

Characteristics of the Summative Assessment Questions in Arabic Language: Analytical Study Based on Cognitive Levels

Salma Laksmi Benedik^{1*}, Bakri Mohammad Bakheet², Nur Hasan³

¹ Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia; salmabenedik@gmail.com

² Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia; bakribakheet@gmail.com

³ Universitas Islam Malang, Indonesia; nur.hasan@unisma.ac.id

ABSTRACT

This study aims to identify the characteristics of the end-of-year summative assessment questions (ASAT) in Arabic language subjects based on cognitive levels and question construction. A qualitative approach with a descriptive method was employed in this study. The population and sample in this study consisted of the final-year summative assessment (ASAT) test scripts for Arabic language subjects in grade 4 at MI Manarul Islam. Data collection techniques in this study included observation, interviews, and documentation, which were analyzed using the Miles and Huberman analysis techniques. The results indicate that 35 questions were tested, covering cognitive levels L1, L2, and L3. Level L1 includes cognitive levels C1 and C2. Level L2 includes cognitive level C3. Level L3 includes cognitive levels C4 and C6 only. The material tested consists of two chapters: "Afrādul Ustratī" and "fil baiti". The questions encompass six types: multiple-choice questions, complex multiple-choice questions, true-false questions, matching questions, short-answer questions, and essay questions. The language used in these questions is simple and communicative Arabic.

Keywords: Assessment, Cognitive Levels, Summative

How to Cite Benedik, S. L., Bakheet, B. M., & Hasan, N. (2025). Characteristics of the Summative Assessment Questions in Arabic Language: Analytical Study Based on Cognitive Levels. *Jurnal Al Burhan*, 5(2), 288–301. <https://doi.org/10.58988/jab.v5i2.513>

INTRODUCTION

Arabic is a compulsory subject in schools, especially in madrasah schools under the auspices of the Ministry of Religion (KEMENAG) and schools under the aegis of the Ministry of Education and Culture (KEMENDIKBUD), which are Islamic in genre. This is confirmed by the Decree of the Director General of Islamic Education No. 3211 of 2022, which specifically regulates several provisions regarding Arabic language learning in schools, including the determination of learning outcomes (CP) that must be achieved during the learning process. This includes learning objectives and learning targets, commonly referred to as learning achievement elements (Decree of the Director General of Islamic Education Number 3211 of 2022, tt-a). Therefore, these elements will be incorporated into the learning process.

One of the significant aspects in the learning process is learning evaluation. Where learning evaluation is carried out to measure the abilities of students after they have studied the science concentration. In addition, learning evaluation is also used by teachers as data to inform decisions about improving and enhancing classroom activities. According to Law No. 20 of 2003, article 58 paragraphs 1 and 2, it states that evaluation is a method of monitoring and assessing the process and results of teaching activities carried out by institutions in a continuous, regular, comprehensive, transparent, and systematic manner. In the educational process, several key aspects must be understood to comprehend the learning process. If, in teaching, student elements are used as input, school and class learning are viewed as a process, and graduate competency is the outcome, then assessment occurs both at the beginning of learning, during learning, and at the end of learning (Rani Tri Damayanti et al., 2024). Therefore, evaluation plays a crucial role in the learning process.

Several terms related to evaluation often appear, including evaluation, assessment, measurement, and test. In essence, a test is a tool that contains tasks or questions that students must complete to measure a specific aspect of knowledge or understanding. Measurement is the process of assigning numerical values to the learning outcome process based on particular rules, enabling informed decisions about the learning process and its outcomes. Assessment is an activity that involves interpreting or describing the results of measurements of the learning outcome process in the form of scores, converting them into values based on specific procedures, and making informed decisions. Evaluation is a process of determining values or providing considerations regarding values in the learning process and learning outcomes, which are carried out through learning measurement and assessment activities (Pitaloka et al., 2021). Based on several terms explained, this research primarily focuses on discussing assessment tools or tests within the context of assessment.

One of the functions of learning evaluation is to determine the number of learning outcomes of students. Typically used when advancing to the next class or determining a student's graduation status. In this case, one type of learning outcome evaluation is summative evaluation. Summative evaluation is a type of evaluation used to assess the progress of multiple students or groups of students, which can also be used for reporting to parents, determining class promotion, and other purposes (Makbul et al., 2022).

