JOURNAL OF ENGLISH TEACHING, APPLIED LINGUISTICS AND LITERATURES (JETALL)

English Language Education Study Program, FKIP Universitas Lambung Mangkurat Banjarmasin
Volume 5 Number 2 2022

HIGHER-ORDER THINKING SKILLS ON TEACHERS-MADE TESTS BY ENGLISH TEACHERS OF A SENIOR HIGH SCHOOL IN BENGKULU CITY

Serla Apriana
English Education Study Program of Universitas Bengkulu
serlaapriana@gmail.com

Azhar Aziz Lubis
English Education Study Program of Universitas Bengkulu
azharlubis@unib.ac.id

Dedi Sofyan
English Education Study Program of Universitas Bengkulu
dedi.sofyan73@unib.ac.id

The objectives of this study are to investigate the composition of Higher Order Thinking Skills (HOTS) on the English teacher-made test based on the revised Bloom's taxonomy and investigating the quality of the teacher-made tests viewed from the composition of HOTS. This study was a content analysis that used a triangulation design: data transformation model mixed method approach. The data of this study were teacher-made test sets that consisted of 136 questions. The data were analyzed by understanding the operational verbs used in each item of the teacher-made test based on the cognitive domain of revised Bloom's taxonomy. Then, data were put in a checklist table based in each cognitive domain category. After that, the data were calculated quantitatively and interpreted qualitatively. The results show that the composition of HOTS was lower than the Lower Order Thinking Skills (LOTS) of all questions on the teacher-made tests. The most frequent cognitive domain of the Bloom's taxonomy found on the teacher-made test was C2 (Understand), followed by C4 (Analyze), C3 (Apply), C1 (Remember), C5 (Evaluate), and C6 (Create). Moreover, the quality of the teacher-made tests viewed from the composition of HOTS was in a less proportional criterion. In conclusion, the findings indicate constructing a test based on the principles of HOTS seem still problematic for Senior High School English teachers.

Keywords: *HOTS, teacher-made test, English teacher, revised-Bloom's Taxonomy*

INTRODUCTION

Higher-Order Thinking Skills (henceforth HOTS) are notions of the educational reform based on Bloom's Taxonomy. In revised Bloom's taxonomy by Anderson and Krathwohl (2001), there are six levels of cognitive domain; C1 (Remember), C2 (Understand), and C3 (Apply) called Lower-Order Thinking Skills (henceforth LOTS). While C4 (Analyze), C5 (Evaluate), and C6 (Create) are called HOTS. HOTS are important to apply in learning, particularly in making good questions for a test that contains HOTS. Making questions test is one of the teacher's abilities to evaluate students, and the ability to evaluate students is called pedagogic competence. According to RI Government Regulation No. 19 of 2005 article 28 paragraph 3 concerning National Education Standards states that teachers are expected to have four competencies, including pedagogical competence, personality competence, social competence, and professional competence.

The benefits of HOTS are significant for students and have to be trained and managed well by teachers. In fact, the implementation of HOTS in Indonesia to the learning progress still has some weaknesses. First, the teacher-center is still dominant in the learning process. Teachers as the conveyor of knowledge sometimes eliminate the students' center practice. Second, the educational focus is memorizing. Third, the classic problem, more student achievement scoring systems are based on tests that are tested for low-level cognitive trends. Fourth, teachers still lack HOTS-based question knowledge. Therefore, high order questions need to be applied to students' test items. These are including the ability to conclude, hypothesize, analyze, apply, synthesize, evaluate, compare, imagine, and answer the questions. These skills need high logical thinking.

In addition, teachers should apply higher-order thinking skills in the learning process especially in designing a test. Nevertheless, the teacher is required to design an item test based on the higher-order thinking skill of Bloom's revised taxonomy. The test designed or made by the teacher itself is called a teacher-made test. Arikunto (2005) stated that a "teacher made-test is a test written and made by a teacher in the school, so the validity and reliability of the test are not like the standardized test." The effectiveness of this type of test depends on the skill and ability of the teacher in designing the test. The teacher-made test can enhance the students' thinking skills. Furthermore, the students will be guided to think critically about complexity and various problems.

The development of the HOTS test is not convenient. It requires high logical power. However, this is also a challenge for the teacher, the teacher should be able to get and test students by making a test that contains HOTS compositions. It means that the teacher should be able to develop the test in the form of problem-solving so that students are expected to interpret and study the question presented. And also, the most important thing in higher-order thinking skills is the ability to make a reasonable decision in a complex situation. It emphasizes "learning how" rather than "learning what". Therefore, the efforts to help individuals to achieve their targets require self-awareness which is a part of the efforts of the educators to explore higher-order thinking skills (Sulaiman et al., 2017).

Based on the statements above, students need to have HOTS as a solution to increase their thinking ability to solve the problem faced. Teachers must be able to teach HOTS. More importantly, a teacher must also be able to measure whether or not students have already acquired HOTS. It can be faced from the test used by the teacher to measure student learning outcomes. To construct a good HOTS test item for students, the quality of teachers is an important part of this case. Therefore, in developing the test, they must have a clear knowledge or skills about making good HOTS test for students.

