Eternal: English Teaching Journal
Vol. 16, No. 1, pp. 24-36, February 2025
https://doi.org/10.26877/Eternal.v16i1.892
Received Sep 30, 2024; Revised Nov 18, 2024; Accepted Nov 18, 2024

Analysis of Higher Order Thinking Skills (HOTs) in English Final Examination Questions at Junior High School

[*1]Esti Sapta Lestari[*], [2]Bambang Widi Pratolo
Universitas Ahmad Dahlan
Yogyakarta, Indonesia
2307042010@webmail.uad.ac.id, bambang.pratolo@pbi.uad.ac.id

Abstract. The development of students' Higher Order Thinking Skills (HOTs) is emphasized in the Indonesian educational system. The aim of this research was to examine the HOTs found in the English final exam questions for the 2023–2024 academic year at private schools in Yogyakarta. The test questions were categorized into higher-order thinking skills (HOTs) and lower-order thinking abilities (LOTs) using the updated Bloom's taxonomy framework. The purpose of this study is to determine how much final examination questions at private junior high schools for the academic year 2023–2024 foster higher-order thinking abilities and to pinpoint areas that could use development in order to better support the curriculum objectives. Qualitative content analysis design was used in this research. The data resources of this research were the final exam test questions for the academic year 2023–2024. The revised Bloom's taxonomy framework was used to analyze the data. The findings of this research indicated that the test questions demand more HOTs-type questions since the exams are predominantly LOTs-type questions. Most of the test questions (46.7%) are about applying questions, followed by comprehending (37.8%) and analysis (15.5%). The results revealed that the exam questions lacked HOTs-type questions.

Keywords: Bloom's taxonomy, Higher Order Thinking skills, English test

Introduction

The implementation of the Merdeka Curriculum still provides flexibility for educational units to implement it based on their respective readiness until the 2025/2026 school year (Wahyudin et al., 2024). This means that some schools that haven't implemented the curriculum yet still use the previous curriculum. The 2013 curriculum, recently implemented for Grade IX students in Indonesian junior high schools, promotes Higher Order Thinking Skills (HOTs) among the students. According to Schulz & FitzPatrick (2016), the three highest thinking levels of cognitive talents are called HOTs, include analyzing, evaluating, and producing. Because the goal of education is to improve the caliber of instruction and learning,

[*]*Esti Sapta Lestari, 2307042010@webmail.uad.ac.id*

these critical thinking abilities expand in the area (Driana & Ernawati, 2019). Additionally, it seeks to assist students in becoming more analytical, imaginative, and productive when solving problems (Nguyễn & Nguyễn, 2017). Thus, students are acquainted with both low- and high-level assignments. Thus, the teacher should prioritize these abilities in both teaching and learning in order to meet that goal. According to some experts, asking questions can improve the questioner's skills and encourage students to acquire higher levels of thinking, such as creating in the HOTS framework (Harahap & Astrid, 2021). To some extent, teacher questions, such as Socratic questions, Bloom's Taxonomy, and metacognitive questions, can help students think critically and improve their critical thinking skills (Feng & Wei, 2019). According to some education professionals, asking students HOT questions can help them improve their ability to think (Tyas et al., 2020; Qasrawi & BeniAbdelrahman, 2020). Since they must use more than just their memory to find the solution, they can be more inventive and efficient in their problem-solving.

Higher Order Thinking Skills (HOTs)

Brookhart (2010) defined remembering as the recalling of information. Questions on this level of thinking usually elicit knowledge that students have learned. This level of thinking is essential for understanding other levels of thought. Anderson (2001) defined two different categories of this cognition: recognizing and recalling. The recognizing question often requires students to search their long-term memory for information that relates to the instructions given. Recalling question needs a straightforward recollection of facts.

