.8 \leq \alpha < 1.0$	Very Hihg

(Rusilowati, 2014 p.29)

RESULTS AND DISCUSSION

Result

The results of this study show and discuss two important points, namely; 1) the form of a writing test in the English subject of writing competence in SMK, 2) the validity and reliability of the HOTS writing test instrument in the English subject of writing competence in SMK which was developed based on the construct of the instrument. The explanations of these two points are described in the following:

The test on the writing competence of HOTS-oriented recount text material is a test carried out with formative test activities for students in writing a recount text with high-level ability (HOTS) questions specifications. Instrument development is one of the strategic steps that will determine the quality of research results, because the validity or validity of the data obtained will be largely determined by the quality or validity of the instruments used, in addition to the procedures for collecting the data obtained. Furthermore, the development of the instrument will affect the shape of the previous instrument.

The specifications of the instruments compiled are reviewed from 3 factors, namely external contextual factors, internal attributes, and future test specifications (Rusilowati, 2017: 54). External contextual factors are related to the characteristics of test-takers, time constraints, test administration, and test standardization. Question indicators are adjusted to the abilities of junior high school students. The test maker with a time limit of 90 minutes or 2 hours of lessons with a total of 20 questions. The test administration is structured practically. The teacher simply uses the grading sheet to get the final score. Standardization of the test is done by completing the instrument with instructions for test-takers, instructions for assessors, test material, time used, assessment procedures have been prepared so that they can measure well.

The internal attributes of the test instrument that are compiled taking into account the suitability of the test with the material. The test material is adjusted to the curriculum and arranged in the form of a grid. The essay test was chosen as a form of test following what the teacher used to measure learning outcomes for Basic Competencies. Assessors or teachers can easily understand the assessment guidelines. The assessment categories are clear and are divided into 3 score categories, each of which has a different score range.

The construct validity of the instrument was carried out to determine the factors formed from several dimensions that could be revealed, obtained through the EFA (Exploratory Factor Analysis) approach using the help of the IBM SPSS Statistic v24 application program. essay writing test items to measure high thinking ability (HOTS). Through factor analysis, the researcher identifies a structure and then determines how much of the variable can be explained by each dimension. After the dimensions and explanations of each variable are known, the main objective of factor analysis can be carried out, namely summary of data and data reduction Mustaqima suggests several important values that need to be considered in interpreting the output of exploratory factor analysis, namely the KMO and Bartlett's Test values, Anti Image Matrix, Total Variance Explained, Screen Plot, Component Matrix, and d Rotated Component Matrix (Mustaqima et al., 2018)

Conceptually, the items for the essay writing recount text test to measure high thinking skills (HOTS) in recount text materials

were developed by constructs of 3 indicators, namely Analyzing, Evaluating, and Creating. One HOTS indicator consists of 1 to 2 items, so 3 indicators consist of 6 items of developed thinking skills. The syntax used to input data in the IBM SPSS Statistic v24 application program from 6 questions is described in Table 2.

Three indicators are explained in syntax code 1 and syntax 2 AN = Analyzing consisting of item number 1, EV = Evaluating item number 2 and 6, and CR = Creating item numbers 3,4, and 5.

Table 2. SPSS Syntax Code

Items Num	Syntax **1**	Syntax **2**
1	AN_1	But1_Analyzing
2	EV_2	But2_Evaluating
3	CR_3	But3_Creating
4	CR_4	But4_Creating
5	CR_5	But5_Creating
6	EV_6	But6_Evaluating

The KMO value varies from 0 to 1. If the KMO value is > 0.5 then factor analysis can be carried out, but if the KMO value is < 0.5 then the factor analysis cannot be continued (Ghozali, 2016, p378).

Table 3. KMO and Bartlett's Test Scores

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.628
Bartlett's Test of Sphericity	Approx. Chi-Square	45.988
	df	15
	Sig.	.000

Based on the output of IBM SPSS Statistic v24 in table 3. the value of KMO and Bartlett's Test obtained a value of 0.628. This explains that all of the instrument items meet the criteria because the KMO and Bartlett's Test scores are 0.5. Therefore, it can be concluded that the results of the analysis of instrument development of recount text writing test HOTS have met the criteria and requirements for further analysis using factor analysis using the help of the IBM SPSS Statistic v24 application program. the next step is to analyze the correlation between instrument items.