Assessment is divided into four types, namely formative assessment, summative assessment, placement assessment, and diagnostic test. Formative assessment is an evaluation that aims to determine student achievement in the learning process. Summative assessment is an evaluation that assesses the value of students who have completed their learning. Typically, this assessment occurs at

the end of the semester or at the end of the year. Placement assessment is an evaluation used to understand the condition of students and their personalities during the teaching and learning process. A diagnostic assessment is an evaluation conducted to determine the need for guidance or counseling services. This diagnostic test is typically performed in cases where students experience difficulties, such as the reason why students often skip math lessons (Handayani et al., 2022). Of the several types, this study focuses on summative assessment.

Summative evaluation in educational institutions is carried out at least twice a year. In the independent curriculum, summative evaluation is referred to as SAS (Summative Assessment at the End of Semester) and SAT (Summative Assessment at the End of Year). From year to year, including elementary schools and their equivalent, the questions tested in the summative evaluation are coordinated by the Teacher Working Group (KKG) or the Madrasah Working Group (KKM). However, it has received an appeal, so that this will not be done again. The appeal is stated in a circular from the Director of Curriculum, Facilities, Institutions, and Students (KSJK), which was directly signed by M. Ishom Yusqi as the Director of KSJK. The circular is addressed to the Regional Offices of the Ministry of Religion of Provinces throughout Indonesia. The contents of the circular are a prohibition on coordinating the preparation and duplication of exam questions in madrasahs (Khoeron, 2021).

MI Manarul Islam Malang is one of the madrasahs in Malang City, which is equivalent to an Elementary School, where the school implements personal writing of final summative questions. The writing is done by the subject teachers themselves, including the Arabic language subject. According to the results of an interview with one of the teachers, the writing of the questions was done because of the inconsistency between the questions coordinated by the local KKG or KKM with the material achieved by the students. This is because each madrasah has a different background of student abilities and facilities. Therefore, this madrasah does the writing independently.

This study is considered significant since, after initial observation, the questions initially appear difficult due to their use of Arabic. According to field data, many children aged seven to 12 years, or equivalent to elementary school age, still cannot read Arabic writing properly and correctly. However, the questions prepared by this madrasah as a whole use Arabic. Are the questions in line with what the students have learned? Therefore, the study carries the characteristics of end-of-year summative questions in the Arabic subject of grade 4 MI Manarul Islam.

Moh conducted previous research. Bunny Andaru Bahy and Primasti Nur Yusrin Hidayanti (2023) entitled "Analysis of Characteristics of Arabic Language Questions at Elementary School Level," which resulted in findings in the form of

two difficult questions, 16 moderate questions, and 16 multiple-choice questions with easy questions. The level of difficulty of the essay questions is one difficult category, eight moderate categories, and one easy category. In the long essay, none are classified as difficult, three as mild, and two as easy; the multiple-choice question item discrimination power yielded two very good category questions, nine good questions, seven sufficient category questions, and two poor questions. Furthermore, in the short essay questions, one very good item was obtained, one long essay item, three very good category items, and no questions in the good category, none in the sufficient category, and two items in the poor category. What distinguishes this study from previous studies is its unique context. This study focuses not only on the characteristics of the question construction. But it also discusses the cognitive level of the question items used.

In more detail, this study focuses on understanding the characteristics of test questions, especially in Arabic subjects. With the mastery of test question characteristics, it can serve as a reference material for improving or developing the quality of learning and skills of students. Specifically, the purpose of this study is to determine the characteristics of end-of-year summative questions (SAT) in Arabic subjects for grade 4 of MI Manarul Islam Malang. In detail, the research has the following objectives: (1) to determine the characteristics of the cognitive level of end-of-year summative questions (SAT) in Arabic subjects for grade 4 based on Bloom's Taxonomy, (2) to determine the construction of end-of-year summative questions (SAT) in Arabic subjects for grade 4 based on the rules of writing questions. Thus, the study's results are expected to contribute to the development of learning evaluation questions and the implementation of the learning process, particularly in Arabic subjects, thereby improving the Arabic learning system and enhancing students' Arabic skills.

METHOD

The approach in this research employs a qualitative method because the data obtained are in qualitative form, namely, written words. (Textbook of Qualitative Research Methodology, tt). The type of research used is qualitative descriptive, which aims to describe the state of something at the time of the research. (Bahy & Hidayanti, 2023). The object used in this study is the End-of-Year Summative Assessment (ASAT) question script for the Arabic language subject in grade 4. The population and sample used in this study are the end-of-year summative questions (SAT) for the Arabic language subject in grade 4 at MI Manarul Islam Malang for the 2023/2024 academic year, which were created by the Arabic language subject teacher for grade 4 at MI Manarul Islam.