Anderson (2001) identified comprehension as the ability to understand knowledge and the cognitive processes of interpreting, exemplifying, classifying, summarizing, inferring, comparing, and explaining. Interpreting is a question that requires students to translate one sort of representation into another. When students are required to find a specific example of a topic, this is categorized as exemplifying. Next, classifying questions require students to make an assumption that something belongs to a certain group. A summarizing subcategory includes questions that ask students to abstract from a general subject to a specific issue. Inference occurs when a question prompts students to draw conclusions based on supplied facts. The comparison sub-category asks students to compare two thoughts or things. The last subcategory, explanation, requires students to build a general notion or idea.

According to Anderson (2001), applying competence involves putting it into practice in many ways. This level's questions assess students' comprehension of knowledge in a specific circumstance (Virranmäki et al., 2020). Anderson (2001) identified two sub-categories of this cognition: execution and implementation. The executing category questions demand students to apply previously acquired procedures to a new problem related to the course. Implementing questions encourage students to apply the strategy to unfamiliar tasks.

Krathwohl (2002) stated that analyzing involves gathering information from several sources. Meanwhile, Virranmäki et al. (2020) stated that analysis questions require students to identify and evaluate relevant material. Virranmäki et al. (2020) identified three sub-categories within this cognitive process: attributing, organizing, and differentiating. Differentiating questions require students to identify important and irrelevant material. Organizing questions require students

to identify how components align with the arrangement. Attributing questions challenge students to identify the core concept, value, or aim of the material.

According to Krathwohl (2002) and Brookhart (2010), evaluating information involves assessing quality, efficiency, and consistency. Anderson (2001) identified checking and criticizing as sub-categories of this thought process. Checking assignments might require students to find solutions to problems (Anderson, 2001). Critiquing questions ask students to evaluate the product or information.

Brookhart (2010) defined creating as students' capability for arranging existing knowledge into a new product. According to Brookhart (2010), creating questions required students to create several answers or strategies for issue solving. This thought process includes three sub-categories: creating, planning, and organizing. According to Anderson (2001), these subs connected and inspired kids to be creative. In creating, students were asked to create plausible answers to problems (Krathwohl, 2002). During the planning phase, students explore issue solutions and create action plans. Finally, during the creating session, the students develop solutions to the challenge provided.

Previous research focused on HOTs in question items created by book writers and the government (Putra & Abdullah, 2019; Narwianta et al., 2019; Ilham et al., 2020). However, the study found that HOTs in teacher-created test questions were distinct (Akbariah, 2018; Wisrance & Semiun, 2020). The analysis of Higher Order Thinking Skills or HOTS in English final examination questions at private junior high school in Yogyakarta can be done by examining the proportion and nature of HOTS in the test items. Some studies center on HOTS in any educational setting. For example, a survey of HOTS in the assessing instrument has been conducted in mathematics, showing the proper statistical verification as well as the reliability in test HOTS items (Sausan et al., 2023). A similar example was the reading comprehension test items on a high school book inspection showing the lower distribution of HOTS- item in comparison to Lower Order Thinking Skills, and in the, measurement of HOTS question use on students' writing ability case that explained the significant change of writing descriptive text students when they were given questions of HOTS type (Yudha, 2023) and (Ernawati, 2023).

The studies that were previously mentioned looked into the HOTs question items that the government and book authors created. Still, there isn't study looking at HOTs in English Final Examination Questions at private junior high school especially in Yogyakarta. By examining the English Final Examination Questions at private junior high school Yogyakarta for the academic year 2023–2024, this study seeks to close this gap. Two main questions are at the core of the study of the English Final Examination Questions at a private junior high school in Yogyakarta. In order to determine how well these exam questions measure students' comprehension and proficiency in English, it first looks at the general level and format of the questions. The second aim of the research is to evaluate how much Higher Order Thinking Skills (HOTS) are integrated into these exam questions. This will help determine how well they may push students to think critically and solve problems by going beyond simple retention and comprehension.

Method

In order to investigate the Higher Order Thinking Skills (HOTs) in the English final exam questions, this study used a qualitative content analysis design. The English final exam questions for junior high school students at private junior high school in Yogyakarta for the academic year 2023–2024 served as the study's data source. The information was gathered from the final exam question sets for private junior high school students in Yogyakarta for the 2023–2024 academic year. The revised Bloom's taxonomy framework (producing (Anderson, 2001; Brookhart, 2010). was used for analysis.