Anti-Image Correlation Matrix

In addition to the KMO and Bartlett's Test values, factor analysis was continued because it was strengthened by the correlation value of each item of more than 0.5. Table 4. shows the correlation value of each item in question.

Table 4. The Correlation Value of Each Item on the Instrument Development of Recount Text Writing Test HOTS

		But1_ Anal	But2_ Eval	But3_ Creat	But4_ Creat	But5_ Creat	But6_ Eval
Anti-image Correlation	But1_Analyzing	**.505[a]**	-.715	-.101	.033	-.043	.376
	But2_Evaluating	-.715	**.626[a]**	-.061	-.296	.001	-.426
	But3_Creating	-.101	-.061	**.567[a]**	.293	-.007	-.453
	But4_Creating	.033	-.296	.293	**.710[a]**	-.375	-.142
	But5_Creating	-.043	.001	-.007	-.375	**.786[a]**	-.253
	But6_Evaluating	.376	-.426	-.453	-.142	-.253	**.604[a]**

a. Measures of Sampling Adequacy(MSA)

Based on Table 4. in the anti-image column, it can be seen that the correlation value between items has a correlation value > from 0.5, therefore factor analysis can be continued by including all items.

Total Variance Explained

Total variance explained shows the percentage of total variance that can be explained by the diversity of the formed factors. To see how many factors are formed in factor analysis, it can be seen in Table 4.7. Results of Analysis of Total Variables Explained by Factor Analysis.

Table 5. Results of Analysis of Total Variance Explained Factor Analysis

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	2.865	47.749	47.749
2	1.136	18.938	66.687
3	1.020	16.995	83.683
4	.465	7.743	91.426
5	.343	5.711	97.137
6	.172	2.863	100.000

Extraction Method: Principal Component Analysis.

Table 5. shows that 3 components are formed and can represent the number of indicators. The 6 items analyzed turned out to have eigenvalues above 1 there were 3 components, meaning that the 6 indicators could be grouped into 3 groups of factors. Component 1 has a value of 2.865 and can explain the variance of 47.749 then Component 2 has a value of 1.136 and can explain the variance of 18.938 and Component 3 has a value of 1.020 and can explain the variance of 16.995 thus the three components can explain the variance of 83.683. To determine how many components/factors are used to explain the total diversity, judging from the large eigenvalues, components with eigenvalues > 1 are the components used. Overall, the results of the total variance can be seen in Figure 2.

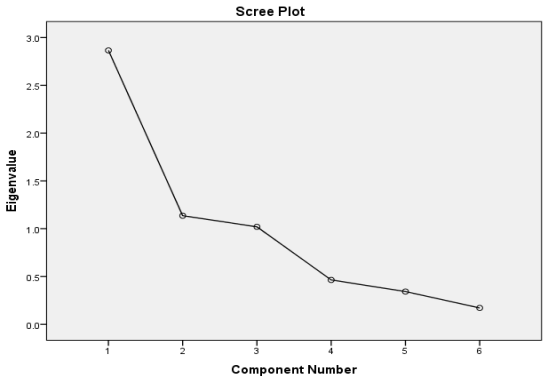

Figure 2. Scree Plot Factor Analysis

Figure 2. The scree plot shows that there are 3 points or 3 coordinates that are above the value 1 and other points are below the value 1. This illustrates that 3 components have an eigenvalue above the value 1, meaning that there are only 3 factors formed and it is best to summarize the six items of the instrument of recount text writing test HOTS.

Rotated Component Matrix

A rotated component matrix is an output that describes the grouping of items in the instrument recount text writing test HOTS into several factors. In this study, the nine instrument items after being extracted became 3 components. After the rotation of the factors using the varimax method, we get table 4.8 Rotated Component Matrix. There is a difference in the value of the correlation of variables with each factor before and after the varimax rotation. It can be seen that the rotating loading factor has given meaning as expected and each factor can be interpreted clearly. The results of the rotated component matrix analysis are presented in Table 6.