The data sources used in this study are three things, namely observation, interviews, and documentation. Observations can be made regarding the

characteristics of each question item and the cognitive level of the questions being tested. The data obtained from the interview provides the reasons for choosing the form of questions and the process of selecting materials in compiling questions, which will be conducted with the Arabic language subject teacher. The documentation data were obtained from the Arabic language question script for grade 4 MI Manarul Islam Malang, along with its accompanying equipment, including grids and answer keys. However, from this data, it cannot be directly concluded that a finding has been made, as a technical process of data analysis is required to conclude the study (Miles et al., 2014).

This research was analyzed using the interactive data analysis method, a technique developed by Miles and Huberman. (Qomaruddin & Sa'diyah, 2024). The stages (see Figure 1) are explained in detail as follows: Data reduction involves selecting, focusing, simplifying, and transforming raw data obtained from on-site written notes. At this stage, questions are chosen, and those with cognitive levels 1, 2, and 3 are included. Additionally, the types of questions used in the final summative questions are also recorded. The presentation of data can be in the form of descriptive narratives, tables, or charts. Miles and Huberman stated that the most widely used presentation of qualitative data is narrative text. (Saleh et al., tt, 2017). In this study, the presentation of data employs a combination of images and narratives to convey the data obtained. Drawing conclusions where conclusions are drawn from several data obtained, and an explanation of the characteristics of the end-of-year summative questions (SAT) in the Arabic subject for class 4 of MI Manarul Islam Malang.

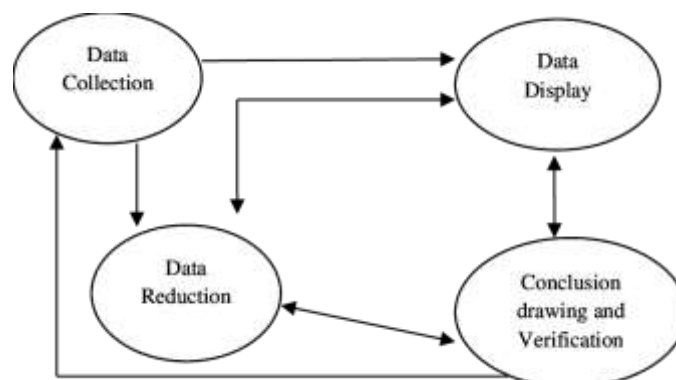


Figure 1. Stages of Data Analysis of Miles and Huberman's Theory

RESULTS AND DISCUSSION

RESULTS

Characteristics of Questions Based on Cognitive Level

The number of questions tested is 35, with details of the questions that fall into L1, specifically 22 questions. L1 questions consist of C1 and C2. From the questions that fall into C1, the operational verb used is "write" and there is one

question, namely question number 31, the operational verb "mention" as many as 12 questions, namely questions number 1, 4, 5, 11, 12, 23, 25, 26, 27, 28, 29, and 34, and the operational verb "pair" as many as five questions, namely questions number 3, 10, 18, 19, and 20. The questions that fall into C2 use the operational verb to interpret as many as four questions, namely questions number 21, 30, 32, and 35.

The following questions are included in the cognitive level group C1 (see Figure 2). The questions ask students to identify the number of family members mentioned in a provided text. This is clearly written in the sentence *أُسْرَتِي مِنْ ... أَنْفَارٍ* (tatakawwanu usratī min ... anfarin) or if translated, "my family consists of ... people". In this question, the answer choices are presented in Arabic numerals.



Figure 2. ASAT Arabic Questions Class 4 No. 1

Below is an example of a question that falls into group C2 (see image 3). The question falls into cognitive level C2 because it asks students to "interpret" a word from Arabic into Indonesian. The image shows that the question asks students to interpret the word "أُسْرَتِي" (usratī) into Indonesian with the wording "... معنى أسرتي" (ma'na usratī ...), which means "the meaning of the word usratī is ..." and the answer is written in the column provided.

