Lughawiyah Journal of Arabic Education and Linguistics, 7(2) - Dec 2025 113-125

Lughawiyah
Journal of Arabic Education and Linguistics
E-ISSN 2715 - 8101
P-ISSN 2715 - 8098

Development of Arabic Assessment for Tarkib and Qira'ah Based on Higher Order Thinking Skill

Ade Sri Wahyuni [1], Rahmawati [2], Yelfi Dewi [3], Fitri Alrasi [4], Bambang [5], Meliza Budiarti [6]

[1] *Universitas Islam Negeri Imam Bonjol Padang, Indonesia*
[2] *Universitas Islam Negeri Imam Bonjol Padang, Indonesia*
[3] *Universitas Islam Negeri Sjech M. Djamil Djambek Bukittinggi, Indonesia*
[4] *Universitas Muhammadiyah Sumatera Barat, Indonesia*
[5] *Universitas Muhammadiyah Sumatera Barat, Indonesia*
[6] *Universiti Pendidikan Sultan Idris, Malaysia*

Corresponding Author: Ade Sri Wahyuni, E-mail; ade.sri@uinib.ac.id

Received: July 19, 2025	Revised: August 22, 2025	Accepted: Nov 25, 2025	Online: Dec 30, 2025

ABSTRACT

In Senior High School, Arabic language assessments mostly focus on memorizing grammar rules (*tarkib*) and recalling information in reading comprehension (*Qira'ah*). This assessment does not improve students' abilities to analyze complex texts, evaluate arguments, or apply grammatical knowledge in new contexts. In Eleventh Grade, the Tasawwuq chapter of the Arabic textbook includes only 21% Higher-Order Thinking Skills questions, while 79% focus on Lower-Order Thinking Skills, such as remembering and understanding. This study aims to develop an Arabic assessment focused on tarkib and qiraah skills. The ADDIE model was used in five steps: analysis, design, development, implementation, and evaluation. This study used a qualitative approach to collect data, including documantation and questionnaires. Data analysis used descriptive statistics. In the analysis stage, LOTS questions were identified more than HOTS questions. For the student needs analysis, a questionnaire was distributed to 60 respondents, resulting in an average score of 4.19. The design stage involved developing an outline and question matrix, creating HOTS-based questions, formulating a scoring system, and establishing a rubric. During the development stage, questions were constructed in accordance with HOTS criteria, incorporating eight specific HOTS indicators. Implementation was conducted at Madrasah Aliyah Negeri 4 Agam in the eleventh grade. Evaluation results demonstrated that the Arabic assessment for Tarkib and Qira'ah, based on Higher Order Thinking Skills, effectively enhanced the critical thinking skills of eleventh-grade students at the state madrasah aliyah, as indicated by a significance value (2-tailed) of less than 0.05.

Keywords: *Arabic Assessment, Arabic Language, Higher Order Thinking Skill*

Journal Homepage https://ejournal.uinmybatusangkar.ac.id/ojs/index.php/lughawiyah

How to cite: Wahyuni, S. A., Rahmawati, Rahmawati., Dewi, Y., Alrasi, F., Bambang, Bambang., & Budiarti, M. (2025). Development of Arabic Assessment for Tarkib and Qira'ah Based on Higher Order Thinking Skill. *Lughawiyah Journal of Arabic Education and Linguistics*, 7(2), 113-125. http://dx.doi.org/10.31958/lughawiyah.v7i2.16052

Published by: Universitas Islam Negeri Mahmud Yunus Batusangkar, Indonesia

INTRODUCTION

21st-century Arabic language education demands Higher-Order Thinking Skills (HOTS) to develop students' analytical and applicative competencies beyond rote memorization (Wahyuni et al., 2023). The Ministry of Education and Culture explains that HOTS is an instrument designed to measure the ability to transfer, process, and apply information (Gozali et al., 2021); identify relationships between pieces of information (Pisriwati et al., 2024); use information to solve problems (Isnaeni et al., 2021); and critically examine ideas or information (Fanani et al., 2018). Based on Bloom's Taxonomy, the levels of ability included in HOTS are analyzing, evaluating, and creating (Tuela & Palar, 2022). If the ability level is still at the level of remembering, understanding, and applying, then the ability is still in the Lower Order Thinking Skill (LOTS) category.

