

Cognitive Levels in English Summative Assessment Items Under the Merdeka Curriculum

 <https://doi.org/10.31004/jele.v11i5.2710>

*Masintan, Ira Maisarah^{ab} 

¹²Universitas Bengkulu, Bengkulu, Indonesia

Corresponding Author: intanmas122@gmail.com

ABSTRACT

Summative assessment in the Merdeka Curriculum is carried out as a basis to ensure the overall achievement of learning objectives. This study examines the cognitive levels in English summative assessment constructed by teachers under the Merdeka Curriculum. This study employed a descriptive analysis with 30 test items for Grade X and 30 test items for Grade XI, which were constructed by two English teachers at Senior High School Kota Bengkulu. Each test item was examined and classified according to the cognitive process levels of the Revised Bloom's Taxonomy to identify the distribution of Lower-Order Thinking Skills (LOTS) and Higher-Order Thinking Skills (HOTS) in the assessments. The study found that English summative assessment test items developed by teachers are still dominated by lower-order thinking skills (LOTS). Most of the questions are at the Remembering level (C1), while questions that measure higher-order thinking skills (HOTS) are only found at the Analyzing (C4), Evaluating (C5) levels, and Creating (C6) levels, with relatively fewer numbers. Overall, HOTS elements have appeared in the analyzed assessments, but their proportion is still lower compared to LOTS. Therefore, English teachers need to increase the use of HOTS-based questions in summative assessments so that they are more in line with the objectives of the Independent Curriculum in developing students' critical and higher-order thinking skills.

Keywords: *Cognitive Levels, English Summative Assessment, Test Items, Merdeka Curriculum*

Article History:

Received 05th June 2026

Accepted 30th July 2026

Published 05th September 2026



INTRODUCTION

Curriculum is a guideline to conduct teaching and learning process in the classroom. In recent years, Indonesia has implemented the newest curriculum, namely the Merdeka Curriculum. The Merdeka curriculum evaluates earlier curricula. The Merdeka curriculum was developed as a foundation for a more adaptable curriculum that emphasizes student skills, character development, and significant subjects. In addition to emphasizing important content, numeracy, and differentiated instruction, the teaching and learning process is integrated with Pancasila students' profiles (Jusuf & Sobari, 2022). According to Darmawan & Winataputra (2020), the Merdeka Curriculum seeks to strengthen student independence and facilitate student-centred learning by emphasizing empowerment and the development of 21st century skills. The curriculum has an important role in organizing a teaching and learning activities, one of which is for student assessment.

Assessment is an integral component of the teaching process. It can provide information that both enhances instructions and promote learning (Zahidannisa & Hikmat, 2024). It provides students with an overview of their progress in a course, identifying individual strengths and weakness, and ultimately serves as a measure of whether students are achieving the learning objectives of the course. Assessment is the process of measuring or evaluating various aspects such as individuals, properties, performance, or ideas. Levy-Feldman (2025) stated that assessment is an integral core component of the educational process (learning teaching assessment), influencing learning and teaching. Its impact extends beyond these core elements, shaping decisions that affect students, teachers, administrators,

policymakers, and the wider community. From this perspective, assessment can be understood as a systematic process that involves collecting, analyzing, and interpreting data to make informed judgments about the quality of an individual or an object. Assessment in the Merdeka curriculum is not only limited to measuring the final results of student learning, but also evaluating the learning process experienced by students (Zahidannisa & Hikmat, 2024). Therefore, teachers are required to have adequate knowledge of student development in order to ensure that students are effectively engaged in the learning process.