Tabel 6. Result of *Rotated Component Matrix*[a]

	Component		
	1	2	3
But1_Analyzing	.070	.959	.035
But2_Evaluating	.417	.781	.296
But3_Creating	-.065	.144	.928
But4_Creating	.837	.294	-.085
But5_Creating	.831	.080	.179
But6_Evaluating	.559	.075	.703

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

Based on Table 6 showing the results of the rotation factor, it can be seen that the grouping of instrument items into factors and the magnitude of the loading factor obtained is shown in Table 4.8, it can be seen that the determination of the input of instrument items to certain factors follows the magnitude of the correlation between variables and factors, namely those with large correlations. Thus, the formed factors and their items are shown in Table 7.

Table 7. Grouping Instrument Items into Factors

Item Num	Instrument Items	Factor	Factor Correlation Value	Variable	Construct Validity Coefficient
4	Creating_2	1	**.837**	Item 4	.834
5	Creating_3		**.831**	Item 5	
1	Analyzing	2	**.959**	Item 1	.870
2	Evaluating_1		**.781**	Item 2	
3	Creating_1	3	**.928**	Item 3	.815
6	Evaluating_2		**.703**	Item 6	

Based on Table 7, it can be seen from the rotation process that the distribution of the six instrument items is clearer and the formation is real. For the components formed in factor 1, there are 2 variables named Item 4 and Item 5, which are declared as constructs with a construct validity value of 0.834, instrument items consist of items 4 and 5 with the following item sentence editorial (4) Creating_2, (5) Creating_3. While factor 2 consists of items 1 and 2 are named Item 1 and Item 2 and are declared a construct with a construct validity value of 0.870 with the editorial points (1) Analyzing and (2) Evaluating_1, then factor 3 is named Item 3 and Item 6 and declared a construct with a construct validity value of 0.815 consisting of items 3 and 6 with the sentence editor points (3) Creating_1 and (6) Evaluating_2.

The instrument recount text writing test HOTS in English subjects for vocational high school students was tested for reliability using the help of the IBM SPSS Statistic v24 application program. Based on the analysis, the results are obtained in Table 8.

Table 8. Reliability Analysis Data

Test	**Alpha Cronbach's**	**N of items**
Test Instrument	0.771	6

Table 8. shows the reliability coefficient of the expert agreement test of 0.771 so it can be interpreted that the instrument is reliable.

Discussion

The results obtained in the preparation of the background and formulation of the problem became the basis for the instrument recount text writing test HOTS in English subjects for vocational high school students. The researcher compiled the specifications for the recount text writing test HOTS as follows: The test is in the form of essay test questions, working time is 90 minutes, measuring Core Competence 3 and Basic Competence 3.7. which consists of 3 indicators, vocational high school English subject for class X. The instruments developed are in the form of a grid, 6 questions, scoring guidelines, and an assessment sheet.

Researchers write or develop instruments based on indicators that have been set by the government and developed on grids, questions, scoring guidelines, and assessment sheets. The researcher arranges 1-2 items for each indicator and there is 1 item for the analysis level, 2 items for the evaluation level, and 3 items for creation. The scoring guidelines were developed into 3 categories of answers for each item, with details of the Orientation (opening) scoring category having a score range of 0-2 and the Sequences of Events scoring category having a score range of 1-4. The scoring guide contains an answer key as a guide for categorizing student answers. The design of the instrument recount text writing test HOTS in English subjects for vocational high school

students was continued by testing the validity of the instrument.

The instrument recount text writing test HOTS in English subjects for vocational high school students was analyzed using EFA to determine the construct validity of the instrument compiled. This research is in line with the research of Pratiwiningtyas, Susilaningsih, & Alreadya (2017) that the sociology learning assessment module that is oriented to higher-order thinking skills must meet the feasibility aspect from an expert's point of view. The study covers aspects of material suitability, construction suitability, and language suitability. The expert test instrument consists of some prerequisite items that must be assessed and analyzed. Experts and practitioners also provide suggestions and input on the design of the developed instrument.