Reviewing and evaluating the English final exam questions for private junior high school students in Yogyakarta for the academic year 2023–2024 was part of the data collection procedure. The knowledge component and cognitive process were used to categorize the questions producing (Anderson, 2001; Brookhart, 2010)). The six thinking categories that make up the cognitive process are recall, comprehension, application, analysis, evaluation, and creation. Lower Order Thinking Skills (LOTs) are the first three thinking levels—remembering, comprehending, and applying—while Higher Order Thinking Skills (HOTs) are the final three—analyzing, assessing, and producing (Anderson, 2001; Brookhart, 2010).

The test questions were coded and categorized using the updated Bloom's taxonomy framework as part of the data analysis procedure. The cognitive process and knowledge factor served as the foundation for the coding method (Anderson, 2001; Brookhart, 2010). Creswell (2002) provided an analysis approach that was used to examine the data. The findings portion included a presentation and interpretation of the data analysis results.

Findings and Discussion

There was only one section of the final examination for Muhammadiyah junior high school in Yogyakarta for the 2023–2024 academic year. The section was in the form of multiple-choice questions and there were 45 questions. The Revised Bloom Taxonomy's table of operational verb cognitive process domain provided six cognitive domains. These were classified as Lower-Order Thinking Skills (LOTS) and High Order Thinking Skills (HOTS). The findings include determining the composition of the cognitive levels as well as comparing the cognitive levels of Lower-Order Thinking Skills (LOTS) and High Order Thinking Skills. The table below displays the questions from test items that were categorized using cognitive domain keywords.

Table 1. The classification of the question

No	The Question	Keywords of cognitive process domain	Level
1	What is the relationship between the speakers?	inferring	C2
2	Why did Haris say, 'I wish I could, but something unexpected has come up'?	deconstructing	C4
3	The appropriate expression to complete the text is ….	implementing	C3
4	Where does the conversation probably take place?	inferring	C2
5	From the dialog, we may infer that ….	deconstructing	C4

6	*"I think my competitors showed their best performance".* The underlined word has a similar meaning to ….	interpreting	C2
7	What is the appropriate expression to complete the dialog?	implementing	C3
8	What is the best expression based on the picture?	implementing	C3
9	"Yeah, I agree. That doesn't sound like fun to me." What is Dony's intention by saying the expression?	deconstructing	C4
10	"I don't think so, though I prefer a more relaxing holiday." What does the expression mean?	interpreting	C2
11	Where does the dialog take place?	inferring	C2
12	"Well, I am with you then". What does the expression mean?	interpreting	C2
13	"She prefers a purse, right?" The underlined word refers to ….	inferring	C2
14	Arrange the sentences into a good dialog.	executing	C3
15	What is the product used for?	concluding	C2
16	From the label we know that ….	concluding	C2
17	Why does the writer write the label?	deconstructing	C4
18	What should people do before using the product?	concluding	C2
19	What is the writer's intention of writing the text?	deconstructing	C4
20	What will likely happen if people ignore the storage information of the label?	carrying out	C3
21	Which statements are in line with the text?	deconstructing	C4
22	How much calories if we consume two packs of potato crisps?	impementing	C3
23	When do we add milk?	inferring	C2
24	What will probably happen if we do not whirl the ingredients?	carrying out	C3
25	What does the text tell us about?	inferring	C2
26	What will probably happen if we blend the ingredients for one minute?	carrying out	C3
27	Who will probably be interested in reading the text?	deconstructing	C4
28	How long do you think we need to make Fruitcake cookies? It needs about ….	implementing	C3
29	What does the text tell us about?	comprehending	C2
30	What should you do before pulling down the handle on the right side?	comprehending	C2
31	What cooking utensils do we need most in the second step?	comprehending	C2
32	How many cups of cocoa powder if we want to make pancakes in four portions of the recipe?	implementing	C3
33	The correct words to complete the dialog is ….	implementing	C3
34	The correct words to complete the dialog is …. .	implementing	C3
35	The correct words to complete the dialog is ….	implementing	C3
36	The correct words to complete the dialog is ….	implementing	C3
37	The correct word to complete the text is ….	implementing	C3
38	The correct word to complete the text is …	implementing	C3
39	The correct words to complete the dialog is ….	implementing	C3
40	Putri . . . the flowers when her grandmother . . . her yesterday.	implementing	C3
41	The correct arrangement of the words below is	implementing	C3