Unfortunately, looking at reality now, many educational institutions only emphasize memorizing skills (LOTS). Besides, higher-order thinking skill is very important to be applied in the aspects of knowledge. It is important to be embedded in students, considering that the rapid development of technology requires every individual to mobilize his or her mind and all the potential to be able to survive and compete. Educational institutions, which only emphasize memorizing make students uncustomed to thinking critically in accepting material given. As the result, the habit of students who only memorize without developing arguments will continue in college and even in the real world of work.

HOTS-based learning has been introduced in line with the implementation of the 2013 Curriculum, but in reality, there are still many educators who do not understand and have not

implemented it. This is because, during the training, the instructors transferred theories solely without giving examples or practices of implementation of HOTS learning (Warmadewi et al., 2019).

There were some previous studies related to this present research such as teacher-made test (Amaliyah, 2018; Pratiwi et al., 2019; Utami et al., 2019; Putri et al., 2020; Syahdanis et al., 2021), national examination test (Ilham et al., 2020), reading exercises on English textbooks ('Ilma, 2018) and language skill tasks on English textbook (Febriyani et al., 2020). These studies reported that the number of the questions that contained HOTS was lower than LOTS. These findings indicate the implementation of HOTS in the English teaching and learning process such as on test and exercises in Indonesia is still low. This case is caused by teachers' factor (teachers' knowledge of HOTS) and students' factor (students' knowledge of HOTS and their score achievement will be low if the questions have many more HOTS questions).

Different from this previous study, this study analysed the HOTS proportion and the previous research only took one grade for their study while this study took all grades. Also, this study took a different place that might be the result of the HOTS test that has been applied by the teacher also will be different, which is in previous research the results were 65% test that has made by the teacher applied LOTS. Therefore, it concludes that the implementation of HOTS in the test to stimulate students' critical thinking is still far from what is expected. In addition, this study used the proportion of HOTS as the standard of the HOTS test, whether the test made by the teachers have been achieved from the proportion that should be or not.

RESEARCH METHODOLOGY

This study used mixed-method approach. The type of mixed method used in this study was Triangulation Design: Data Transformation Model by following Creswell and Clark (2007, p. 63) which made an attempt to integrate the quantitative data into qualitative analysis as shown in Figure 9.

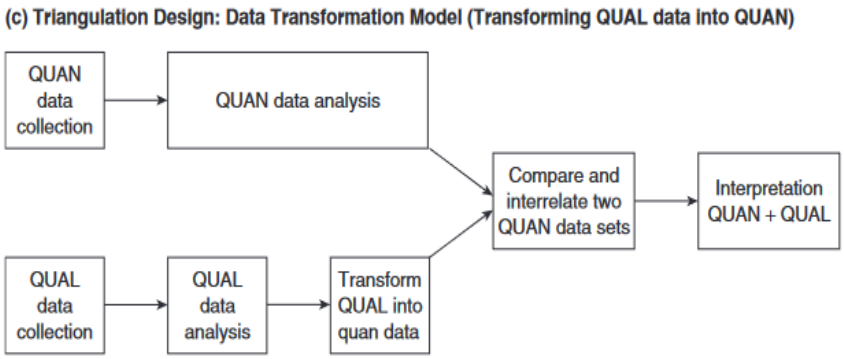

Figure 1: A Mixed Method ("QUAN": quantitative; "QUAL": qualitative)

In accordance with the principle aim of this current research, that was, to specify the composition of HOTS and the quality of teacher-made test viewed from the composition of HOTS, five teacher-made test sets were constructed by five English teachers as the primary research materials. Then, those data were counted based number of question so that the data

consisted of 136 questions. Thus, this process was categorized as transforming qualitative data into quantitative.

The data analysis of this research was twofold. On the one hand, the quantitative analysis focuses on the investigation of HOTS's composition on the teacher-made test based on the revised Bloom's taxonomy that consisted of six cognitive domains. A simple statistical analysis was used to find out the frequency and percentage of each cognitive domain of the revised Bloom's taxonomy on the teacher-made tests. It was for answer the research question number 1. On the other hand, the qualitative analysis consisted of the assessment of the teacher-made tests. It was done by comparing the composition of LOTS and HOTS questions on the teacher-made test. The proportion of LOTS and HOTS was calculated. Then, the quality of the teacher-made tests was interpreted qualitatively based on the Riduwan (2013) as presented in the Data Analysis Technique. Moreover, in presenting the research findings, the numerical (quantitative) and interpretation analysis (qualitative) were done in this research. Also, the results were supported with examples. The examples were explained detailly with comprehensive reasons and analysis.

In conclusion, this content analysis design can be both quantitative and qualitative or mixed methods. This design would give advantages to the quality of the results of this research. It is in line with Sándorová's (2014) opinion who mentions, "high-quality studies combine both qualitative and quantitative analysis of texts and mixing methods is generally acknowledged as an effective way to ensure the trustworthiness of the research in terms of validity and reliability. Thus, based on the explanation above, it is clear that this research was mixed-method with Triangulation Design: Data Transformation Model.

