EVALUATING ENGLISH TEACHERS' TEST ITEMS CONSTRUCTION BASED ON THE REVISED BLOOM'S TAXONOMY

Skolastika Neldis[1], Stanislaus Guna[2], Ely Heldydiana Selamat[3*]

[123]Universiatas Katolik Indonesia Santu Paulus Ruteng,
astineldis@gmail.com, gunastanis@gmail.com,
ely.selamat@gmail.com

*Corresponding Author

Abstract: This study aims to find out how senior high school English teachers constructed the final English item tests for grade XII Language Program students under the cognitive domain of the revised Bloom's taxonomy. This study utilized a descriptive quantitative approach. The source of data for this study was 50 multiple-choice items from final English tests for grade XII that were constructed by six English teachers. The result of this research indicated that out of 50 item tests, 39 items or 78% were LOTs-based, 4 items or 8% in MOTs-based, and 7 items or 14% in HOTs-based. Besides, the tests only applied four of six levels of cognitive domain, such as remembering (C1), understanding (C2), applying (C3), and analysing (C4). It is proposed that future test constructions have a fairer distribution of item types across all six cognitive levels. This will allow for a more accurate evaluation of the broad range of cognitive abilities that students possess. With this approach, not only will assessments be more closely aligned with educational standards, but it will also create a more comprehensive evaluation of the learning and comprehension of students.

Keywords: *EFL Teachers, English test, HOTs, MOTs, LOTs*

INTRODUCTION

Students in this 21st century are required to possess skills such as communication, collaboration, critical thinking, and creativity (Rachmawati & Purwati, 2021). In light of this, the Indonesian Ministry of Education and Culture tries to make progress education's quality through the implementation of 2013 curriculum which focuses more on the improvement of those skills (Khaeruddin *et al*. 2023: 47). Therefore, schools have to deal with the necessity of integrating high-order thinking skills to help students get ready for the workforce of the 21stcentury. As the result, schools are expected to stimulate students not merely on the level of remember, recall, understand, and apply the same routine steps of activities, but also able to develop students who can think creatively and critically.

Bloom developed three domains in his taxonomy, such as cognitive domain, affective domain and psychomotor domain. Each of these domains measures a particular object and they are ordered from simple level (low-order level) to complex level (high-order level) (Al-Khayyat, 2020). Bloom's taxonomy of cognitive domain widely use in educational field, it consist of six levels namely, knowledge, comprehension, application, analysis, synthesis, and evaluation. In 1990, Lorin Anderson revised Bloom's taxonomy of 1956 to meet the 21st century pedagogical areas. The revision was about the terminologies, which is from noun to verb form. In the revision of Bloom's taxonomy, the six level of cognitive domain became remembering, understanding, applying, analyzing, evaluating, and creating. Low-order thinking skills (LOTs) consist of remembering and understanding, medium-order thinking skills (MOTs) includes applying, and high-order thinking skills (HOTs) include analyzing, evaluating, and creating (Kemp *et al*. 2022).

Tyas *et al*. (2020) defined HOTs as critical thinking and creative thinking. Teachers are expected to motivate students to think critically and creatively by incorporate and use them in teaching learning process as well as in doing the evaluation. Teachers can use the revision of Bloom's taxonomy as a

guide when developing learning objectives and assessment. Utilizing it may develop a well-structured curriculum and serves as the foundation in designing instruments for assessment that accurately evaluate the achievement of learning objectives (Freihat, 2020:67).

However, in Indonesian context it is still difficult to apply merely HOTs in formulating questions. This can be approved by the study conducted by Fitriani (2019), who investigate the distribution of HOTs in test items made by teachers. The result revealed that LOTs were dominantly used in formulating the test. Another researcher is Fitriani (2021) who carried out a study on analyzing the use of LOTs and HOTs on lectures' questions in university classroom. The result show that even though the lectures utilized HOTs in giving questions, however the number are relative low compare to LOTs questions. Therefore, school examination item tests must apply the combination of all the cognition levels. Formulate all the questions in HOTs alone might be difficult for students. Meanwhile, elaborate those in LOTs and MOTs alone also will not trigger students' higher thinking skills.