42	The correct arrangement of the words above is ….	implementing	C3
43	My grandparents . . . a new coffee shop in their village since 2019.	implementing	C3
44	The text mostly tells the readers about ….	concluding	C2
45	From the text, we know that ….	inferring	C2

Result of Data

The results of document analysis using an updated Bloom's taxonomy framework derived from Anderson and Krathwohl (2001) and Brookhart (2010) showed that the test questions prepared by teachers consisted mostly of LOTs-type questions (57.7%). The remaining 42.3% were HOT-type questions. Table 4.3 shows the percentages of LOT and HOT questions.

Table 2. The frequency of Cognitive Skills on The Question of Test Items

Cognitive Skills	Levels	Item Test	Number
Create		0	0
Evaluate	HOTS	0	0
Analyze		2,5,9,17,19,21,27	7
Apply		3,7,8,14,20,22,24,26,28, 32 – 43	21
Understand	LOTS	1,4,6,10 – 13, 15,16,18,23,29, 30,31, 44,45	17
Remember		0	0

Table 3 The Percentage of Each Cognitive Levels

Criterion	Note	Test Item	Percentage of the total score (%)	Total
C6	HOTS	0	0	15.5 % HOTS
C5	HOTS	0	0	
C4	HOTS	13	7/45 x 100 % = 15.5 %	
C3	LOTS	13	21/45 x 100 % = 46.7 %	84.5 LOTS
C2	LOTS	8	17/45 x 100 % = 37.8 %	
C1	LOTS	0	0	
The total number of items		45	The total number of percentages	100 %

According to Table 4.3, LOTs-type questions represent 84.5%, or 26 test items, of the total data (45 items). They are made up of 17 items (37.8% comprehension) and 21 items (46.7%) of application. LOTs question types are evenly divided from C1 to C3 levels, but in varied percentages. Furthermore, the scope of HOTs-type inquiries encompasses 15.5%, or six questions, which include 17 items (37.8%) of analysis.

Lower-order Thinking Skills in the Test Questions

Table 2 shows that the LOTs questions had a major percentage. This conclusion supports previous researches on the cognitive levels of teacher-created English exams (Akbariah, 2018; Wisrance & Semiun, 2020). In this research, the majority of the test questions developed are classified as understanding and applying, requiring students to respond to the questions by interpreting, exemplifying, categorizing, summarizing, inferring, comparing, explaining, executing, or implementing information. These mental processes are classified as low-order thinking. However, all of these are required components of the cognitive

process before achieving a high level of thinking (Virranmäki et al., 2020). However, even though LOTs questions had the highest rate, questions in the area of application were the most common in this study. Below are several examples of understanding questions taken from the final examination test.

Excerpt 1

```
… dialog…
Where does the dialog take place?
A. In the store
B. In the garden
C. In the canteen
D. In the schoolyard
```

This question is at the understanding level. This question requires the student to comprehend the dialog's content, assess contextual clues within the dialog, and choose the most appropriate setting from the alternatives provided. The cognitive processes involved include interpreting: The students were asked to figure out the sense of the dialog's context, inferring: to draw a logical conclusion about the setting based on the dialog's content, exemplifying: to recognize an instance of a general concept (matching the dialog to a specific setting), and classifying: to determine that the dialog belongs to a specific category of setting. This question is effective because it challenges students to demonstrate knowledge of the dialog's content, use contextual cues to infer information that is not directly stated, and apply their understanding to select the best appropriate response from the possibilities provided.