The objects of this study were teacher-made tests taken from the mid-term test questions and written by five English teachers at a favorite Senior High School in Bengkulu City. The distribution of those tests is presented in the following table;

Table 1. The Distribution of the Teacher Made Tests

No	Teacher-Made Tests	Grade	Number of Questions
1	Test A	X MIPA	30
2	Test B	X IPS	26
3	Test C	XI MIPA and IPS	25
4	Test D	XII MIPA	25
5	Test e	XII IPS	30
Total			136

MIPA= Natural Sciences, IPS = Social Sciences

As presented in Table 1, those tests were used in the Mid-test to test students' English learning outcomes at every grade Natural Sciences and IPS. However, for grade XI, there was only one teacher-made test because it was used for both grade XI IPA and XI IPS. The total of questions made by them was 136 questions. The types of the questions were 125 multiple choice questions, 1 completion item question, and 10 essay questions. Moreover, a purposive sampling technique was used to determine the objects of this study. As stated by Sugiyono (2013) purposive sampling is a technique of taking data sources with certain consideration certain. The consideration applied in determining the objects of this research was as follows;

1) The Senior High School was one of the favorite schools in Bengkulu city and Bengkulu province. Moreover, this school was more welcome for the researcher in doing the research than other schools.
2) This school was applied the 2013 curriculum and also this school had implemented HOTS in the teaching and learning process of English subject.
3) The teachers were certified teachers. They have been teaching English more than 10 years. Those teachers have been categorized as experienced teachers (Jasrial et al., 2021), it assumes that they have more teaching abilities and higher performance levels (Rakib et al., 2016), including in making question tests.

The data of this study were analyzed by using keyword analysis table as a guideline in determining the level of thinking skill in those test items. The keyword analysis table was adapted from Southeastern University site. The table provided the understanding of each level of thinking skill from the lower to higher order thinking skill. It also provided some action verbs or keywords that related to each level and provide some examples of questions based on the level of thinking skill. To classify all the test items the researcher needed this table as a guideline.

After classifying all the test items, the frequency of each group of cognitive domains of the revised Bloom taxonomy were calculated. Then, the composition of HOTS on the teacher made-test was calculated. Next, data were drawn which explained previous discussion based on research question to get the conclusion. The last, the quality of the composition of HOTS on the teacher made test was interpreted by using the worthiness scale criteria of Riduwan (2013, p. 22).

Table 2. The Proportional Scale Criteria of HOTS's Composition on Teacher Made Test

No	Percentage Range	Category
1	81% - 100%	Verry Proportional
2	61% - 80%	Proportional
3	41% - 60%	Fair
4	21% - 40%	Less Proportional
5	0 – 20%	Least Proportional

Riduwan (2013, p. 22)

As shown in Table 2, there were five criteria of the quality of the teacher-made test. Those criteria were adapted from Riduwan (2013, p. 22) by changing the criteria from Very Good into Very Proportional, Good into Proportional, Fair into Fair, Poor into Less Proportional, and Very Poor into Least Proportional. The aim of the researcher changed the criteria of the quality of teacher-made Test based on the composition of HOTS was language problem. The category of the previous criteria used too strong judgment to interprete the quality of the teacher-made tests. The criteria of Riduwan (2013) can be used for assessing quality of test and all level of test because this theory was designed for assessment to see the worthiness of a test including the quality of a test based on HOTS proportion. Moreover, this theory had been used by previous researchers in journals to interpret the worthiness of HOTS in a test such as Wahyuni et al. (2019) and Septiarini and Puspasari (2020).

JOURNAL OF ENGLISH TEACHING, APPLIED LINGUISTICS AND LITERATURES (JETALL)
English Language Education Study Program, FKIP Universitas Lambung Mangkurat Banjarmasin
Volume 5 Number 2 2022

To ensure the validity of the analysis, the researcher involved a co-rater to avoid subjective judgment. The co-rater was an alumna of the English Education Study Program who experienced in conducting the same research in terms of analyzing HOTS in the tests or questions. Moreover, the reliability of the analysis between the rater and co-rater was determined by using the Cohen Kappa agreement. The agreement scales of Cohen's Kappa were 0-20 was poor, 0,21-0.40 was fair, 0.41-0.60 was moderate, 0.61-0.80 was good, and 0-81-1.00 was very good (Arikunto, 2010). The results of the inter-ratters' reliability are presented in the following table;

Table 3. The Results of Inter-ratters' Reliability

No	Teacher-Made test	Kappa Value	Category
1	Test A	1.00	Very Good
2	Test B	0.96	Very Good
3	Test C	0.97	Very Good
4	Test D	0.97	Very Good
5	Test E	0.88	Very Good
Average		0.96	Very Good

As presented in Table 3, the results of the inter-ratters' reliability in analyzing the data were in a very good category. The average of the Kappa Value was also in a very good category. It means that the result of the analysis was reliable.

RESULT AND DISCUSSION

The Proportion of HOTS on Teachers-Made Tests

This first question aimed at investigating the proportion of HOTS on the teacher-made tests constructed by five English teachers at a favorite school in Bengkulu City. The results of the analysis of HOTS are presented in the following table.