The instrument recount text writing test HOTS in English subjects for vocational high school students was declared constructively valid. The construct validity used Exploratory Factor Analysis (EFA) with the aid of the SPSS Program. The next step on instrument development recount text writing test HOTS for vocational high school students is to test the writing test instrument in the field with a sample of 30 students at SMK Ma'arif NU 1 Semarang in class X. The results of field trials show that the KMO score (Kaiser Meyer Olkin Measure of Sampling) on the 6 items of the instrument recount text writing test HOTS for vocational high school student that was developed has a KMO value > 0.05, which is 0.628, which means that all of these items are suitable for factor analysis by the theory put forward, namely if the KMO value is > 0.5 then factor analysis can be carried out, but if the KMO value is <0.5 then the factor analysis cannot be continued (Ghozali, 2016, p. 378).

The analysis of the instrument test items in the field by looking at the results of the rotated component matrix as a whole, the items of instrument recount text writing test HOTS in English subjects for vocational high school students have a loading load of > 0.5. In the instrument recount text writing test HOTS in English subjects for vocational high school students, 3 components are formed and can represent the number of indicators. The 6 items analyzed turned out to have eigenvalues above 1 there were 3 components, meaning that the 6 indicators could be grouped into 3 groups of factors. Component 1 has a value of 2.865 and can explain the variance of 47.749 then Component 2 has a value of 1.136 and can explain the variance of 18.938 and Component 3 has a value of 1.020 and can explain the variance of 16.995 thus the three components can explain the variance of 83,683. To determine how many components/factors are used to explain the total diversity, judging from the large eigenvalues, components with eigenvalues > 1 are the components used.

Similar research on the use of EFA to look for formed factors, the study resulted in a rotated component matrix value greater than 0.5 indicating that the instrument formed certain factors, such as the component of the teacher's task in teaching, a teacher in educating, a teacher in training and directing, teachers in guiding and teachers in assessing and evaluating. The results of the factor analysis show a fairly good fit model, meaning that it can be concluded that (1) the primary school teacher performance assessment instrument consists of five components, namely teaching, educating, training and directing, guiding, and assessing, and evaluating, (2) each component teacher performance assessment instruments can be translated into several relevant indicators as described in the research results (Sadtyadi & Kartowagiran, 2014, pp.297-300)

This shows that overall, the items of instrument recount text writing test HOTS in English subjects for vocational high school students have a good construct validity value and can measure every cognitive ability and show that the construct validity of each instrument development recount text writing test HOTS for vocational high school students is good, because construct validity is reflected in the extent to which the factor content of the analysis results is by the underlying theory.

The reliability results of instrument recount text writing test HOTS in English subjects for vocational high school students were compiled based on Table 4.15 and analyzed using the Cronbach Alpha reliability test to obtain a fairly good reliability coefficient. The instruments compiled must have a level of reliability and consistency of measurement (Azwar, 2016:111). In the small-scale test instrument, the reliability coefficient is 0.771.

CONCLUSIONS AND SUGGESTIONS

Based on the results of the analysis above, it can be concluded that the form of the

instrument produced in this study is an instrument development recount text writing test HOTS for vocational high school students at Basic Competence (KD) 3.7. Instruments are development by guidelines and signs from experts to produce standardized instruments. The instruments that are arranged consist of a grid, questions, scoring guidelines, and scoring sheets.

The resulting instrument has gone through testing starting from the stage of determining the specifications of the test instrument then proceeding to the stage of writing the test instrument and analyzing the validity and reliability of the instrument. The results of testing instrument development recount text writing test HOTS for vocational high school students were valid and reliable can provide an assessment that is appropriate to the ability of students in the cognitive domain (Test) in learning English in the writing aspect of recount text material.

The results of the Rotated Component Matrix factor analysis on the test items on instrument recount text writing test HOTS for vocational high school students were formed by 3 factors. Based on the reliability analysis of the field test instrument items analyzed using Cronbach's Alpha, the results obtained are quite high reliability in the implementation of instrument measurements in the field so that the results of the development of the resulting assessment instrument are reliable.