Assessing students' knowledge and skills is important in the educational environment and it is always attached to the teaching learning process. Giving tests is a technique of assessing students' knowledge and skills. Dwiyono *et al*. (2021) stated that test is conducted to measure students' outcomes in particular subject. In addition, Khasanah (2020) said that the aim of examination is to have a picture of students' achievement during their study process which includes the aspect of knowledge, attitude and skills. Therefore, it supposedly to encourage the improvement of their HOTs in the formulation of the questions. Indonesian government as well encourage teachers to provide students with HOTs-based exam appropriate to Bloom's taxonomy and in teaching and learning process through the implementation of the 2013 curriculum . If HOTs is applied, it means that the exam test leads students to analyze, evaluate, and create solution to problem rather than just ask students to memorize or other level of questions that belong to lower-order thinking skills.

Senior High School is the level of education that prepares students to continue their education to a higher level or entering the

job markets. Learning English is necessary to enhance one's career prospects in this century, keeping up with technology advancement and pursuing academic goals that allow a person to access global information (Bayaydah & Altweissi, 2020). English is also being tested in the final exam. Schools have the independence to assess students' competencies, whether in the form of written tests or other form of assessments. Typically, a number of questions are used to figure out students' comprehension level such as questions about the main idea, vocabulary in context, detail information, reference, and inference (Nisa & Helmanda, 2020:74). It is good to apply HOTs in formulating those questions. Although there is a need for HOTs-based assessment, still problems regarding the integration of HOTs in teaching learning process, especially in assessment are found.

In the previous years, the final exam of senior high school was carried out in the form of a National Standardized School Examination or USBN (Ujian Sekolah Berstandar Nasional) (Dwiyono *et al*. 2021). However, in 2019 Minister of Education, Culture, Research and Technology Nadiem Makarim launched the Merdeka Belajar Episode 1 policy. One of the important points of this new policy is the elimination of the National Standardized School examination and the decision regarding student's graduation is returned to the school through school exam (Kemendikbud, 2022). In 2023, most of the schools utilized the written test for final school exam. For English subject, it consists of 50 multiple-choice items and focuses on reading as well as language elements like grammar and vocabulary (Uswatunnisa, 2020).

There have been studies regarding to revised Bloom's taxonomy over item tests. The first previous study was carried out by Chandio *et al*. (2021). This study attempts to investigate the manifestation of Bloom's taxonomy in the annual question papers made by the Boards of Intermediate and Secondary Education (BISE) of Sindh province of Pakistan for English subject. The result show that 74% of the questions refer to LOTs whereas only 26 % are used HOTs. Moreover, there is another pervious research conducted by Belarbi and Bensafa (2020). In this study, the researchers

analyzed the realization of Bloom's taxonomy on Algerian EFL baccalaureate exam. The result revealed that the exam does not stimulate students' higher-order thinking skills and they suggested for the test makers to improve the quality of the EFL baccalaureate exams with elaborate more on HOTs. Finally, a research done by Setiyana and Muna (2019) who evaluate the manifestation of HOTs based on revised Bloom's taxonomy in university' final test item. The result indicated that the most dominant used was LOTs (98%) and the used of HOTs was only 2 %.

From those previous studies can be seen that various studies have been conducted in various context of investigating the manifestation of cognitive domain of Bloom's taxonomy in item tests. However, there is few studies on the context of the distribution of cognitive domain of revised Bloom's taxonomy in Indonesian senior high school's final item test of English. Thus, the current research fills this gap by going depeer about this topic entitled:"The Analysis of Final English Item Tests At Senior high school Based on Bloom's Taxonomy". The aim of this study is to evaluate the distribution of LOTs, MOTs, and HOTs in English item tests constructed by the senior high school English teachers, for grade XII Language Program students in final exam.