Table 4. The Proportion of Teacher-Made Test Based on the Cognitive Skills of Revised Bloom's Taxonomy

No	Tests	f/p	LOTS			HOTS			Total
			C1	C2	C3	C4	C5	C6	
			Remember	Understand	Apply	Analyze	Evaluate	Create	
1	Test	f	0	26	2	1	0	1	30
	A	p	0%	87%	7%	3%	0%	3%	100%
2	Test	f	1	22	0	3	0	0	26
	B	p	4%	84%	0%	12%	0%	0%	100%
3	Test	f	1	13	0	10	0	1	25
	C	p	4%	52%	0%	40%	0%	4%	100%
4	Test	f	1	11	3	8	2	0	25
	D	p	4%	44%	12%	32%	8%	0%	100%
5	Test	f	1	18	0	9	1	1	30
	E	p	3%	60.5%	0%	30.5%	3%	3%	100%

Total	f	4	90	5	31	3	3	136
	p	3%	66%	4%	23%	2%	2%	100%

f= Frequency, p= Percentage, Total N = Total frequency of all question of every teacher

As presented in Table 4, the tests that were made by the English teachers were categorized as LOTS and HOTS. The most frequent teacher-made test composed by the English teachers was LOTS questions. It means that the questions that the teacher made have not required the ability to think critically and creatively. The most frequent cognitive domain used in the tests was C2 (Remember). The examples are as followings;

Excerpt 1

Essay

1. What is general statement?
2. In the description of a report text, what kind of information can you include in it?
3. In what tense is a report text written?

(Test C Question No 3, Grade XI IPA/IPS)

Question number 3 in Example 8 was categorized as C1 (Remember) based on the revised Bloom's taxonomy. It was because this question asked the students to remember the type of tense used in the report text. They did not need to think to answer this question, but they just recalled their memory. They would answer based on what they had learned about the report text. This type of question was still in the LOTS category.

Excerpt 2 – Understand (C2)

Mr. Rizky : Hello, Miss Lily. Would you like a cup of tea? I'm just making some.
Miss Lily : Oh, yes please, that would be lovely.
Mr Rizky : How do you take it?
Miss Lily : With milk and sugar, please.
Mr. Rizky : Here you are.
Miss Lily : Thank you.

1. The underlined expression expresses ...
   A. Offering help
   B. Accepting an offer
   C. Offering something
   D. Declining an offer
   E. Refusing an offer

(Test A, Question No 1, Grade X Sciences)

The question in Excerpt 1 was categorized as the understand (C2) level. In this question, the students were asked to understand the underlined expression. Then, they were asked to

classify the underlined expression by choosing one of the best answers in the option. Thus, the cognitive process in this question was classifying the type of expression in a conversation.

Excerpt 3 – Apply (C3)

I was sleeping when somebody knocked the door.

13. The correct phrase of past perfect tenses is ...
    a. I was slept when somebody knocked the door.
    b. I had slept when somebody knocked the door.
    c. I have slept when somebody knocked the door.
    d. I has slept when somebody knocked the door.

(Test A, Question No 12, Grade X IPA)

The example in Excerpt 2 was categorized as the apply (C3) level. The teacher provided a sentence. The cognitive process applied in this question was procedure knowledge in composing a correct sentence by using the past perfect tense. The students were asked to apply it correctly. Thus, in this question, they would choose a correct sentence in the form of the past perfect tense by choosing the best answer in the option.

Excerpt 4 – Analyze (C4)

John : Why was Erika absent yesterday
Ani : Her father passed away
John : I am deeply sorry to hear that. Why didn't you tell me before?
Ani : I have just heard about it.

8. From the dialogue we can conclude that...
    a. Ani happy to hear the news
    b. Erika absent because her father passed away
    c. Ani tells the sad news to John
    d. Erika goes out of the school
    e. Ani and John is not join the class today

(Test A, Question No 8, Grade X IPA)

The question in Excerpt 3 above was categorized as the Analyze (C4). In this question, the students would read the content of the conversation. After that, they applied the cognitive process in terms of the ability of concluding the content of the conversation. The last, they would choose the best answer in the option based on the conclusion that they made after reading and analyzing the conversation.

Excerpt – C5 (Evaluate)

13. What can you learn from the text?
    a. We should help others in need
    b. We should live a modest life
    c. We should keep our health well
    d. We should take care of ourselves well
    e. We should concern about the surrounding.

(Test D, Question No 13, X1I IPA)

The question in the Excerpt 16 was categorized as C5 (Evaluate) which was in the HOTS category. This question asked the students to take the lesson from the text. To answer this question, the students needed to evaluate, criticize, and interpret the content of the text carefully as part of the cognitive process. After that, they would decide the lesson from the text from social point of view.

Excerpt 6: Create (C6)

B. Write the correct answer of the blank sentences below!

1. Write two sentences of simple past tense!

(Test A, Question No 1 – Essay, Grade X IPA)

The question in Excerpt 4 was in the form of essay. It was categorized as the create (C6) level because the students were asked to create sentences in the form of the simple past tense. The cognitive process applied in this question was the ability of making or creating own correct sentences based on the structure of the simple past tense.