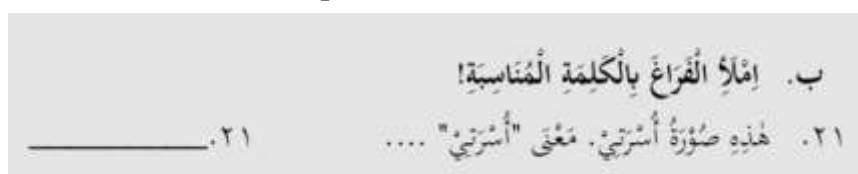


Figure 3. ASAT Arabic Questions Class 4 No. 21

L2 is one level higher than L1. This year-end summative assessment consists of six questions for L2. At the cognitive level, according to Bloom's Taxonomy, only level C3, with the operational verb "to determine," is included in L2. The questions included in this cognitive level group are questions 6, 7, 13, 14, 15, and 22. The number of questions included in the L2 cognitive level group is the smallest compared to the L1 and L3 cognitive level groups.

The following is an example of a question that falls into cognitive level C3 (see Figure 4). In this question, students are asked to determine two predetermined answers. According to the question maker, the correct way to answer this type of question is by placing a check mark (✓) in the provided column. The question shows the command to determine two correct answers if preceded by a noun isim ḍamir “هي” (hiya), which means “she (female) as the key to answering the question.

٦. هذه أختي، اسمها صافيا. هي (اختر الإجابتين الصحيحتين)

☐ خياط ☐ مدرس ☐ خطاط ☐ مدرّسة

Figure 4. ASAT Arabic Questions Class 4 No. 6

L3 is the highest cognitive level, also known as higher-order thinking skills (HOTS) type questions. (Religious Studies & Lifestyle, 2022). Bloom's Taxonomy Cognitive Levels included in L3 are C4, C5, and C6. However, the observation results revealed that only cognitive levels C4 (analyzing) and C6 (creating) were used. The operational verb used in C4 is "analyzing," which is found in questions 2, 8, 9, 16, 17, and 24. The operational verb used in C6 is "composing," which is found only in question 33.

The following is an example of a question that falls into cognitive level C4 (see Figure 5), which asks students to analyze the truth of a sentence or statement. From the question, there are three sub-statements in one question, in which students are asked to analyze the sentence. The analysis is carried out based on grammar or tarāḳib to verify the truth in terms of sentence structure, the correct use of isim isyārah (pronouns), the use of correct harakat, and the proper use of words such as muzakkar (words indicating men) and mu'annas (words indicating women). If the sentence is declared correct, then give a check mark (✓) in the column “الصحيحة” (ash-shahihah). If the sentence is declared wrong, then check the column “الخاطئة” (al-khathī'ah).

٢. ضع علامة (✓) في العبارة الصحيحة أو الخاطئة!

الخاطئة	الصحيحة	العبارة
		هذه أمي، اسمها سيدة حميدة.
		هذا أبي، اسمه سيد حسن.
		هذا أخي الكبير، اسمه قاتع.

Figure 5. ASAT Arabic Questions Class 4 No. 2

The following questions are included in Cognitive Level C6 (see Figure 6). The operational verb used is "menyusyun". In its application to the question, several words in Arabic are not in the correct order, and students are asked to

rearrange these words to form a complete sentence. It can be seen from the question that there are several words, namely اسمه (ismuhu), أحمد (ahmadu), هذا (hāzyes), جدي (jaddī), طبيب (thabībun), and هو (huwa). The keyword is the word هذا (hāzyes). Then the students arrange the words into complete sentences correctly.

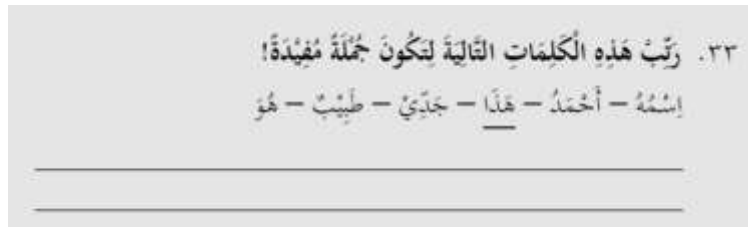


Figure 6. ASAT Arabic Questions Class 4 No. 33

From the several presentations above, in preparing this test, all five levels of Bloom's taxonomy are included. Namely C1, C2, C3, C4, and C6. This was also confirmed by the author of the questions, who is also the teacher of Arabic language subjects, that in preparing these questions, it is recommended to create questions that cover all cognitive levels. The aim is to assess the students' understanding of the material taught by the teacher. Additionally, it aims to enhance the quality of students' learning.