Bloom's Taxonomy defines the "understanding" level as the process of generating meaning from instructional information, which might be oral, written, or visual in type. This question works effectively since it requires students to understand the dialogue and analyze its context in order to establish the setting. It's worth mentioning that, while this question is mostly at the "understanding" level, it also contains a hint of the "applying" level, as students must use their learning to choose the proper answer from many options. The multiple-choice style (A, B, C, D) is used for "understanding" level questions since it tests comprehension without demanding higher-order thinking abilities such as analysis or assessment.

Excerpt 2 has another question that measures remembering.

```
"I think my competitors showed their best performance".
The underlined word has a similar meaning to ….
A. pals
B. rivals
C. mates
D. friends
```

This question is also on the understanding level. This question asks the student to understand the meaning of the term "competitors" in context, identify synonyms or words with comparable meanings, and choose the best acceptable synonym from the possibilities provided. The cognitive processes involved in this question include interpreting: The students understand the meaning of "competitors" in the given context, then compare: the students relate the meaning of "competitors" to the given options, then classify: the students identify which word belongs to the same category of meaning as "competitors," and finally explain: the students can articulate why the chosen word has a similar meaning. This question is also effective because it requires students to demonstrate an understanding of vocabulary in context, recognize relationships between words

with similar meanings, and apply their knowledge of word meanings to select the best synonym.

The questions in the two examples above require students to respond using basic cognitive processes. According to (Wilson, 2016), students solve two subcategories of understanding using the two sample questions presented previously. The first sub-category involves inferring. The second subcategory is interpreting.

The next category level of thinking, based on Table 3, is application. This sort of thought process leads to 46.7%, among others. There are 21 items out of 45 questions that assess application abilities. This cognition contains two subcategories: execution and implementation (Anderson, 2001). All of the subcategories are necessary for effective teaching and learning (Virranmäki et al., 2020). The following passages provide instances of applying questions drawn from the final assessment examinations.

Excerpt 3

```
The appropriate expression to complete the text is ….
A. Congratulations on your achievement
B. Thanks for being the representative of our city
C. I hope you will be the winner of the competition
D. I wish I was there to support you in the competition
```

This question is at the applying level. This question also requires the students to understand the context of the conversation, comprehend the tone and flow of the dialogue, apply knowledge of appropriate expressions in English conversation, and select the most suitable expression to fit the given context. The cognitive processes involved are executing, implementing, selecting and carrying out. In the executing process, the students apply their knowledge of conversational norms to a familiar task (completing a dialogue). Next, implementing, in this process students use their learned information about expression and context in a specific situation. Then, selection, students choose the most appropriate response from given options and last is carrying out, in this phase students perform the task of completing the dialogue with the correct expression. This question is effective because it requires students to understand the context and tone of the conversation, apply their knowledge of appropriate expressions in English, consider the relationship between the speakers and the situation and choose the most suitable response to maintain the flow of the conversation.

The "applying" level in Bloom's Taxonomy involves carrying out or using a procedure in a given situation. This question fits well because it asks students to apply their knowledge of language and social norms to select the most appropriate expression in a specific context. The multiple-choice format (A, B, C, D) in this case requires students to not just recall information, but to apply their understanding to choose the best fit for the conversation. It's worth noting that while this question primarily fits in the "applying" level, it also incorporates elements of the "understanding" level, as students need to comprehend the dialogue's context and meaning. However, the need to use this understanding to select the appropriate response pushes it into the "applying" category.

Excerpt 4

```
Arrange the sentences into a good dialog.
(1) Rury: “Do you like it? I think it’s too sweet.”
(2) Rury: “I think so. Let’s try this.”
(3) Moana: “Absolutely, I agree with you.”
(4) Moana: “Sure”
```

(5) Moana: "Look at this chocolate tart. It looks so delicious"
A. 5 - 2 - 4 - 1 - 3
B. 5 - 2 - 3 - 1 - 4
C. 5 - 1 - 3 - 2 - 4
D. 5 - 1 - 4 - 2 - 3

This question is also at the applying level. Like the previous question, the question above also requires student to understand and comprehend. This question also asks the students to apply their knowledge of the dialog, especially the structure and the coherence. The cognitive processes involved are executing, implementing, organizing, and sequencing. In the organizing process, students structure the dialog in a logical and coherent manner and in the sequencing process, students put the sentences in the correct order to form a meaningful conversation. The multiple-choice format (A, B, C, D) in this case requires students to not just recall information, but to apply their understanding to choose the best sequence for the conversation.