Based on the results of this study, the researcher suggests that the English teacher in conducting the assessment of Basic Competence 3.7 must have an assessment instrument, especially a standardized essay writing test instrument to measure higher-order thinking skills (HOTS). Teachers must also use standardized assessment instruments or instruments that have passed validity and reliability tests in conducting assessments. The scores obtained with standard instruments (valid and reliable) will provide useful information for English subject teachers to take appropriate steps in the learning process seen from the scores obtained by students in the cognitive domain test on Basic competence 3.7. Teachers must understand the thinking ability of students to know the right learning model in planning learning. Teachers need to improve students' thinking skills which were previously only low-level thinking skills (LOTS) to be upgraded to higher-order thinking skills (HOTS).

ACKNOWLEDGMENTS

The authors would like to thank Graduate School Universitas Negeri Semarang for the funding and all resources provided to conduct this research. Funding statement: The research is funded under Research Project DIPA Universitas Negeri Semarang Number: SP DIPA-023.17.2.677507/2021, November 23, 2020 in accordance with the Letter of Assignment for the Implementation of the UNNES DIPA Research Fund Research in 2021. Number: 68.19.5/UN37/PPK.5.1/2021, May 19, 2021.

Sincere gratitude also goes to anonymous reviewers and editors who have provided constructive feedback so that this manuscript looks worth reading and citing.

REFERENCES

Anazifa, R. D., & Djukri, D. (2017). Project-Based Learning and Problem-Based Learning: Are They Effective to Improve Student's Thinking Skills? *Jurnal Pendidikan IPA Indonesia*, *6*(2), 346. https://doi.org/10.15294/jpii.v6i2.11100

Arafah, S. F., Ridlo, S., & Priyono, B. (2012). Pengembangan LKS Berpikir Kritis Pada mteri Animalia. *Unnes Journal of Biology Education*, *1*(1), 47–53.

Arifin, Z. (2009). *Evaluasi Instruktusional*. Bandung: Remaja Karya.

Azwar, S. (2015). *Penyusunan Skala Psikologi Edisi 2*. Yogyakarta: Pustaka Belajar.

Azwar, S. (2016). *Penyusunan Skala Psikologi Edisi 3*. Yogyakarta: Pustaka Pelajar.

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. https://doi.org/10.1080/01443410.2011.638619

Darmadi, Hamid. (2010). *Kemampuan Dasar Mengajar*. Bandung: Alfabeta

Day, S. C., Norcini, J. J., Diserens, D., Cebul, R. D., Schwartz, J. S., Beck, L. H., … Elstein, A. (1990). The Validity of an Essay Test of Clinical Judgment. *Academic Medicine*, *65*(9), S39-40.

Ellery, K. (2008). Assessment for Learning: A Case Study Using Feedback Effectively in an Essay-style Test. *Assessment & Evaluation in Higher Education*, *33*(4), 421–429. https://doi.org/10.1080/02602930701562981

Faizah, U., Prastiwi, M. S., Subekti, N., Setyowati, D. L., Rachmadiarti, F., & Kuncjoro, S. (2015). Teaching Materials Model-based Problem Based Learning (PBL) to Habituate Students' Conservation Awareness. *In Procceding ICCBL 2015* (pp. 1–3). Semarang: International Conference on Conservation for Better Life.

Fuadi , Sumaryanto, T., & Lestari, W. (2014). Pengembangan Instrumen Penilaian Aspek Psikomotor Pembelajaran IPA Materi Tumbuhan Hijau Kelas V Berbasis Kompetensi Pendekatan Sea Berwawasan Konservasi. *Innovative Journal of Curriculum and Educational Technology*, *3*(2). Retrieved from https://journal.unnes.ac.id/sju/index.php/ujet/article/view/9827

Gareis, R.C. (2007).”Reclaiming an Important Teacher Competency: The Lost Art of Formative Assessement”. *J Pers Eval Educ*. No. 20. Hal. 17–20

Ghozali, Imam. 2016. *Aplikasi Analisis Multivariete Dengan Program IBM SPSS 23 (Edisi 8). Cetakan ke VIII*. Semarang: Badan Penerbit Universitas Diponegoro.