This finding is similar to the previous studies' findings in terms of teacher-made test (Amaliyah, 2018; Pratiwi et al., 2019; Utami et al., 2019; Putri et al., 2020; Syahdanis et al., 2021), national examination test (Ilham et al., 2020), reading exercises on English textbooks ('Ilma, 2018) and language skill tasks on English textbook (Febriyani et al., 2020). They reported that the questions of the test made by teacher and exercises on the textbook made by the authors were dominated by LOTS questions. In other words, the quality of their tests as instrument to assess the students' learning outcomes that require the ability to think creatively, critically, and logically has not fulfilled a criterion of a good test.

Regarding the proportion of the questions that contained HOTS is lower than LOTS, the most frequent cognitive domain used in the test was C2 (Understanding). It was in the LOTS category. This is because this type of test is dominantly designed in the form of multiple-choice questions. Martinez (1999; Magno, 2003) confirm the previous statement that the multiple-choice questions often assess low thinking. This type of test will be difficult to evaluate students' higher levels of thinking, such as C4 (analyze), C5 (evaluate), and C6 (create) (Gareis & Grant, 2015). Moreover, multiple-choice questions are hard to encourage productive, innovative, or creative thinking, and students will only choose the options presented in the test. However, it does not mean that the multiple-choice question cannot measure the higher-order thinking skills of the students in the tests.

In constructing or composing a test, the multiple-choice questions can measure higher levels of thinking. Liu (2009) suggests that some techniques can be applied by the test makers in constructing the multiple-choice questions to measure students' higher levels of thinking are using a combination of tests format (for example, combining multiple-choice with responses that are constructed), providing factual statements and ask students to analyze, and provide data and ask questions students to develop hypotheses. Based on these techniques, the English teachers have applied those techniques in some questions, but it is dominated by the cognitive domain of C4 (Analyze).

Another possible reason that causes lower the composition of HOTS than LOTS on the teacher-made test is from the teachers' and students' aspects. Based on their study, Syahdanis et al. (2021) reported that the reasons come from the teachers' and students' aspects. In other words, their knowledge of HOTS is low. From the teachers' aspect, they still need more professional HOTS training. This way is quite effective to improve teachers' understanding of HOTS and have better competency standards especially in constructing the tests that contained HOTS from the students. The test of grade X did not contain 25% of HOTS composition in the tests that they made must follow this activity, but the composition of the cognitive domain of C4 (Analyze), C5 (Evaluate), and C6 (Create) must be balanced proportion. Moreover, from the students' aspect, the students will get difficulty answering the questions on the test if the HOTS questions are applied too many in the test. It will influence the students' scores in doing the test. Therefore, those reasons probably influence the composition of HOTS lower than LOTS on the teacher-made tests.

A lower number of HOTS questions was also found by Febriyani et al. (2020). This study found that most questions in the compulsory English Textbook for the twelfth grade of Indonesian Senior High Schools used the cognitive domain in terms of C1 (remembering). It means that the authors of the textbook have not realized the importance of HOTS in English learning. The difference between this previous study with this present research finding is the cognitive domain of C2 (Understand) is dominantly on the teacher-made test. In this cognitive domain, the students were asked to understand the texts, dialogues, or linguistics knowledge by reading them. Then the students will find the information questions that contained the Understanding (C2).

In conclusion, the results of this present research show that the English teachers need to increase the number of questions that contain HOTS, especially in the cognitive domain of C5 (evaluate) and C6 (create). The findings of this present research relate to previous studies (Jailani & Retnawati, 2016; Retnawati et al.; 2016) that demonstrated some teachers have difficulties in implementing learning or assessment model that fits the demands of the

Curriculum. These cases will influence the teaching and learning process. It is because the awareness of HOTS by the teachers on the teacher-made test can demonstrate that they are ready to make changes or improvements during the learning process (Retnawati et al., 2018). This awareness will improve the quality of the teaching and learning process of English at the senior high school to achieve the goal of the 2013 curriculum to prepare every individual to have critical and creative thinking, collaboration, and good communication skills in facing the global era (Yunita et al., 2020). Thus, implementing HOTS in the teaching and learning process such as in the test will help improve students' knowledge of HOTS so that they will be familiar with and easy to apply HOTS in their life.

The Quality of the Teacher-Made Test

This question aimed at investigating the quality of the teacher-made tests viewed from the compositions of HOTS. After analyzing the data and calculating the composition of HOTS on the teacher-made tests, the quality of the teacher-made tests was interpreted by using the proportion scale criteria. It was adopted from the evaluation scale criteria of Riduwan (2013, p. 22). The results are presented in the following table.

Table 5. The Proportional of HOTS on the Teacher Made Test

No	Test	N	f/p	The Composition of Teacher-Made Test		Interpretation
				LOTS	HOTS	
1	Test A	30	f	28	2	Least Proportional
			p	93.33%	6.67%	
2	Test B	30	f	23	3	Least Proportional
			p	88.46%	11.54%	
3	Test C	30	f	14	11	Fair
			p	56%	44%	
4	Test D	30	f	15	10	Less Proportional
			p	60%	40%	
5	Test E	30	f	19	11	Less Proportional
			p	63.30%	36.70%	
Total		136	f	99	37	Less Proportional
			p	72.80%	27.80%	

f= frequency, p= percentage, N= total questions

As shown in Table 5, the proportion of HOTS on the teacher-made test was lower than HOTS. It indicates that the quality of the teacher-made tests was less proportional based on the proportion scale criteria that was adopted from Riduwan (2013, p. 22). This finding shows that teachers' knowledge of HOTS, especially in constructing a test that contains HOTS is still low. This result does not support the study conducted by Yunita et al. (2020) who found that Senior High School English teachers in Bengkulu province have had very good knowledge of HOTS. It was shown by the result of the protocol test on HOTS. Almost all of them could answer the questions that contained HOTS. It is because the survey

conducted by Yunita et al. (2020) was for all English teachers in Bengkulu province, while this present study was conducted in a school in Bengkulu City.