Assessment consists of assessment as learning, assessment for learning, and assessment of learning, in which these three types of assessment enable teachers to gather evidence and make judgments about students' competencies. In learning, assessment as learning is a learning process used to reflect on the learning process and functions as formative assessment (Sufyadi et al., 2021). Assessment for learning, on the other hand, is an evaluation used to improve learning. This assessment serves as formative assessment. From the results of formative assessment, teachers obtain information about the need for improving learning the next day by designing learning that is positive, supportive, and meaningful (Ardiansyah et al., 2023). Assessment of learning, as an assessment at the end of learning, is used to evaluate learning. It is usually carried out at the end of a learning session. The function of Assessment of Learning is as a summative assessment. Summative assessment can be conducted at the end of a topic or semester. The purpose of this summative assessment is to measure students' learning achievement over a certain period, based on achievement criteria set by educators (Sutriyanti et al., 2024). In the learning process, assessments are designed and implemented according to their function, with flexibility in determining assessment techniques and timing to achieve learning objectives effectively. The learning process supports the development of students' competencies and character comprehensively. Assessments are designed to be fair, proportional, valid, and reliable to reflect learning progress and make decisions about the next steps.

Summative assessment in the Merdeka Curriculum is carried out as a basis to ensure the overall achievement of learning objectives. Summative assessment is conducted at the final stage of learning or carried out jointly for two or more learning objectives, in accordance with school policies and teacher considerations (Wulandari et al., 2024). At the elementary and secondary school levels, summative assessment aims to evaluate the achievement of learning objectives or student learning outcomes, which are used to determine promotion to the next grade or completion of a teaching unit by comparing learning outcomes with the criteria for achieving learning objectives. In early childhood education, the purpose of summative assessment is different, where the information obtained is used to understand the student's developmental achievements, not to determine grade promotion or graduation, and also serves as a report on learning achievements and information about the growth and development of students (Maisyaroh et al., 2023). By conducting an assessment at the end of the learning, teachers can have a better understanding of the students they teach and evaluate their achievements related to certain material.

In English Language Teaching and Learning, summative assessment serves as a tool to measure students' learning outcomes and provides evidence of their achievement of the expected competencies outlined in the Merdeka Curriculum. The assessment should include Higher Order Thinking Skills (HOTS) based test items that assess higher cognitive processes, particularly analyzing, evaluating, and creating, as outlined in the revised Bloom's Taxonomy (Bajo, 2025; Anderson & Krathwohl, 2001). Integrating Bloom's taxonomy in language assessment coincides with the current implementation of the Merdeka curriculum in Indonesia, which emphasizes student-centred learning, character development, literacy, numeracy and 21st-century competencies such as critical thinking and creativity (Irwan et al., 2024). Higher level questions require complex skills in applying, analyzing, evaluating, or generating. Test Items at a higher level of taxonomy are generally best suited to encourage students to think more deeply and critically, solve problems, encourage discussion, and challenge students to find information on their own. Lower level problems are memorization, comprehension and lower level applications of taxonomy. Typically, lower-level items are appropriate for

assessing a student's preparation and understanding, diagnosing student strengths and weaknesses, and assessing and/or summarizing content. The more advanced students can understand, the more deeply they can develop critical thinking (Alghifari et al., 2022).

HOTS has been one of the main educational goals considered crucial for students to guide their idea generation. HOTS is needed by the students to respond to the real-world demand in their future life. As a consequence, teaching HOTS to students is also a demand for teachers (Allen & Lee, 2017). For that purpose, teachers need to allocate much time to develop qualified materials in developing students' higher-order thinking. There must be a few efforts to acquire critical thinking inserted in teaching and learning processes. The importance of HOTS skills for students' thinking development requires teachers to be creative and innovative in providing teaching strategies and techniques for the learning process. Teachers can use various techniques or methods to support and facilitate students thinking skills. Questioning strategies can also be acquired as a way of proposing the HOTS concept in the learning process (Wiyaka et al., 2020).

However, there are challenges among English language classroom concerning the distribution of HOTS in the assessment items. (Faradella et al., 2024) discovered that the dominant test items found in language assessment are characteristically LOTS focusing on recall and memorisation. (Ramadhana et al., 2018) also found some weaknesses of the application HOTS based test items. First, teacher-centre is still dominant in learning process. Teacher 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 knowledge of HOTS-based questions.

Although these studies have highlighted the limited integration of HOTS in English language assessment, they primarily focused on the implementation of HOTS or teachers' challenges in developing assessment items. Few studies have systematically examined the cognitive distribution of teacher-developed English summative assessment items under the Merdeka Curriculum using the Revised Bloom's Taxonomy as an analytical framework. Moreover, empirical evidence describing the balance between Lower-Order Thinking Skills (LOTS) and Higher-Order Thinking Skills (HOTS) in teacher-developed summative assessments remains limited, particularly in Indonesian senior high school English classrooms.