Goodwin, S. (2016). A Many-Facet Rasch Analysis Comparing Essay Rater Behavior on an Academic English Reading/Writing Test used For Two Purposes. *Assessing Writing*, *30*, 21–31. https://doi.org/10.1016/j.asw.2016.07.004

Handayani, R., & Priatmoko, S. (2012). Pengaruh Pembelajaran Problem Solving Berorientasi HOTS (Higher Order Thinking Skills) Terhadap Hasil Belajar Kimia Siswa Kelas X. *Jurnal Inovasi Pendidikan Kimia UNNES*, *6*(2), 954–965. https://journal.unnes.ac.id/nju/index.php/JIPK/article/view/6532

Jailani, J., Sugiman, S., & Apino, E. (2017). Implementing the Problem-based Learning in Order to Improve the Students' HOTS and Characters. *Jurnal Riset Pendidikan Matematika*, *4*(2), 247. https://doi.org/10.21831/jrpm.v4i2.17674

Juhanda. (2016). Analisis Soal Jenjang Kognitif Taksonomi Bloom Revisi Pada Buku Sekolah Elektronik (BSE) Biologi SMA. *Jurnal Pengajaran Matematika Dan Ilmu Pengetahuan Alam*, *21*(1).

Katoningsih, S., Soegito, A. T., Sumaryanto, T., & Susanto, M. (2018, September). The Implementation of Mobile Learning for Reading Literacy (Morlisa) for English Teacher training in Semarang. In *International Conference on Science and Education and Technology 2018 (ISET 2018)*. Atlantis Press. https://doi.org/10.2991/iset-18.2018.95

Khumaedi, M (2012). Reliabilitas Instrumen Penilian Afektif Mata Pelajaran Aqidah Akhlak. *Journal of Educaional Research and Evaluation, 3* (1), 37-40.

Krathwohl, D. R., & Anderson, L. W. (2010). Merlin C. Wittrock and the Revision of Bloom's Taxonomy. *Educational Psychologist*, *45*(1), 64–65. https://doi.org/10.1080/00461520903433562

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

La Fave, L. (1966). Essay vs Multiple-choice: Which Test is Preferable? *Psychology in the Schools*, *3*(1), 65–69. https://doi.org/10.1002/1520-

6807(196601)3:1<65::AID-PITS2310030117>3.0.CO;2-Y

Mardapi, D. (2008). *Teknik Penyusunan Instrumen Tes dan Nontes*. Yogyakarta: Mitra Cendekia.

Mustaqima, Z. Z. (2018). *Faktor-Faktor yang Mempengaruhi Pemilihan Moda Angkutan Umum Berbasis Online di Jakarta Selatan* (Doctoral dissertation, Universitas Brawijaya).

Noviani, Y., Hartono, & Rusilowati, A. (2017). Analisis Pola Pikir Siswa Dalam Menyelesaikan Soal Sains Ditinjau dari Kemampuan Berpikir Kritis dan Kreatif Serta Literasi Sains. *Journal of Innovative Science Education*, *6*(2), 147–154.

Pratiwiningtyas, B. N., Susilaningsih, E., & Sudaha, I. M. (2017). Pengembangan Instrumen Penilaian Kognitif untuk Mengukur Literasi Membaca Bahasa Indonesia Berbasis Model Pirls pada Siswa Kelas IV SD. *Journal of Educational Research and Evaluation*, *6*(1), 1–9. https://doi.org/10.15294/jrer.v6i1.16199

Rahmawan, E. F., Sumaryanto, T., & Supriyadi, S. (2016). Pengembangan Instrumen Penilaian Kinerja Kemampuan Bernyanyi Berbasis Android. *Journal of Educational Research and Evaluation*, 5(1), 81-89. https://doi.org/10.15294/jrer.v5i1.14888

Ritdamaya, D., & Suhandi, A. (2016). Konstruksi Instrumen Tes Keterampilan Berpikir Kritis Terkait Materi Suhu dan Kalor. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, *2*(2), 87–96. https://doi.org/10.21009/1.02212

Rusilowati, A. (2014). *Pengembangan Instrumen Penilaian*. Semarang: UNNES PRESS

Rusilowati, A. (2017). *Pengembangan Instrumen Penilaian*. Semarang: Unnes Press.