In addition, among five teacher-made tests, there was one of them that contain the composition of HOTS as a fair proportional criterion, while others were in a less proportional and least proportional category. It means that this teacher has understood the principle and characteristics of HOTS questions when constructing a test. This test will improve students' critical thinking skill. It is supported by the statement of Serevina et al. (2019) who state that a good instrument to evaluate the students' outcome in learning is an instrument that contains good questions to test cognitive, affective, and psychomotor abilities of the students. It will assess how far the students master the materials that have given by giving some problems and they can solve those problems by using their critical thinking.

Moreover, the fair proportional criterion of teacher-made test was found at grade XI IPA and IPS. The teacher made more HOTS questions compared to other teachers who teach grade X and XII. It is possibly because Grade XI has more English learning materials than other grades (X and XII). The students of grade XI study some reading texts (e.g., personal letter, exposition text, report text, and explanation text), expressions (e.g., persuading someone, expressing happiness, and expressing sympathy), and grammar (complex sentence, gerunds, and infinitive) that challenge students' critical thinking skills. Utilization of HOTS in the teaching and learning process can be influenced by the use of materials that require students' higher thinking skills (Huda et al., 2021). The materials that contain HOTS also influence the teachers' in constructing a test that contain more composition of HOTS. Meanwhile, grade XII does not have many learning materials.

Regarding the less and the least proportional of HOTS's composition constructed by two of five English teachers at grade X (X IPA and X IPS), it shows that those tests only require lower thinking skills of the students in doing the tests. It happened possibility caused by the teachers have difficulty in constructing the test that contain HOTS questions. In this case, the teachers also need to apply the higher thinking skills in constructing the test. It means that HOTS is difficult for them because HOTS main characteristics of HOTS are critical and creative (Conklin & Manfro, 2012). They cannot adequately distinguish specific keywords for a certain level from HOTS questions (Afifah & Retnawati, 2019) such as operational verbs of C4, C5, and C6. As a results, they construct the test with the same way in terms of type of test (multiple choice) and kinds of questions (C2 questions).

The less and the least proportion of HOTS' composition on the teacher-made test will affect the way of thinking of the students either in their learning process or in their daily life. It influences the development of cognitive process dimension of students in learning English (Agung et al., 2017), while HOTS is needed by every individual in facing the global era because it requires critical and creative thinking, collaboration, and good communication skills (Yunita et al., 2020). Therefore, these skills must be taught and introduced to the students at schools during the teaching and learning process and on the test as a form of critical and creative thinking practical for students.

Moreover, HOTS has been embodied in the 2013 curriculum. It means that the teachers cannot ignore the implementation of HOTS in the education field. The less proportional of HOTS on the teacher-made tests means that the teachers have not been brave to implement it for their students. In the 2013 curriculum, the type of test is suggested to be used by the teachers is authentic assessment and it must contain HOTS. However, the

unexpected results of this present research are also in line with the results of study conducted by Hajaroh and Adawiyah (2018) which shows that teachers are still having difficulties in implementation authentic assessment. Similarly, Riadi and Hilal (2017) reported the results of their research that not all teachers make assessment instruments for the affective, cognitive and psychomotor domains. This finding shows that the assessment constructed by the teachers need to be an improvement in terms of cognitive assessment. It aims to achieve the goals of the implementation of HOTS in a test. The assessment-based HOTS or test-based HOTS as embodied in the 2013 curriculum becomes an ability that must be possessed by students. However, based on the finding of this present research, the implementation of HOTS at school has not fulfilled the expectation of the government. It causes the way of thinking of students will be low and it is not suitable with cognitive domain of HOTS based on the Bloom's taxonomy. Kusuma et al. (2017) state that HOTS is able to lead the individual to have the ability to apply knowledge, skills, and values in reasoning, reflection, problem-solving, decision making, innovating, and creating new things. Thus, by implementing HOTS on the test, the teachers have helped the government in improving the students' critical thinking in their daily life.

Some strategies have been done by Indonesian government for implementing learning processes and assessments based on HOTS at school. However, some factors cause the low achievement of HOTS in Indonesia. They are; Firstly, Indonesian students are still not familiar to do HOTS questions on the tests so that they teachers often use LOTS questions on the test. Secondly, many teachers find it difficult to arrange questions HOTS so that it uses existing and previously created questions that are still in the LOTS category which resulted in students not trained to solve HOTS-based questions.