In this case, the explanation of the characteristics in terms of question construction will be divided into three parts: material, construction, and language use. The explanation is as follows:

1. Material

In the final year summative question (SAT) for grade 4 MI Manarul Islam Malang, from the observations obtained, the material tested in this question is only two chapters from the total material that should be studied in the even semester, namely three materials (see table 1), but what is tested is chapter 4 with the title "أفراد الأسرة" (afrādul usratī) and chapter 5 with the title "في البيت" (fil baiti). According to the interviews obtained, this question is deliberately testing only two chapters because of the form of optimization in studying the material in a shorter time.

Additionally, the questions align with the question indicators (see Figure 7). The questions test the material in Chapter 4. This is proven by the presence of the word "أمي" (ummī) in the questions. In this chapter, there are about 11 questions that test Chapter 4. The questions are questions number 1, 3, 4, 5, 10, 21, 22, 23, 25, 31, and 32, which occupy the L1 cognitive level; questions number 6, 7, and 22 occupy the L2 cognitive level; and questions number 2, 8, 9, 24, and 33 occupy the L3 cognitive level.

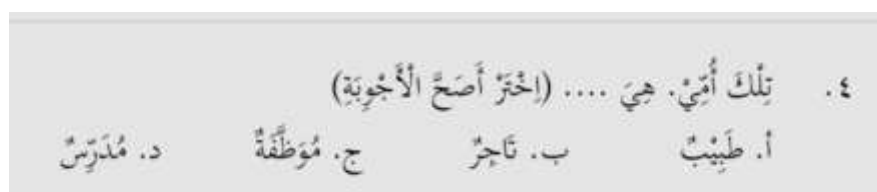
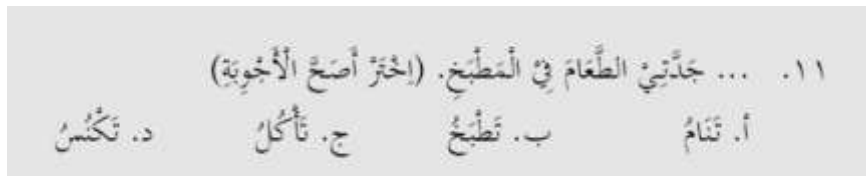


Figure 7. ASAT Arabic Questions Class 4 No. 4

The following are examples of questions that test Chapter 5. This is proven by the sentence "في المطبخ" (fil mathbakhi), which means "in the kitchen" and the kitchen is located inside the house. The questions that test chapter 5 are numbers 11, 12, 18, 19, 20, 26, 27, 28, 29, 30, 34, and 35, which occupy the L1 cognitive level; questions number 13, 14, and 15, which occupy the L2 cognitive level; and questions number 16 and 17, which occupy the L3 cognitive level.

**Figure 8.** ASAT Arabic Exam Question for Grade 4 No. 11

The learning outcomes for the questions compiled here include four things written on the question grid sheet, namely: (1) Reading and understanding very simple discourse in the form of written text or visual text about "family members" with the sentence pattern ismul isyārah – ismul mufrad to understand the explicit information from very simple discourse; (2) Produce and present vocabulary that is grammatically and contextually appropriate to the topic of "family members" with the sentence pattern ismul isyārah – ismul mufrad to express ideas in writing and orally; (3) Reading and understanding very simple discourse in the form of written text or visual text about "at home" with the sentence pattern ismul isyārah – ismul mufrad to understand the explicit information from very simple discourse; and (4) Produce and present vocabulary that is grammatically and contextually appropriate to the topic "at home" with the sentence pattern ismul isyārah – ismul mufrad to express ideas in writing and orally.

Based on the written learning outcomes, the compilation of these questions encompasses the elements of reading, listening, writing, and presenting. This is because the learning outcomes prioritize reading and understanding a discourse, as well as presenting vocabulary. The two learning outcomes listed in the final year assessment (ASAT) question grid for Grade 4 MI Manarul Islam Malang are part of the elements of reading, listening, writing, and presenting.