METHOD

This research used descriptive quantitative to answer the research question of evaluating the distribution of LOTs, MOTs, and HOTs in English item tests constructed by the senior high school English teachers, for grade XII Language Program students in final exam. This approach is considered appropriate because it can help the researchers to identify the types of test items based on the revised Bloom Taxonomy by calculating the percentage and frequency of each category. To this end, the data were gathered from the six English teachers by purposefully selected the test items from the final test document. Moreover, the data were obtained from the reading tests sections on English subject constructed by the English teachers for grade XII Language Program students. In analyzing the

data, the researchers listed the reading tests into checklist table. The example of the checklist table used is seen in table 1.

Table 1. Checklist Table

No	**Question**	**Cognitive Domain of Revise Bloom's taxonomy**					
		Lower-Order Thinking Skill (LOTs)		Medium-Order Thinking Skills (MOTs)	Higher-Order Thinking Skills (HOTs)		
		C1	C2	C3	C4	C5	C6

Modified from The Revised Bloom's Taxonomy (2000)

The questions listed on the checklist table were checked following the indicators from analysis card of Anderson and Krathwohl by Ardiana *et al*. (2020).

RESULT

The result of the analysis of the cognitive domain, which has been applied in final English item tests constructed by the teachers at grade XII Language Program students, is shown in table 2.

Table 2. The Frequency of Cognitive Domain in Final English Item Test

Cognitive Domain		Frequency
LOTs	C1 Remember	26
	C2 Understand	13
MOTs	C3 Apply	4
HOTs)	C4 Analyze	7
	C5 Evaluate	-
	C6 Create	-
	Total	**50**

Based on table 2, there were 26 items of question in remembering level (C1), 13 items of understanding level (C2), 4 items of applying level (C3), 7 items of analyzing level (C4), and no questions applied evaluating (C5) and creating (C6). The most applied cognitive level is remembering (C1) with 26 items of the total questions. Besides, questions that belong to LOTs were dominantly utilized rather than those belong to MOTs or HOTs. The manifestation of the cognitive domain presents in the chart 1.

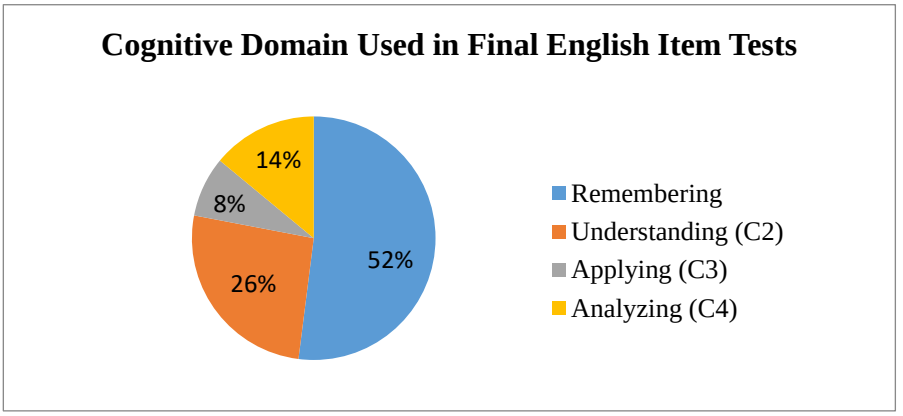

Chart 1 Cognitive Domain Used in Final English Item

Chart 1 revealed that from the six levels of cognitive domain in the revised Bloom's taxonomy, such as remembering (C1), understanding (C2), applying (C3), analyzing (C4), evaluating (C5), and creating (C6), only four levels namely remembering (C1), understanding (C2), applying (C3), analyzing (C4) were applied in constructed the final English item tests.

The English teachers at Senior high school applied LOTs, MOTs, and HOTs in constructing final English item tests. However, the distribution of questions among these three thinking skills category were unequal. Moreover, the last two levels of HOTs category were not being used in the questions. Questions were mainly constructed in LOTs which covered more than half of the total questions. HOTs came as the second in place, even though it was merely used analyzing level (C4). The percentage of LOTs, MOTs, and HOTs used by the teachers in constructing the English item tests is seen in the table 3.