Addressing this gap is important because the Merdeka Curriculum explicitly promotes higher-order thinking, yet little is known about whether teacher-developed summative assessments reflect these curriculum expectations. Therefore, this study aims to analyze the cognitive levels represented in English summative assessment test items developed by senior high school English teachers under the Merdeka Curriculum based on the Revised Bloom's Taxonomy. Specifically, this study examines the distribution of cognitive levels from Remembering (C1) to Creating (C6) to determine the extent to which the assessments incorporate LOTS and HOTS.

METHOD

This study used a descriptive study which aimed to analyze the cognitive levels found in English summative assessment test items developed by teachers under the Merdeka Curriculum. This research used test items designed by English teachers in Senior High School Kota Bengkulu. The data for this study were obtained from a final examination prepared by two English teachers at SMA Negeri Kota Bengkulu. There were 30 test items for Grade X, which were constructed by Teacher 1, and 30 test items for XI students, which were constructed by Teacher 2.

Before collecting the data, the researcher requested permission from the teachers to obtain copies of the assessment tests used in their classes. The documents collected consisted of English summative assessment test items administered to students from different grade levels in the 2025/2026 academic year. The analysis was carried out using Anderson and

Krathwohl's Revised Bloom's Taxonomy (2001), which classifies cognitive processes into six levels: Remembering (C1), Understanding (C2), Applying (C3), Analyzing (C4), Evaluating (C5), and Creating (C6). Each test item was examined and classified according to these cognitive process levels to identify the distribution of Lower-Order Thinking Skills (LOTS) and Higher-Order Thinking Skills (HOTS) in the assessments. A checklist was used to facilitate the classification process. To enhance the credibility of the analysis, a co-rater independently classified the test items, and any discrepancies were resolved through discussion until consensus was reached.

FINDINGS AND DISCUSSION

Findings

This section presents the findings of the study concerning the cognitive levels represented in teacher-developed English summative assessment test items under the Merdeka Curriculum. The findings are organized based on the classification of test items according to the Revised Bloom's Taxonomy.

Table 1. Cognitive Levels Analysis of Test Items for X Grade Students

No.	Test Item	C1	C2	C3	C4	C5	C6
1	Choose the correct form of the possessive pronoun.	✓					
2	What is the correct form of the comparative adjective for "big"?	✓					
3	Identify the correct sentence using the past perfect tense.	✓					
4	Choose the correct sentence with the correct use of a conjunction.	✓					
5	What is the correct form of the verb "to do" in the past tense for the pronoun "he"?	✓					
6	Identify the correct sentence with the correct use of an auxiliary verb.	✓					
7	Choose the correct sentence with the correct use of a modal verb.	✓					
8	What is the correct form of the superlative adjective for "good"?	✓					
9	Identify the correct sentence with the correct use of the gerund form of the verb.	✓					
10	Choose the correct sentence with the correct use of the reflexive pronoun.	✓					
11	Which sentence is in the Simple Present Tense?		✓				
12	What is the correct form of the verb "to be" in the Simple Present Tense for "I"?	✓					
13	Which sentence is NOT in the Simple Present Tense?		✓				
14	What is the correct form of the verb "to have" in the Simple Present Tense for "she"?	✓					
15	Which sentence is in the Simple Present Tense negative form?		✓				
16	What is the correct form of the verb "to go" in the Simple Present Tense for "we"?	✓					
17	Which sentence is a question in the Simple Present Tense?		✓				
18	What is the correct form of the verb "to play" in the Simple Present Tense for "they"?	✓					
19	Which sentence is in the Simple Present Tense affirmative form?		✓				
20	What is the correct form of the verb "to do" in the Simple Present Tense for "you"?	✓					
21	Read the following paragraph and identify the main reason why the writer prefers online learning.				✓		
22	Choose the correct verb form to complete the sentence. "Rina her homework before she watched television last night."			✓			
23	After reading the dialogue, determine the relationship between the speakers and explain your answer based on the evidence provided.				✓		
24	Read the text and identify which statement best represents the author's opinion.				✓		
25	A student writes: "Social media always has a negative impact on teenagers." Evaluate this statement and choose the most reasonable argument.					✓	
26	Read the advertisement and judge whether it is effective in attracting customers. Give the best reason for your answer.					✓	
27	Which of the following conclusions is best supported by the information presented in the text?					✓	