Based on the results of the second question, it assumes that the English teachers still implement the less proportional HOTS questions on their tests. It is caused by several reasons as explained above. The important thing needs to know is teachers are agents who transfer the skills and knowledge of HOTS to the students. The assessment or test is also given by them so that they influence the achievement of the HOTS ability in the education field. Therefore, constructing a test that contain many HOTS questions is important to be done by them to achieve the goal of the implementation of HOTS.

Based on the findings of this research, some strengths and weaknesses are found in this research. In terms of the weaknesses, this research did not explore the reasons for the less appearance of HOTS in the English teacher-made test. Then, this research was conducted at one school that involved only five Senior High School English teachers. Next, the analysis of the teacher-made test is not viewed from the length of teaching experiences of English teachers. Further, the analysis focuses on the teacher-made test in the form of Mid Term Test. The last is the criteria for a good test proposed by Riduwan (2013). This criterion is limited to the number of the composition of HOTS questions in all tests, not viewed from the proportion based on each cognitive domain of bloom's taxonomy. The strength of this research is the samples involved in this research were from a good reputation of Senior High School in Bengkulu. As a result, this finding gives a description and an evaluation of the knowledge of the English teachers on HOTS.

The findings of this research provide pedagogical significance. This research presents the theory and analysis of HOTS of the Teacher-Made Test. Hopefully, it can provide broader insight to the English teachers at the same level of education in Bengkulu on HOTS

knowledge from the revised edition of Bloom's Taxonomy and in designing a better good test for their students.

CONCLUSION

This research has analyzed the proportion of HOTS on the teacher made test by English teachers at a favorite Senior High School in Bengkulu City in the academic year of 2021/2022. Based on the results, the proportion of HOTS on the teacher-made tests is lower than the LOTS. Among all cognitive domain levels, C2 (Understand) is the most frequent cognitive domain level found on the teacher-made test which is in the LOST category. Meanwhile, the composition of HOTS questions on the teacher-made test is dominated by the cognitive domain of C4 (Analyze), while other cognitive domains of HOTS such as C5 (Evaluate) and C6 (Create) are rarely found on the teacher made test. Moreover, the quality of the teacher-made tests viewed from the proportion of HOTS is in a less proportional criterion. One of five English teachers-made test has the quality in a fair criterion, two teacher-made tests are in a less proportional criterion and the others are in a least proportional criterion

REFERENCES

Amaliyah, A. (2018). An analysis of multiple choice items made by teacher based on Bloom's revised taxonomy theory at tenth grade of SMAN 1 Sidoarjo [*An Undergraduate Thesis, Universitas Islam Negeri Sunan Ampel*].

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of bloom's taxonomy of educational objectives*. New York: Allyn & Bacon.

Arikunto, S. (2005). *Dasar-dasar evaluasi pendidikan*. Jakarta: Bumi Aksara.

Conklin, W., & Manfro. (2012). *Higher Order Thinking Skills to develop 21st century learners*. Huntington: Shell Education Publishing, Inc.

Creswell, J. W., & Clark, P. V. L. (2007). *Designing and conducting mixed methods research*. Thousand Oaks CA: Sage Publications.

Febriyani, R. A., Yunita, W., & Damayanti, I. (2020). An Analysis on higher order thinking skill (HOTS) in compulsory English textbook for the twelfth grade of Indonesian Senior High Schools. *Journal of English Education and Teaching*, *4*(2), 170-183. https://doi.org/10.33369/jeet.4.2.170-183.

Gareis, C. R., & Grant, L. W. (2015). *Teacher-made assessments: How to connect curriculum, instruction, and student learning.* (2nd edition). Oxford, United Kingdom: Routledge.

Hajaroh, S., & Adawiyah R. (2018). Kesulitan guru dalam mengimplementasikan penilaian autentik. *Midad Jurnal Jurusan PGMI, Vol 10* No 2. 131-152.

Huda, M., Purnomo, E., Anggraini, D., & Prameswari, D. H. (2021). Higher order thinking skills (HOTS) dalam materi dan soal pada buku pelajaran Bahasa Indonesia SMA

terbitan Kemendikbud RI. *Prasi: Jurnal Bahasa, Seni, dan Pengajarannya, Vol 16* No 2, 128-143. http://dx.doi.org/10.23887/prasi.v16i02.40671.

'Ilma, F. (2018). Higher order thinking skill questions in reading exercises: An analysis of reading exercises in Bright course book for seventh grade of Junior High School students 2013 curriculum published by Erlangga in the year of 2016 [*An Undergraduate Thesis, Universitas Islam Negeri Walisongo*].

Jailani, & Retnawati, H. (2016). The challenges of junior high school mathematic teachers in implementing the problem-based learning for improving the higher-order thinking skills. *The Online Journal of Counseling and Education, Vol 5* No 3, 1–13.

Jasrial, D., Yunita, W., Sukma, B., & Septianggaira, A. (2021). Appreciating students' responses: Verbal and non-verbal compliments used by English teachers in classroom. *Pedagogy: Journal of English Language Teaching, Vol 9*, No 2. 165-176. https://10.32332/joelt.v9i2.3650.

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), Vol 07* No 01, 26– 32. https://doi.org/10.9790/7388-0701052632

Liu, X. (2009). Essentials of science classroom assessment. Thousand Oaks, California: Sage Publications.