Table 3.The percentage of Cognitive Domain in Final English Item Tests

Cognitive Domain		Frequency	Percentage	
LOTs	C1 Remember	26	52%	78%
	C2 Understand	13	26%	
MOTs	C3 Apply	4	8%	8%
HOTs)	C4 Analyze	7	1 4%	14%
	C5 Evaluate	-	-	
	C6 Create	-	-	
	Total	**50**	**100%**	**100%**

Table 3 presented that the item tests with LOTs question are 78% or equivalent with 39 items, MOTs questions are 8% or 4 items, and HOTs questions are 14% or 7 items. It is shown that the percentage in the usage of LOTs in item for final English tests for grade XII students is higher than MOTs and HOTs.

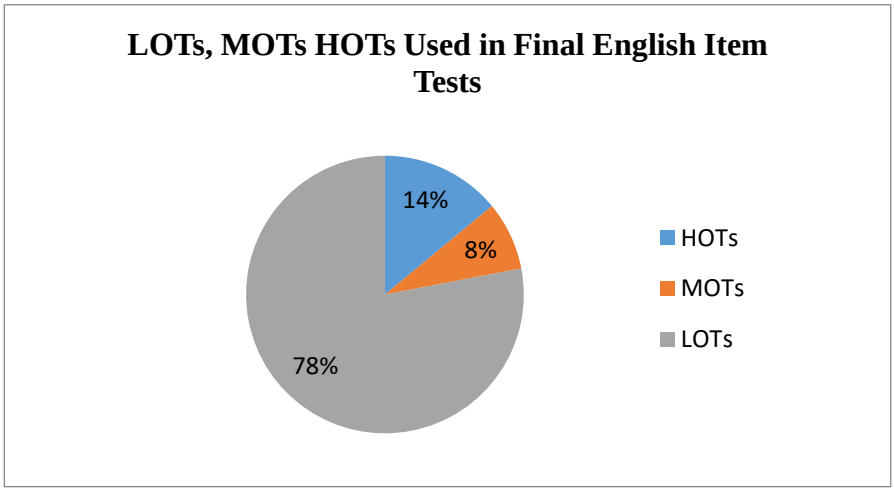

Chart 2 LOTs, MOTs HOTs Used in Final English Item Tests

Chart 2 revealed that the application of LOTs question items in the final English tests constructed by the teachers for grade XII Language Program students is more significant than the use of MOTs and HOTs. There is a big gap between the applications of these three categories of cognitive domain. It shows that most of the

questions applied are to trigger student's ability to recognize, recall and understand the materials that they have learned previously. The questions that may train students' ability to apply the knowledge they have acquired is only 8 % and 14% of questions can trigger students' higher-order thinking skills.

DISCUSSION

This study sought to identify how English teachers constructed the final English tests for grade XII Language Program students. Based on the analysis done by the researchers, it found that there is imbalance distribution of the six cognitive domains of the revised of Bloom's taxonomy. Remembering and understanding levels were more noticeable in the item tests, and there were a few questions that belong to applying and analyzing levels. The most utilized level in the tests was remembering (C1) with 26 items or 52%, the second was understanding (C2) with 13 items or 26%, the third was analyzing (C4) with 7 item or 14%, and the last was applying (C3) with 4 items or 8%.

This study showed the distribution of cognitive levels in the final tests of English constructed by the English teachers was in LOTs level. This is in accordance with the study of Muslih*a et al*. (2021), who found that teachers dominantly utilized LOTs-based questions than MOTs and HOTs questions. However, the result of this study in contrast with the study of Purbosari *et al*.(2021),who found that MOTs-based questions were dominantly used by the teachers in constructing item tests.

Indonesian Ministry of Education and Culture, through the implementation of 2013 curriculum expects to apply 10-15% HOTs-based assessment. The result of this study presents that 14% of the item tests were constructed in HOTs. This number meets the expectation of Indonesian Ministry of Education and Culture who expects to apply 10-15% HOTs in constructing item tests; however it does not reach the expectation from Programme for International Student Assessment (PISA) to apply 20% of HOTs in item tests.