Cognitive Levels in English Summative Assessment Items Under the Merdeka Curriculum

28	Evaluate the actions of the main character in the story. Which action was the most appropriate and why?	✓
29	Based on the information in the text, create an alternative ending that is consistent with the storyline.	✓
30	Imagine you are the main character in the story. Write a short response explaining how you would solve the problem differently.	✓

Regarding to table 1, there were thirty test items as summative assessment for grade X which has been analyzed by the researcher. The analysis findings indicate that the analyzed questions cover six cognitive levels, ranging from Remembering (C1) to Creating (C6). However, most of the questions are still at the Remembering (C1) level. Questions at this level generally require students to recall grammar rules and language forms that have been learned, such as the use of possessive pronouns, verb forms, conjunctions, modal verbs, and various types of tenses. In addition, several questions are categorized into the Understanding level (C2) because they require students to comprehend and recognize language usage in a specific context. There is also one question that falls under the Applying level (C3), which is a question that requires students to apply their grammar knowledge to complete sentences with the correct verb forms.

At the HOTS level, several questions are found that fall into the categories of Analyzing (C4), Evaluating (C5), and Creating (C6). Questions at the Analyzing level (C4) ask students to identify reasons, determine relationships between characters in a dialogue, and find the author's opinion in a text. Meanwhile, questions at the Evaluating level (C5) require students to make judgments about a statement, argument, or available information. As for questions at the Creating level (C6), they require students to generate new ideas, such as creating an alternative ending to a story or providing alternative solutions to a problem. In general, the analysis findings show that summative assessment questions are still dominated by lower-order thinking skills (LOTS), although some questions have accommodated higher-order thinking skills (HOTS). To obtain a clearer picture of the distribution of each cognitive level, the frequency and percentage of items at each level are presented in Table 2.

Table 2. Summary of Cognitive Levels Analysis of Test Items for X Grade Students

Cognitive Level	Frequency	Percentage (%)
Remembering (C1)	15	50%
Understanding (C2)	5	17%
Applying (C3)	1	3%
Analyzing (C4)	3	10%
Evaluating (C5)	4	13%
Creating (C6)	2	7%
Total	30	100

Table 2 shows that the test items were dominated by Lower-Order Thinking Skills (LOTS). Specifically, Remembering (C1) accounted for 15 items (50%), Understanding (C2) for 5 items (17%), and Applying (C3) for 1 item (3%). In contrast, Higher-Order Thinking Skills (HOTS) consisted of Analyzing (C4) with 3 items (10%), Evaluating (C5) with 4 items (13%), and Creating (C6) with 2 items (7%). Overall, 21 items (70%) were classified as LOTS, whereas 9 items (30%) were categorized as HOTS. Below are the findings of test item analysis for grade XI students.

Table 3. Cognitive Levels Analysis of Test Items for XI Grade Students

No	Test Item	C1	C2	C3	C4	C5	C6
1	The underlined words show	✓					
2	What does the underlined sentence express?	✓					
3	Is Pete inviting Kate?	✓					
4	Why does Kate accept Pete's invitation? What evidence from the dialogue supports your answer?				✓		
5	Who is the writer of the letter?	✓					

Cognitive Levels in English Summative Assessment Items Under the Merdeka Curriculum