It is worth noting that while this question primarily fits in the "applying" level, it also incorporates elements of the "analyzing" level, as students need to examine how the different parts of the dialogue relate to each other. However, the primary task of arranging the sentences places it more firmly in the "applying" category.

Higher-order Thinking Skills in the Test Questions

Table 3 shows that HOTs-type questions have 15.5%, or seven items of the 45 questions classified as a higher level of thinking. In this study, all exam questions generated are classified as analyzing. Furthermore, the analytical questions fall within the deconstruction subcategory. It may be considered that the exam questions designed include a small number of HOT questions. The level of higher-order thinking skills of all the questions in this test is analyzing. The questions are categorized into the deconstructing and structuring subcategory. Excerpt 5 is an example of analyzed queries classified as deconstructing.

Excerpt 5

Why did Haris say, 'I wish I could, but something unexpected has come up?'
A. He wanted to join his friend watching the movie.
B. He couldn't join his friend watching the movie.
C. He expressed his hope for his sister's recovery.
D. He did not know something unexpected had come up.

The question above is categorized in "analyzing" level with the the deconstructing sub-category. This question requires the students to examine Haris's statement in context, break down the components of the statement, infer the underlying reasons for Haris's response, interpret the implied meaning behind the words, and select the most appropriate explanation from given options. The cognitive processes involved are differentiating, organizing, attributing, inferring, and interpreting. In the differentiating process, students distinguish between the literal statement and its implications. Next, the organizing process, where students determine how the parts of the statement relate to each other and the context. In the attributing process, students deconstruct to determine Haris's perspective and intentions. Then, students draw a conclusion about Haris's situation based on the given information in thr inferring process. And the last is interpreting where students understand the meaning behind the idiomatic expression "something unexpected has come up".

This question is effective because it requires students to look beyond the surface meaning of the statement, consider the context and implied information,

analyze the speaker's intentions and circumstances, draw conclusions based on limited information and choose the most logical explanation from given options.
In Bloom's Taxonomy, the "analyzing" level involves dividing down content into its constituent parts and establishing how the parts connect to one another as well as to a larger structure or purpose. This question fits well because it asks students to deconstruct Haris's statement, analyze its components, and understand the underlying reason or situation.

The multiple-choice format (A, B, C, D) in this case requires students not just to analyze the statement, but also to evaluate the given options and select the most appropriate one, which adds an element of complexity to the task. It's worth noting that while this question primarily fits in the "analyzing" level, it also incorporates elements of the "understanding" level (as it requires comprehension of the statement's meaning) and the "evaluating" level (as students need to assess the validity of each option).

The final test questions were analyzed using an updated Bloom's taxonomy framework, revealing lower levels of cognitive cognitions. However, higher-order thinking is limited. The test questions reflect a greater degree of comprehension and application. Only a few questions concern the analysis of cognitions. Only seven of the 45 questions assessed higher-order thinking skills. The increasing number of exam items encourages critical thinking. It seeks to provide pupils more alternatives for answering increasingly challenging issues. Nguyễn & Nguyễn (2017) suggested that teaching students how to solve complex issues might improve their learning outcomes. Learning a language can help students enhance their language skills by helping them to gather information, generate ideas, and make interpretations (Gil-Glazer et al., 2019).