In constructing the tests, the teachers applied only four of six levels of cognitive domain in the revised Bloom's taxonomy, such as remembering, understanding, applying, and analyzing. This indicates that not all the levels of cognitive domain can be assessed by multiple-choice item tests. This is in line with the result of the study of Musliha*et al*. (2021) who discovered that the high frequency of LOTs-based item English tests constructed by the teachers due to the use of multiple-choice items. Evaluating and creating were not being assessed in this current study. The elimination of national examination by the Indonesian Ministry of Education and Culture in 2019 provides more rooms for the school to feel free in assessing the students. The motif behind this new policy was for the teachers to increase the utilization of HOTs-based assessment. They are free to choose the instruments for the assessment to measures students' knowledge and skills. However, this study presented that English teachers at senior high school still following the tests pattern of national examination, which using 50 multiple-choice items in assessing students' knowledge and skills. As a result, the two last HOTs category namely evaluating and creating, were excluded from the tests.

CONCLUSION

Following a review of the English assessment in the form of final tests constructed by teachers for grade XII Language Program students based on the cognitive domain of the revised Bloom's taxonomy, 50 items were constructed applying various levels of cognitive level. The result of the study demonstrated that only four of six levels were applied: remembering (C1), understanding (C2), applying (C3), and analyzing (C4). LOTs-based items were dominantly used than MOTs and HOTs in the test. There were 39 items or 78% of the items in the HOTs category, which divided into two levels, such as remembering (C1) 26 items or 52% and understanding (C2) 13 items or 26%. There were 4 items or 8% in the MOTs category which include applying (C3). While in the HOTs category only used one level from three, namely analyzing, which reached 7 items or 14% of the total items. This result met the expectation from Indonesian Ministry of Education and Culture to

apply 10-15% HOTs-based item tests but did not meet the expectation of PISA to apply 20% HOTs-based tests.

Upon conducting the research about cognitive domain of Bloom's taxonomy in final English item tests conducted by the English teachers, the researcher suggests some significant things for teachers. Firstly, to combine the use of multiple-choice items with other test items in constructing the final tests in order to assess all the cognitive domain of the revised Bloom's taxonomy. Secondly, to apply more HOTs-based questions to trigger students' motivation in studying English.

REFERENCES

Al-Khayyat, A. S. J. (2020). Investigating the Level of Bloom's Cognitive Verbs Included in Baccalaureate English Language Exams in Iraq. *Journal of University of Shanghai for Science and Technology*, *22*(11), 118–125. https://www.researchgate.net/publication/345318945

Anderson, L.W. & Krathwohl, D. R. (Ed.). (2000). *A Taxonomy for Learning, Teaching, and Assesing: A Revision on Bloom's Taxonomy of Educational Objectives* (1st ed.). pearson.

Ardiana, A. P., & Sabiq, A. H. A. (2020). Hots-Based Analysis on English Reading Comprehension Formative Assessment. *Tarling: Journal of Language Education*, *4*(1), 19–36.

Bayaydah, A. M., & Altweissi, A. I. (2020). International Online Journal of Primary Education. *International Online Journal of Primary Education (IOJPE)*, *9*(2), 197–211.

Belarbi, M. F., & Bensafa, A. (2020). An Evaluation of the Algerian EFL Baccalaureate Exam under the Cognitive Domains of Bloom's Taxonomy. *Arab World English Journal*, *11*(4), 534–546. https://doi.org/10.24093/awej/vol11no4.34

Chandio, M. T., Zafar, N., & Solangi, G. M. (2021). Bloom ' s Taxonomy : Reforming Pedagogy Through Assessment. *Journal of Education and Educational Development*, *8*(1), 109–140.

Duesbery, L., Monegan, J. B., Liu, K., & McCoy, J. (2017). Thinking

Critically about Data Displays. *Journal of Visual Literacy*, *36*(1), 41–54. https://doi.org/10.1080/1051144X.2017.1331681

Dwiyono, Y., Mulawarman, W. G., Pramono, P. O., Salim, N. agus, & Ikhsan, M. (2021). Implementation of National Examination Based on Computer Based Test At Vocational School 1 North Sangatta. *Cypriot Journal of Educational Science*, *16*(1), 86–95. https://doi.org/. https://doi.org/10.18844/cjes.v16i1.5510