6	How could the students' project help solve the unemployment problem in their community?	✓
7	When will the presentation be held?	✓
8	Explain the characteristics that make the text an invitation letter.	✓
9	What day must Mrs. Colling reply to the letter?	✓
10	The dialogue uses expression of	✓
11	From the underlined sentence, we know that	✓
12	Where is the location of PT CSL Indosejahtera?	✓
13	When will the party be held?	✓
14	Based on the information in the letter, explain how you determine the age of the company.	✓
15	What is the position of Alex Manuhutu?	✓
16	What kind of letter is it?	✓
17	When will Mia get married?	✓
18	If Julia arrives at 09.30 a.m., would she be late? Explain your answer based on the dialogue.	✓
19	"Will you come?" The expression is included into	✓
20	Do you think Mia's dress code is appropriate for a wedding ceremony? Give your reason.	✓
21	Our records that you have been a customer of PJ Party Inc.	✓
22	We would like to thank you for your business inviting you	✓
23	Public admission commence at noon.	✓
24	Which response is the most appropriate? Explain why the other options are less suitable.	✓
25	A: "Do you like the game?" B:	✓
26	A: "Will you pick me up this evening?" B:	✓
27	Analyze how the characteristics of gold contribute to its value as an investment.	✓
28	What is IRA gold's form?	✓
29	Why do people prefer investing in gold during uncertain situations?	✓
30	Evaluate the author's argument that gold is a secure investment. Do you agree? Why?	✓

Table 3 shows the findings of the classification of 30 questions based on cognitive levels in the Revised Bloom's Taxonomy. The analysis results show that the most frequently found questions are at the Remembering (C1) level. Questions at this level generally ask students to recall information contained in the text, recognize expressions, and identify simple facts. In addition, some questions fall into the Understanding (C2) level because they require students to comprehend the meaning of expressions or information conveyed in texts and dialogues. At the Applying (C3) level, there is only one question that asks students to use the available information to determine the correct answer. Meanwhile, several questions have been aimed at higher-order thinking skills. Questions at the Analyzing (C4) level ask students to analyze reasons, explain the characteristics of a text, and connect information contained in the reading. At the Evaluating level (C5), students are asked to provide assessments and reasons regarding a situation or opinion based on the available information. In general, questions that fall into the LOTS category are still more numerous compared to HOTS questions. However, the presence of questions at the C4 and C5 levels indicates that teachers have made efforts to integrate higher-order thinking skills into the English summative assessment.

Table 4. Summary of Cognitive Levels Analysis of Test Items for XI Grade Student

Cognitive Level	Frequency	Percentage (%)
Remembering (C1)	15	50%
Understanding (C2)	5	17%
Applying (C3)	1	3%
Analyzing (C4)	5	17%
Evaluating (C5)	4	13%
Creating (C6)	0	0%
Total	30	100

Table 4 shows that Remembering (C1) was the most dominant cognitive level, with 15 items (50%). This was followed by Understanding (C2) and Analyzing (C4), each with 5 items (17%), Evaluating (C5) with 4 items (13%), and Applying (C3) with 1 item (3%). No items were classified as Creating (C6). These findings indicate that the summative assessment for grade XI was still dominated by lower-order thinking skills (LOTS), although several higher-order thinking skills (HOTS) items were also included. Table 4 shows that Remembering (C1) was the most dominant cognitive level, accounting for 15 items (50%). This was followed by Understanding (C2) with 5 items (17%), Applying (C3) with 1 item (3%), Analyzing (C4) with 5 items (17%), Evaluating (C5) with 4 items (13%), while no items were classified as Creating (C6) (0%). Overall, LOTS represented 70% of the assessment items, whereas HOTS accounted for the remaining 30%. Below the chart of the LOTS and HOTS test items of summative assessment developed by English teachers at SMAN Kota Bengkulu.