The use of multiple-choice questions in this study might explain why there were more LOT questions than HOT ones. The final test is completely composed of multiple-choice questions. According to Heyde & Siebrits (2019), multiple-choice questions are frequently employed to assess lower-order thinking skills. Multiple-choice questions may fail to test higher-level thinking abilities such as analysis, assessment, and creativity (Gareis & Grant, 2015).
Questions with multiple options are incapable of encouraging productive, innovative or creative thought. (Morrison, 2018). Gareis & Grant (2015) pointed out that students' reasoning is limited by the possibilities supplied. Multiple-choice questions can help assess higher-order cognitive abilities. A well-structured multiple-choice question may measure higher ordered thinking abilities in medical education; hence it is a favored alternative (Javaeed, 2018).

Conclusion

The data analysis of exam questions results in the following conclusions. The examinations mostly consist of LOTs-style questions. To successfully achieve the curriculum targets of improving students' higher-order thinking skills, the exams should focus on HOT skill. Second, it's important to discuss the percentage of HOTs questions to include in assessments. LOT skills are necessary for developing higher-order cognitive abilities. It's important to debate and examine the optimal balance of LOTs and HOTs questions in examinations. The study revealed to a significant difference in the categories of thinking skills evaluated by the English final exam questions at a private junior high school in Yogyakarta. Furthermore,

the study showed that 84.5% of the total items were Lower Order Thinking Skills (LOTS) questions, whereas just 15.5% were HOTS questions. This difference highlights the urgent need for educational reform which puts an emphasis on analytical and critical thinking abilities rather than remembering and understanding. This analysis suggests that in order to motivate students to engage with deeper cognitive processes, examination frameworks should include more HOTS. By doing this, teachers may foster an atmosphere in which students not only retain knowledge but also use it to analyze, assess, and produce. This change is crucial for equipping students to solve challenging issues and practice critical thinking in real-life situations.

References

Aditomo, A., Syahril, I., Yuliati, K., Suryani, N., Handayani, F., Muhammad, H., & Zulfikri. (2024). Kurikulum Merdeka. Pusat Kurikulum dan Pembelajaran, Badan Standar, Kurikulum, dan Asesmen Pendidikan, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi.

Akbariah, Y. (2018). Cognitive Skills in Summative Test Items of Secondary Vocational Schools Across Curriculum (School-Based Curriculum and Curriculum 2013) (Doctoral dissertation, Diponegoro University).

Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives: Complete Edition*. New York: Longman.

Brookhart, S. M. (2010). How to Assess Higher-Order Thinking Skills. *Educational Leadership, 67*(6), 34-39.

Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative*. Upper Saddle River, NJ: Prentice Hall.

Driana, E., & Ernawati, E. (2019). Teachers' understanding and practices in assessing higher order thinking skills at primary schools. *Acitya: Journal of Teaching and Education, 1*(2), 110–118. https://doi.org/10.30650/ajte.v1i2.233

Ernawati, E. (2023). Analisis Soal Ujian Matematika Berkategori Higher Order Thinking Skills (HOTS) Siswa SMP dengan Analisis Rasch Model. *Jurnal Didaktika Pendidikan Dasar, 7*(1), 277-294. doi: https://doi.org/10.26811/didaktika. v7i1.1027

Feng, Z., & Wei, W. (2019). Study on Cultivating Students' Critical Thinking Ability Through Higher Order Questioning. *Proceedings of the 4th International Conference on Contemporary Education, Social Sciences and Humanities (ICCESSH 2019).* https://doi.org/10.2991/ICCESSH-19.2019.170.

Freahat, N. M., & Smadi, O. M. (2014). Lower-order and higher-order reading questions in secondary and university level EFL textbooks in Jordan. *Theory and Practice in Language Studies, 4*(9), 1804-1813. https://doi.org/10.4304/tpls.4.9.1804-1813

Gareis, C., & Grant, L.W. (2015). Teacher-Made Assessments: How to Connect Curriculum, Instruction, and Student Learning (2nd ed.). Routledge. https://doi.org/10.4324/9781315773414

Gil-Glazer, Y. A., Walter, O., & Eilam, B. (2019). PhotoLingo—Development and Improvement of Higher-Order Thinking and Language Skills Through

Photographs. *Journal of Education, 199(1)*, 45-56. https://doi.org/10.1177/0022057419843523

Harahap, A., & Astrid, A. (2021). Boosting Classroom Interaction Based on Higher Order Thinking Skills (HOTS) in English Learning for Beginners. *International Journal of Language Education, 5*(1), 477-489. doi: https://doi.org/10.26858/ijole.v5i1.15211

Heyde, V. V. D., & Siebrits, A. (2019). Higher-order e-assessment for physics in the digital age using Sakai. *The Physics Teacher, 57*(1), 32-34. https://doi.org/10.1119/1.5084925.