Table 1. Mapping of Questions on Lesson Material

Material	Cognitive Level	Question Number
Chapter 4 "afrādul usrati"	L1	1, 3, 4, 5, 10, 21, 22, 23, 25, 31, and 32
	L2	6, 7, and 22
	L3	2, 8, 9, 24, and 33
Chapter 5 "fil baiti"	L1	11, 12, 18, 19, 20, 26, 27, 28, 30, 34, and 35
	L2	13, 14, and 15
	L3	16, and 17

2. Construction

In terms of question construction, the end-of-year summative assessment (ASAT) questions for the Arabic language for grade 4 MI Manarul Islam Malang consist of several types. There are six types of questions used in the End-of-Year Summative Assessment, namely (1) multiple choice, there are five questions, (2) complex multiple choice, there are five questions, (3) matching, there are five questions, (4) true-false, there are five questions, (5) short answer, there are 10 questions, and (6) description/essay, there are five questions. So, the number of questions tested is 35. Of all the questions, some use a stimulus and others do not. The stimulus used can be in the form of images, simple text, or a combination of text and images.

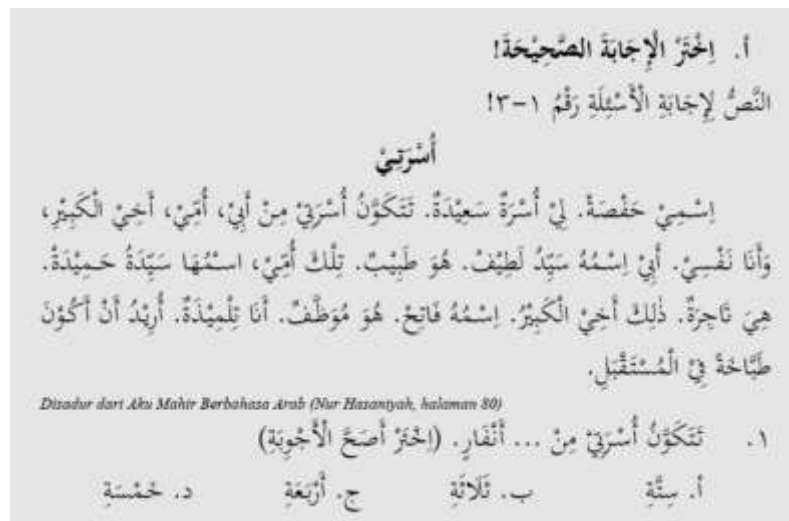


Figure 9. Arabic ASAT Questions Class 4 No. 1

The following is an example of a question that uses a stimulus (see Image 9). Based on the image, it can be seen that the construction of the question utilizes a stimulus in the form of a simple script titled "Azarita". Additionally, the questions include clear command words (instructors use these at the beginning of the reading and in the questions). The instructions at the start of the reading are "page 1-3" (an-nash li ijabatil as'ilati roqmu 1-3), which means "text to answer questions 1-3." In this case, it is already clear that the instructions are that the text is used to answer questions 1-3. Additionally, instructions are available for each number. In question number 1, there is an instruction that reads, "اختر أصح الأجوبة" (ikhtar ashohhal ajwibah), which means "choose the most correct answer!". The answer choices provided are homogeneous, consisting only of numbers. However, they are not sorted from largest to smallest or vice versa. This question also does not contain the statement "all of the answer choices above are wrong" or "all of the answers above are correct."

3. Language

In terms of language usage, the ASAT Arabic questions for Grade 4 at MI Manarul Islam Malang are taught in Arabic as the language of instruction. Starting from question number 1 to question number 35, all use Arabic. The sentences used in the questions are simple sentences and easy for students to understand. The number of words used in each question ranges from 4-45 words.

DISCUSSION

According to the Directorate General of Teachers and Education Personnel of the Ministry of Education and Culture, there are three cognitive levels in compiling questions. These levels are categorized as follows: L1, which encompasses knowledge and understanding; L2, which includes application questions; and L3, which involves reasoning questions. The purpose of categorizing cognitive levels into three parts is to facilitate the interpretation of Operational Verbs (KKO) at the cognitive level of Bloom's Taxonomy theory. L1 includes the cognitive levels of knowing (C1) and understanding (C2), L2 includes C3 (applying), L3 includes C4 (analyzing), C5 (evaluating), and C6 (creating) (W. Setiawati et al., nd). Based on this finding, it can be inferred that the final year summative questions (SAT) in the Arabic subject of grade 4 at MI Manarul Islam Malang cover all existing cognitive levels, namely L1, L2, and L3. In addition, it is also more specifically emphasized on the use of operational verbs (KKO) that align with the levels in Bloom's taxonomy. Although the number of questions between L1, L2, and L3 is not balanced, it is still comprised of a fairly large number of questions. This indicates that the question writer pays close attention to the type of cognitive level in the questions as a means of measuring students' high-level understanding abilities. Additionally, it is an effort to enhance the quality of student learning.