Fitriani, S. S., Yusuf, Y. Q., & Zumara, A. (2021). The Use of Cognitive Domain in Questions: The Perception of Students and Lecturers of Public Universities in Aceh. *Journal of Language and Linguistic Studies*, *17*(1), 122–138. https://doi.org/10.17263/jlls.903359

Fitriani, W. W. (2019). The Realization of HOTS on Summative Test Items Designed by English Teacher Group Discussion. *English Focus*, *2*(2). https://doi.org/10.24905/efj.v2i2.64

Freihat, N. M. (2020). Analyzing the Thinking Skills of the Jordanian GSCE's EFL Reading Comprehension Test Items. *Journal of Education and Practice*, *11*(18), 67–73. https://doi.org/10.7176/jep/11-18-08

Kemendikbud. (2022). *Pentingnya Ujian Sekolah dan Asesmen National*. Kemendikbud. ditpsd.kemendikbud.go.id

Kemp, D. J. D. P. F. M. P. R., Morrell, A. D. L. M. J., & Halse, O. (2022). Assessment of Factual Recall and Higher - Order Cognitive Domains in An Open - Book Medical School Examination. *Advances in Health Sciences Education*, *27*(1), 147–165. https://doi.org/10.1007/s10459-021-10076-5

Khaeruddin, K., Indarwati, S., Sukmawati, S., Hasriana, H., & Afifah, F. (2023). An Analysis of Students ' Higher Order Thinking Skills Through the Project-Based Learning Model on Science Subject. *Jurnal Pendidikan Fisika Indonesia (JPFI)*, *19*(1), 47–54. https://doi.org/10.15294/jpfi.v19i1.34259

Khasanah, U. (2020). *Analisis Pelaksanaan Ujian Nasional Berbasis Komputer (UNBK) pada Madrasah Aliyah Di Kota Tangerang Selatan* [Universitas Islam Negeri Syarif Hadayatullah]. https://repository.uinjkt.ac.id/dspace/bitstream/123456789/50303/1/tesis USWATUN KHASA.pdf

Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis* (2nd ed.). SAGE Publications, Inc.

Musliha, S., Sudana, D., & Wirza, Y. (2021). The Analysis of Higher Order Thinking Skills (HOTs) in the Test Questions Constructed by English Teachers. *Proceedings of the Fifth International Conference on Language, Literature, Culture, and Education (ICOLLITE)*, *595*, 610–617.

Nisa, R., & Helmanda, C. M. (2020). Analysis of Reading Comprehension Final Test at English Department of Muhammadiyah Aceh University. *Gatsempana English Education Journal (GEEJ)*, *7*(1), 72–85.

Purbosari, D., Endah, R. S. D., & Sulianto, J. (2021). Analisis Soal Penilaian Akhir Semester (PAS) Tema 6 ditinjau dari High Order Thingking Skills (HOTS) Pada Kelas VI SD Negeri 1 Klewor Tahun Ajaran 2019/2020. *DIKDAS MATAPPA: Jurnal Ilmu Pendidikan Dasar*, *4*(3), 376–384.

Purwasih, J. H. G. (2020). Kendala Calon Pendidik dalam Membuat Soal Pilihan Ganda Higher Order Thinking (HOT). *Jurnal Sosial Humaniora (JSH)*, *13*(1), 12–22.

Rachmawati, D. L., & Purwati, O. (2021). EFL Teachers' Attitudes and Competence in Developing HOTS-Based Formative Assessment. *JEES (Journal of English Educators Society)*, *6*(2), 184–196. https://doi.org/10.21070/jees.v6i2.1060

Setiyana, R., & Muna, H. (2019). Evaluating Bloom's Taxonomy in Test Items. *Proceeding of the 2nd Educatinal Conference (EEIC)*, 79–83.

Tyas, M. A., Nurkamto, J., & Marmanto, S. (2020). Cultivating Students' Higher-Order Thinking Skills in EFL Classes: The Role of the Teacher and the Textbook. *International Online Journal of Education and Teaching*, *7*(1), 267–276. http://iojet.org/index.php/IOJET/article/view/773

Uswatunnisa, U. (2020). An Analyzis of English National Exam: Test of English Proficiency for Students. *ELT Worldwide*, *7*(1), 62–69.