Ilham, N. W., Jabu, B., & Korompot, C. A. (2020). analysis of higher-order thinking skills (HOTS) items in senior high school english national examination 2019. *ELT Worldwide, 7*(2).

Javaeed, A. (2018). Assessment of Higher Ordered Thinking in Medical Education: Multiple Choice Questions and Modified Essay Questions. *MedEdPublish, 7*. https://doi.org/10.15694/mep.2018.0000128.1.

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice, 41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Morrison, D. (2018). The SQE and creativity: a race to the bottom? *The Law Teacher, 52*(4), 467-477. https://doi.org/10.1080/03069400.2018.1530337

Narwianta, N., Bharati, D. A. L., & Rukmini, D. (2019). The evaluation of higher order thinking skills in English school nationally standardized examination at state senior high school 6 Semarang. *English Education Journal, 9*(3), 316-326. https://doi.org/10.15294/eej.v9i3.30937

Nguyễn, T. M. T., & Nguyễn, T. T. L. (2017). Influence of explicit higher-order thinking skills instruction on students' learning of linguistics. *Thinking Skills and Creativity, 26*, 113-127. https://doi.org/10.1016/j.tsc.2017.10.004

Putra, T. K., & Abdullah, D. F. (2019). Higher-order thinking skill (HOTS) questions in English national examination in Indonesia. *The Journal of Educational Development, 7*(3), 178-185. https://doi.org/10.15294/jed.v7i1.30878

Qasrawi, R., & BeniAbdelrahman, A. (2020). The Higher and Lower-Order Thinking Skills (HOTS and LOTS) in Unlock English Textbooks (1st and 2nd Editions) Based on Bloom's Taxonomy: An Analysis Study. *International Online Journal of Education and Teaching, 7*(3), 744-758.

Sausan, K., Rohmatillah, R., & Rini, Y. P. (2023). HOTS Based on Revised Bloom's Taxonomy: Analysis in English Examination Test Items Used in High School Level. *English Education: Jurnal Tadris Bahasa Inggris, 16*(1), 102-118. doi: https://doi.org/10.24042/ee-jtbiv16i1.16103

Schulz, H. W., & FitzPatrick, B. (2016). Teachers' Understandings of Critical and Higher Order Thinking and What This Means for Their Teaching and Assessments. *Alberta Journal of Educational Research, 62*(1), 61–86. https://doi.org/10.11575/ajer.v62i1.56168

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.

Virranmäki, E., Valta-Hulkkonen, K., & Pellikka, A. (2020). Geography tests in the Finnish matriculation examination in paper and digital forms–an analysis

of questions based on revised Bloom's taxonomy. *Studies in Educational Evaluation, 66*, 100896. https://doi.org/10.1016/j.stueduc.2020.100896

Wilson, L. O. (2016). Anderson and Krathwohl Bloom's taxonomy revised understanding the new version of Bloom's taxonomy. *The Second Principle, 1*(1), 1-8.

Wisrance, M. W., & Semiun, T. T. (2020). LOTs and HOTs of teacher-made test in junior high school level in Kefamenanu. *JEE (Journal of English Education), 6*(2), 62-76. https://doi.org/10.30606/jee.v6i2.574

Yudha, R. P. (2023). Higher Order Thinking Skills (HOTS) Test Instrument: Validity and Reliability Analysis with The Rasch Model. *EduMa: Mathematics education learning and teaching, 12*(1), 21-38. doi: https://doi.org/10.24235/eduma.v12i1.9468