Learning materials encompass the knowledge, skills, and attitudes that students must acquire to meet the predetermined competency standards. (Djumingin et al., nd, 2022). When writing questions, determining the material is the first step that must be taken. By selecting the material, the question compiler can evaluate the question indicators that can be written in the question grid. (Rahmi Nur Fauziah et al., 2020). Based on the existing findings, it can be inferred that the end-of-year summative assessment (ASAT) questions assess the materials that students have studied. However, the questions tested do not cover all the materials identified by the Directorate General of Islamic Education in 2022, which include three learning materials. But only two learning materials are tested. This meets the provisions for compiling existing evaluation questions.

This question tests only two skill domains: reading skills, also known as *mahārah al-qirā'ah*, and writing skills, also known as *mahārah al-kitābah*. This

demonstrates that both skills can be objectively measured through writing. This is in accordance with the elements outlined in the Decree of the Director General of Islamic Education Number 3211 of 2022, which divides the elements of Arabic language learning into four categories: listening, speaking, reading, and writing. (Decree of the Director General of Islamic Education Number 3211 of 2022, tt-b).

The form of written test questions is classified into two forms, namely objective and non-objective. Objective forms include multiple choice questions, true-false questions, and matching questions. Non-objective questions include short-answer questions and essay questions. In compiling questions, it is necessary to pay attention to several rules in terms of construction, namely (1) the formulation of the question sentence must use question words that lead to detailed answers such as why, describe, explain, compare, connect; (2) make clear instructions on how to work on the questions; (3) make scoring guidelines; and (4) other things that accompany the questions such as tables, figures, graphs, maps, must be presented clearly (S. Setiawati & Lapasau, 2022).

Based on the existing findings, the final year summative assessment (ASAT) questions for the Arabic language in grade 4, as seen in MI Manarul Islam, should focus on the types of questions available. Starting from objective and non-objective questions used in the questions. Furthermore, the questions also have clear instructions and are equipped with images, text, or both as stimuli to help students understand the questions well. This indicates that question writers pay close attention to several aspects of question construction rules.

In terms of language, the rules for writing short answer questions are: (1) using direct questions or imperative sentences; (2) questions must be clear to get a short answer; (3) the length of sentences that students must answer for each question must be relatively the same; (4) avoiding sentences or phrases that are taken directly from textbooks; and (5) creating scoring guidelines. (S. Setiawati & Lapasau, 2022). According to the results obtained, the questions do use Arabic as a whole. Starting from questions, answers, to instructions for completing all tasks, the system uses Arabic. However, the questions use communicative, simple language and have been studied by students. In addition, the length of the sentences used is also relatively the same. The stimuli taught also use simple language. This indicates that, in compiling the questions, the author also pays close attention to existing language rules.

CONCLUSION

The questions tested in the final-year summative assessment (ASAT) for grade 4 at MI Manarul Islam Malang totaled 35 questions. The questions were divided into three sections: multiple-choice, short-answer, and essay groups. The questions tested reached the three cognitive levels, namely L1, L2, and L3. In L1,

which covers cognitive level C1, there were 18 questions, and in cognitive level C2, there were four questions. In level L2, which covers C3, there were six questions. In level L3, which covers C4, there were six questions, and in C6, there was one question. In terms of material, this question tests only two chapters, namely, chapter 4 with the title "The Greatest of All" and chapter 5 with the title "In the name of Allah", and learning outcomes are taken from the elements of reading, listening, and writing-presenting. In terms of question construction, the final-year summative assessment (ASAT) questions for grade 4 MI Manarul Islam Malang utilize six types of questions: multiple-choice, complex multiple-choice, true-false, matching, short-answer, and description. Of the 35 questions tested, some use stimuli and others do not. The stimuli used are in the form of simple text, images, and some are a combination of text and images. These questions are written with clear command word guidance. In questions that have answer choices, the answer choices are homogeneous and do not contain the statement "all of the above answers are wrong" or "all of the above answers are right". The language used in the questions is Arabic from the questions to the answers. Although the questions and answer choices are written in Arabic, they still employ communicative and understandable language.