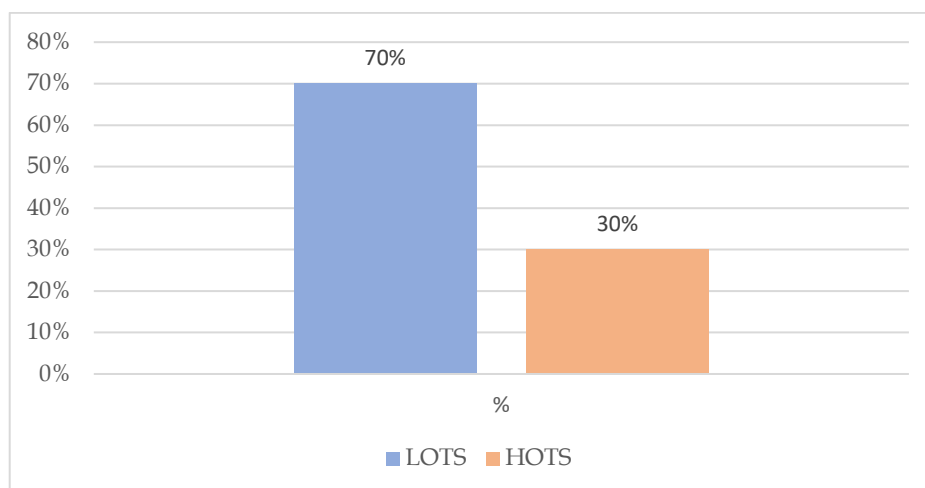


Chart 1. Percentage of LOTS and HOTS in Test Items Summative Assessment

The chart shows that LOTS (Lower-Order Thinking Skills) accounted for 70% of the test items, while HOTS (Higher-Order Thinking Skills) accounted for 30%. This indicates that the assessment was still dominated by questions that measured basic cognitive skills such as remembering, understanding, and applying. Although some items required students to analyze and evaluate information, the proportion of HOTS items was considerably lower than LOTS items.

Discussion

The findings showed that English summative assessment test items developed by teachers are still dominated by lower-order thinking skills (LOTS). Most of the questions are at the Remembering level (C1), while items that measure higher-order thinking skills (HOTS) are only found at the Analyzing (C4), Evaluating (C5) levels, and Creating level (C6) with relatively fewer numbers. Remembering process is the lowest level of cognitive process in education taxonomy. Remembering process is retrieving knowledge that is needed from long-term memory (Anderson & Krathwohl, 2001). The knowledge can be in form of factual knowledge, conceptual knowledge, procedural knowledge, metacognitive, or combination among of those knowledge. The learning condition can be different or same as the situation when the knowledge is taught. Remembering process is very important for meaningful learning and solving some problems that have similarities with the other problems. According to Anderson & Krathwohl (2001), remembering process is divided into two categories. The categories are: (1) Recognizing, Retrieving the information which are needed from long term memory and then comparing with the new information; (2) Recalling, Adopting information which is needed from long term memory as required by assessment.

Analyzing becomes the most dominant level of HOTS found in the test items. HOTS question that measures the dimensions of the C4 thought process is a question that measures

the ability to parse information into parts and determine or explain how the parts are related. The problem with measuring the analysis is when the participant must conclude based on the analysis of parts of the text or stimulus. An example of a question is to find or determine the main ideas, arguments, and assumptions of a text that are not explicitly conveyed; determine or compile evidence that supports and does not support a case description; determine the views of the author of the essay from a certain point of view. Moreover, the principles of HOTS require students to use and manipulate information and ideas by transforming meaning and implication (Parlan & Rahayu, 2021). It reflects when the students combine the facts and their ideas for synthesis, generalization, explanation, and proposed hypothesis.

This findings ties well with previous studies wherein (Santy et al., 2020), revealed that C4 is the most cognitive level which was utilized by English test designers to assess students' English achievement for university students when COVID-19 pandemic. A similar result was also found by (Pratiwi et al., 2020), who concluded that when the tasks of the test were based on critical thinking skills, the students' responses reflected the same manner. (Pratiwi et al., 2020) analyzed the English summative assessment of nine senior high schools in East Java and invited the English test designers to be interviewed. The findings were the school institutions assigned the English teachers to design and develop English assessments based on the HOTS.

This study found that LOTS (70%) is higher than HOTS (30%). This research finding was in line with Bajo, (2025) who also found that a dominance of LOTS-based items (70.6%) over HOTS-based items (29.4%), indicating a gap in the ideal implementation of the Merdeka curriculum, which emphasizes critical thinking, creativity, and independent learning. The primary cause of this discrepancy is often attributed to the lack of teacher training in HTOS-based assessment construction (Rampean et al., 2022). Hence, despite the implementation of the Merdeka curriculum focusing on developing high order skills, the actual assessment practices often fail to accommodate the development of such skills. As a result, most assessments given to the students primarily focusing on the theoretical understanding rather than their ability to apply such knowledge in authentic contexts. This distribution discrepancy has been examined by several studies.