Sasongko, P. (2010). Comparison of The Effectiviness of The Essay Test And Testlets Through The Graded Response Model (GRM) Application. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *14*(2) 1-9.

Suraji, Sumaryanto, T., & Khumaedi, M. (2019). The Analysis of Instrument of The Ability to Acting and Thinking Creatively Based Rasch Model, *Journal of Educational Research and Evaluation*, *8*(10), 48–56. http://journal.unnes.ac.id/sju/index.php/jere

Sadtyadi, H., & Kartowagiran, B. (2014). Pengembangan Instrumen Penilaian Kinerja Guru Sekolah Dasar Berbasis Tugas Pokok dan Fungsi. *Jurnal Penelitian dan Evaluasi Pendidikan*, 18(2), 290-304. https://doi.org/10.21831/pep.v18i2.2867

Susanawati, E., Diantoro, M., & Yulianti, L. (2014). Pengaruh Strategi Projectbased Learning Dengan Thinkquest Terhadap Kemampuan Berpikir Kritis Fisika Siswa SMA Negeri 1 Kraksaan. *Jurnal Pengajaran Matematika Dan Ilmu Pengetahuan Alam*, *18*(2), 207. https://doi.org/10.18269/jpmipa.v18i2.51

Susanto, E., & Retnawati, H. (2016). Perangkat pembelajaran matematika bercirikan PBL untuk mengembangkan HOTS siswa SMA. *Jurnal Riset Pendidikan Matematika*, *3*(2), 189. https://doi.org/10.21831/jrpm.v3i2.10631

Suyata, P. (2006). Tes Bahasa Bentuk Uraian: Upaya ke Arah Kualitas Soal. *Cakrawala Pendidikan: Jurnal Ilmiah Pendidikan*, *2*(2). https://doi.org/10.21831/cp.v2i2.9259

Therrien, W. J., Hughes, C., Kapelski, C., & Mokhtari, K. (2009). Effectiveness of a Test-Taking Strategy on Achievement in Essay Tests for Students With Learning Disabilities. *Journal of Learning Disabilities*, 42(1), 14–23.

Trisnawati, W., Citrasukmawati, A., & Thohir, M. A. (2017). Self Assessment For Student Performance Based on Higher Order Thinking Skills in Physics Learning. *Journal of Education and Learning*, *11*(4), 446–452.

https://doi.org/10.11591/edulearn.v11i4.6456

Widoyoko, E. P. (2016). *Teknik Penyusunan Instrumen Penelitian*. Yogyakarta: Pustaka Pelajar.

Weiss, R. E. (2003). Designing Problems to Promote Higher-order Thinking. *New Directions for Teaching and Learning*, *2003*(95), 25–31. https://doi.org/10.1002/tl.109

Zaini, M. (2015). Hasil belajar dan keterampilan berpikir tingkat tinggi siswa SMA pada pembelajaran biologi menggunakan model pembelajaran berdasarkan masalah. *Jurnal Pendidikan Biologi*, 20(207), 34-43.

Zuhaida, A., Haryani, S., & Susilaningsih, E. (2014). Penyusunan Asesmen Metakognisi Calon Guru Kimia Melalui Pembelajaran Berbasis Masalah. *Jurnal Pendidikan Sains*, *2*(1).

Peraturan Menteri Pendidikan dan Budaya Nasional Republik Indonesia Nomor 66 Tahun 2013 Tentang Standar Proses. *bsnp-indonesia.org/id/wp.../2013/12/Permen-Nomor-66-Tahun-2013.pdf*. (diunduh 19 November 2018)

Undang-Undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional. Jakarta: Departemen Pendidikan Nasional.