REFERENCES

- Bahy, M. B. A., & Hidayanti, P. N. Y. (2023). ANALISIS KARAKTERISTIK BUTIR SOAL BAHASA ARAB TINGKAT SEKOLAH DASAR. ... P, 01(01).
BukuAjar Metodologi Penelitian Kualitatif. (t.t.).
Djumingin, D. H. S., Hum, M., Juanda, D., Pd, S., Hum, M., & Tamsir, N. (t.t.).
PENGEMBANGAN MATERI PEMBELAJARAN BAHASA INDONESIA.
Handayani, F., Maharani, R. A., & Fitria, Y. (2022). Penilaian dan Jenis Tes yang Dibuak Oleh Guru di Tingkat Sekolah Dasar. *Journal of Basic Education*, 6(1), 726-737. <https://doi.org/10.31004/basicedu.v6i1.2033>
Keputusan Direktur Jendral Pendidikan Islam Nomor 3211 Tahun 2022. (t.t.-a).
Keputusan Direktur Jendral Pendidikan Islam Nomor 3211 Tahun 2022. (t.t.-b).
Khoeron, Moh. (2021, November 19). Kemenag dan Forum Guru dilarang Koordinir Penyusunan-Penggandaan Soal Ujian Madrasah. *Kementrian Agama Republik Indonesia*. <https://kemenag.go.id/nasional/kemenag-dan-forum-guru-dilarang-koordinir-penyusunan-penggandaan-soal-ujian-madrasah-tmd6hw>
Makbul, M., S, D. S., & Ahmad, L. O. I. (2022). Pengembangan Evaluasi Formatif dan Sumatif: Development of Formative and Summative Evaluation. *HAWARI: Jurnal Pendidikan Agama dan Keagamaan Islam*, 3(1). <https://doi.org/10.35706/hw.v3i1.6788>

- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (Edition 3). Sage.
- Muhassanah, N., & Hayati, A. (2022). WORKSHOP PENYUSUNAN SOAL HOTS (HIGHER ORDER THINKING SKILLS) BERDASARKAN KISI-KISI UJIAN NASIONAL MATEMATIKA SD. *ADI Pengabdian Kepada Masyarakat*, 2(2), 56–63. <https://doi.org/10.34306/adimas.v2i2.636>
- Pitaloka, D. L., Dimyati, D., & Purwanta, E. (2021). Peran Guru dalam Menanamkan Nilai Toleransi pada Anak Usia Dini di Indonesia. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini*, 5(2), 1696–1705. <https://doi.org/10.31004/obsesi.v5i2.972>
- Qomaruddin, Q., & Sa'diyah, H. (2024). Kajian Teoritis tentang Teknik Analisis Data dalam Penelitian Kualitatif: Perspektif Spradley, Miles dan Huberman. *Journal of Management, Accounting, and Administration*, 1(2), 77–84. <https://doi.org/10.52620/jomaa.v1i2.93>
- Rahmi Nur Fauziah, I., Syihabudin, S., & Sopian, A. (2020). ANALISIS KUALITAS TES BAHASA ARAB BERBASIS HIGHER ORDER THINKING SKILL (HOTS). *لساننا (LISANUNA): Jurnal Ilmu Bahasa Arab dan Pembelajarannya*, 10(1), 45. <https://doi.org/10.22373/lis.v10i1.7805>
- Rani Tri Damayanti, Aisyah Nur Kholifatun Jannah, & Anisa Fadilah Hidayati. (2024). PERAN EVALUASI PEMBELAJARAN PADA MATA PELAJARAN PENDIDIKAN AGAMA ISLAM KELAS 12 DI SMK NEGERI 1 KARANGANYAR. *Journal of Student Research*, 2(1), 205–209. <https://doi.org/10.55606/jsr.v2i1.2598>
- Saleh, S., Pd, S., & Pd, M. (t.t.). *ANALISIS DATA KUALITATIF*.
- Setiawati, S., & Lapasau, M. (2022). *ASPEK BAHASA DAN KONSTRUKSI BUTIR SOAL EVALUASI PADA BUKU TEMATIK KELAS III SEKOLAH DASAR*. 1.
- Setiawati, W., Pd, M., Asmira, O., Ariyana, Y., Bestary, R., Pd, M., & Pudjiastuti, D. A. (t.t.). *Berorientasi Higher Order Thinking Skills*.



© 2025 by the authors. Submitted for possible open-access publication under the terms and conditions of the Creative Commons Attribution (CC-BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).