Based on the findings, it can be seen that the English summative assessment test items were still largely focused on lower-order thinking skills. Most of the questions required students to recall information and demonstrate basic understanding of the material. Although several items were designed to measure higher-order thinking skills, particularly at the analyzing and evaluating levels, their number was still limited. This condition suggests that the assessment has not yet fully supported the development of critical and creative thinking skills expected in the Merdeka Curriculum. Therefore, teachers need to provide more opportunities for students to engage with tasks that require analysis, evaluation, and creation so that assessment can better reflect the learning objectives promoted by the curriculum.

CONCLUSIONS

Regarding to the findings, it can be concluded that English summative assessment test items developed by teachers are still dominated by lower-order thinking skills (LOTS). Most of the questions are at the Remembering level (C1), while questions that measure higher-order thinking skills (HOTS) are only found at the Analyzing (C4), Evaluating (C5) levels, and Creating level (C6) with relatively fewer numbers. Overall, HOTS elements have appeared in the analyzed assessments, but their proportion is still lower compared to LOTS. Therefore, English teachers need to increase the use of HOTS based questions in summative assessments so that they are more in line with the objectives of the Independent Curriculum in developing students' critical and higher-order thinking skills. This study has several limitations. First, the analysis was limited to 60 teacher-developed English summative assessment test items collected from two teachers at one senior high school in Kota Bengkulu. Therefore, the findings cannot be generalized to all English summative assessments implemented under the Merdeka Curriculum. Second, this study focused only on the cognitive levels of the test items based on the Revised Bloom's Taxonomy and did not investigate the teachers' rationale for constructing

the questions or examine how the assessments were implemented in classroom practice. Future studies are recommended to involve a larger number of schools and teachers, combine document analysis with interviews or classroom observations, and explore the factors influencing teachers' development of HOTS-based assessment items.