Magno, C. (2003). The profile of teacher-made test construction of the professors of University of Perpetual Help Laguna. *UPHL Institutional Journal, Vol 1* No 1, 48-55.

Martinez, M. E. (1999). Cognition and the questions item format. *Educational Psychologist, Vol 34* No 4, 207-218. https://doi.org/10.1207/s15326985ep3404_2

Pratiwi, N. P. W., Dewi, N. L. P. E. S., & Paramartha, A. A. G. (2019). The reflection of HOTS in EFL teachers' summative assessment. *Journal of Educational Research and Evaluation, Vol 3* No 3, 127-133.

Putri, A. H, (2020). An analysis of higher order thinking skill (HOTS) and lower order thinking skill (LOTS) in test constructed by twelfth grade English teacher at SMAN 6 Kota Bengkulu [*An unpublished thesis, Universitas Bengkulu*].

Rakib, M., Rombe, A., & Yunus, M. (2016). Pengaruh Pelatihan Dan Pengalaman Mengajar Terhadap Profesionalitas Guru. *Jurnal Ad'ministrare, Vol 3* No 2, 137–148.

Retnawati, H., Djidu, H., Kartianom, Apino, e., Risqa D, & Anazifa. (2018). Teachers' knowledge about Higher-Order Thinking Skills and its learning strategy. *Problems of Education in The 21st Century, Vol 76* No 2,215-230.

Retnawati, H., Hadi, S., & Nugraha, A. C. (2016). Vocational high school teachers' difficulties in implementing the assessment in curriculum 2013 in Yogyakarta Province of Indonesia. *International Journal of Instruction, Vo 9* No 1, 33–48. doi:10.12973/iji.2016.914a.

Ramadhanti, N. (2019). Pengaruh pengalaman mengajar dan banyaknya pelatihan yang pernah diikuti terhadap kinerja guru SD Sekbin 1 Kecamatan Ketanggungan

Kabupaten Brebes [*A bachelor thesis, Universitas Negeri Semarang*]. 1401415197__Optimized.pdf (unnes.ac.id)

Riduwan. (2013). *Dasar-dasar statistik*. Alfabeta.

Riadi, B., & Hilal, I. (2017). Instrumen penilaian buatan guru Bahasa Indonesia SMP di Kota Bandar Lampung. *AKSARA Jurnal Bahasa dan Sastra, Vol 18* No 2,145-153.

Sulaiman, T., Muniyan, V., Madhvan, D., Hasan, R., & Rahim, S. (2017). Implementation of higher order thinking skills in teaching of science: A case study in Malaysia. *International Research Journal of Education and Sciences, Vol 1* No 1, 1-3.

Sándorová, Z. (2014). Content analysis as a research method in investigating the cultural components in foreign language textbooks. *Journal of Language and Cultural Education, Vol 2* No 1, 95-198.

Septiarini, A., & Puspasari, D. (2020). Pengembangan LKPD berbasis HOTS dan inkuiri terbimbing pada Mata Pelajaran Otomatisasi Tata Kelola Humas dan Keprotokolan kelas XII OTKP Semester Gasaldi SMKN10 Surabaya. Jurnal *Pendidikan Administrasi Perkantoran,Vol 8* No 1, 9-21.

Serevina, V., Sari Y. P., & Maynastiti, D. (2019). Developing high order thinking skills (HOTS) assessment instrument for fluid static at senior high school. *IOP Conf. Series: Journal of Physics: Conference Series, Vol 1185* No 2019, 1-9. http://dx.doi.org/doi:10.1088/1742-6596/1185/1/012034.

Sugiyono. (2013). *Metode penelitian kuantitatif, kualittative, dan R & D*. Bandung: Alfabeta.

Syahdanis, J. D., Sofyan, D., Yunita, W. (2021). Analysis of HOTS in English teacher made test. *Jurnal Basis: Jurnal Bahasa dan Sastra Inggris Universitas Putera Batam, Vol 8* No 2, 125-138. https://doi.org/10.33884/basisupb.v8i2.4479.

Utami, F. D. ,Nurkamto, J., & Marmanto, S. (2019). *Higher Order Thinking Skills on Test Items Designed by English Teacher*. *International Journal of Educational Research Review, Vol 4*, 756-765. http://dx.doi.org/10.24331/ijere.629581.

Wahyuni, A. R., Tamrin, A. G., & Nurhidayati, A. Implementasi higher order thinking skills soal ujian dasar-dasar konstruksi bangunan dan teknik pengukuran tanah. *Indonesian Journal of Civil Engineering Education, Vol 5* No 2, 9-14. https://dx.doi.org/10.20961/ijcee.v5i2.43480.

Warmadewi, P. S., Agustini, D., & Wedhanti, N. (2019). Analysis of learning higher order thinking skill (HOTS) toward English subject. *Jurnal Penelitian dan Pengembangan Sains dan Humaniora, Vol 3 No 2*, 134-140. http://dx.doi.org/10.23887/jppsh.v3i2.21278.

Yunita, W., Syahrial., & Hati, G. M. (2020). English teachers' knowledge on higher order thinking skills (HOTS). *English Review: Journal of English Education,* Vol 9 No 1, 205-216. https://doi.org/10.25134/erjee.v9i1.3800.