REFERENCES

- Alghifari, B., Saputra, E., Safitri, W., & Maryansyah, Y. (2022). Higher order thinking skills used by English teachers in designing test items at SMAN Argamakmur. *Teaching English and Language Learning English Journal*, 2(2), 115-130. <https://doi.org/10.36085/telle.v2i2.3763>
- Allen, S. H. L., & Lee, E. W. (2017). Involvement of higher order thinking skills within a preparatory course for the Malaysia university English test. *The English Teacher*, XL, 95-112 <https://doi.org/10.15858/engtea.72.4.201712.53>
- Anderson, L. W., & Krathwohl, D. R. (2001). A taxonomy for learning, teaching, and assessing. In *Longman*.
- Ardiansyah, Mawaddah, F. S., & Juanda. (2023). Asesmen dalam kurikulum merdeka belajar. *Jurnal Literasi Dan Pembelajaran Indonesia*, 3(1), 8-13.
- Bajo, S. M. (2025). The alignment of LOTS and HOTS in summative assessments within the Merdeka curriculum. *Journal of Language, Literature, and Teaching*, 7(1), 37-46 <https://doi.org/10.35529/jllte.v7i1.37-46>
- Darmawan, & Winataputra. (2020). *No Title*.
- Faradella, N. E., Wiyaka, W., & Egar, N. (2024). Teachers' strategies and challenges in conducting HOTS-based activities in Merdeka Curriculum era. *Indonesian Journal of Education and Pedagogy*, 1(2), 75-90. <https://doi.org/10.61251/ijoe.v1i2.64>
- Irwan, I., Arnadi, A., & Aslan, A. (2024). Developing critical thinking skills of primary school students through independent curriculum learning. *Indonesian Journal of Education (INJOE)*, 4(3), 788-803.
- Jusuf, H., & Sobari, A. (2022). Pembelajaran paradigma baru kurikulum merdeka pada sekolah dasar. *Jurnal ABDIMAS*, 5(2), 185-194. <https://doi.org/10.31599/jabdimas.v5i2.1360>
- Levy-Feldman, I. (2025). The role of assessment in improving education and promoting educational equity. *Education Sciences*, 15(2), 224 <https://doi.org/10.3390/educsci15020224>
- Maisyaroh, I., Abdullah, M., & Hadi, M. N. (2023). Model asesmen sumatif dengan menggunakan metode library research untuk mata pelajaran pendidikan agama Islam (PAI) pada kurikulum merdeka. *Asatiza: Jurnal Pendidikan*, 4(3), 274-287. <https://doi.org/10.46963/asatiza.v4i3.1063>
- Parlan, P., & Rahayu, S. (2021). Students' higher-order thinking skills (HOTS) in metacognitive learning strategy. *AIP Conference Proceedings*, 2330(1). <https://doi.org/10.1063/5.0043150>
- Pratiwi, N. P. W., Dewi, N. L. P. E. S., Paramartha, A. A. G. Y., Mustika, N., Nurkamto, J., Suparno, S., Anggita, N. L., Utami, P., Eka, N. L. P., Dewi, S., Gede, A. A., Paramartha, Y., Akib, E., & Muhsin, M. A. (2020). Influence of questioning techniques in EFL classes on developing students' critical thinking skills. *Journal of Education Research and Evaluation*, 5(2), 127. <https://doi.org/10.21070/jees.v5i2.752>
- Ramadhana, N. A., Rozimela, Y., & Fitrawati. (2018). Higher order thinking skills-based questions in the test items developed by senior high school English teachers of Padang. *Journal of English Language Teaching*, 7(4), 720-731. <https://doi.org/10.24036/jelt.v7i4.101757>
- Rampean, B. A. O., Rohaeti, E., & Utami, W. P. (2022). Teacher difficulties for developing higher order thinking skills assessment instrument on reaction rate. *Jurnal Pendidikan Kimia Indonesia*, 6(1), 11-19. <https://doi.org/10.23887/jpk.v6i1>
- Santy, N. P. L., Dewi, N. L. P. E. S., & Paramartha, A. A. G. Y. (2020). The quality of teacher-made multiple-choice test used as summative assessment for English subject. *Prasi*, 15(02), 57. <https://doi.org/10.23887/prasi.v15i02.25560>
- Sufyadi, S., Lmabas, Rosdiana, T., Novrika, S., Isyowo, S., Hartini, Y., Primadonna, M.,

- Rochim, N. A. F., & Rizal, M. L. (2021). Panduan Pembelajaran dan Asesmen Jenjang Pendidikan Dasar dan Menengah (SD/MI, SMP/MTs, SMA/SMK/MA). In *Kepala Pusat Asesmen Dan Pembelajaran Badan Penelitian Dan Pengembangan Dan Perbukuan Kementerian Pendidikan, Kebudayaan, Riset, Dan Teknologi*.
- Sutriyanti, N. K., Luwih, I. M., & Suandewa, I. G. L. A. (2024). Implementasi kurikulum merdeka dalam pembelajaran pendidikan agama Hindu dan budi pekerti pada jenjang SMP di kabupaten bangli. *Jurnal Penelitian Agama Hindu*, 8(2), 193–211. <https://doi.org/10.37329/jpah.v8i2.2857>
- Wiyaka, W., Prastikawati, E. F., & Adi, A. P. K. (2020). Higher-order thinking skills (HOTS)-based formative assessment: A proposed model for language learning assessment. *Vision: Journal for Language and Foreign Language Learning*, 9(2), 115–130. <https://doi.org/10.21580/vjv9i25859>
- Wulandari, Y., Yesi, & Jaya, I. (2024). Asesmen pembelajaran dalam kurikulum merdeka. *Elementary: Kajian Teori Dan Hasil Penelitian Pendidikan Sekolah Dasar*, 7(2), 42–47. <https://doi.org/10.31764/elementary.v1i2.549>
- Zahidannisa, & Hikmat, M. H. (2024). Assessment of language learning based on Merdeka curriculum at an Indonesian junior high school. *Proceedings of the 10th International Summit on Science Technology and Humanity (ISETH 2024)*, 2442–2446 <https://doi.org/10.24256/ideas.v12i2.5572>