However, in the Indonesian context, it is still difficult to apply HOTS in assessments (Setyorini, 2025). In senior high schools, assessments of the four language skills predominantly employ LOTS questions that focus on remembering and understanding (Fayyoumi et al., 2024). The questions are still about who, when, and where; all the cognitive domains are still in the remembering stage (Farros, 2024). It has not yet reached the analyzing, evaluating, and creating stage (Muradi et al., 2020). Program for International Student Assessment (PISA) results confirm that Indonesian students struggle with HOTS assessments, performing below international averages in critical analysis and problem-solving (Riswanda, 2018). The HOTS context should enable students to think critically by presenting solutions (Ragab et al., 2024).

One of the factors for the 21st century is the ability to think in higher order. According Arrias et al., (2019), in the "Learning Handbook Oriented to Higher Level Thinking Skills", critical thinking is a complex thinking process of explaining, inferring, constructing representations, analyzing, and communicating by engaging in basic mental activities. Tyas & Naibaho, (2021) state that HOTS helps prepare students to compete in the era of globalization 5.0. Teachers should guide students to think critically and analytically, foster collaboration, and develop problem-solving skills (Cojorn & Sonsupap, 2024). The primary goal of learning by Higher Order Thinking Skills is to strengthen students' problem-solving abilities and improve their global competitiveness.

Students learning Arabic need to develop four main skills are listening (*istima'*), speaking (*kalam*), reading (*Qira'ah*), and writing (*kitabah*). Arabic teaching standards require that each skill be specified alongside its corresponding percentage. Evaluation is conducted through tests that include both objective and essay questions. However, despite these standards, Arabic textbooks predominantly emphasize LOTS tests, as demonstrated by the eleventh-grade textbook used in Madrasah Aliyah Negeri. In the Arabic textbook by Sari & Amrullah (2020), the "*tasawwuq*" as sub-material is assessed through essays. Fanani et al., (2018) explain that the assessment of HOTS can be done in several ways, including 1) multiple choice, matching, ranked items by selection, 2) short answers and essays to generalization, and 3) giving reasons.

Yasmara et al., 2023 researched Qira'ah exercises in Ministry of Religious Affairs textbooks and identified a few questions on analyzing, evaluating, and creating. Among the 16 question types in the Qiraah test, only 6 are based on HOTS. These 6 question types fall into

three categories: determining the title of a paragraph, answering questions based on the text, and constructing sentences based on pictures. HOTS-based Arabic learning aims to develop individuals who can think critically about global issues and competition. The results show that HOTS-based Arabic learning has a significant effect on students' critical thinking skills. The experimental class scored an average of 68.13, while the control class averaged 54.69. The significance value (2-tailed) is 0.00, which is less than 0.05. This means that HOTS-based learning improves students' critical thinking in Arabic lessons (Mustofa et al., 2022). Research by Ilmiani & Delima (2021), the research used a qualitative approach that categorized HOTS-based Arabic reading skills learning activities into three levels: level of analysis, level of evaluation, and level of creation. Also, research from Haniefa (2022), the assessment of HOTS in the four Arabic language skills can be conducted by aligning assessment formats with the dimensions of the HOTS thinking process: creating, evaluating, and analyzing. These dimensions should then be specified through indicators of competency achievement that elaborate on the basic competencies. Thus, previous studies consistently affirm that integrating HOTS into Arabic learning and assessment not only increases the cognitive level of test items but also significantly enhances students' critical thinking and higher-order language competencies in real learning contexts.

An analysis of eleventh-grade Arabic textbooks, specifically the Tasawwuq chapter, reveals that only 21% of assessment items are HOTS-based, while 79% focus on LOTS. Among the evaluation aspects of the Qira'ah and Tarkib material, the book uses questions with Operational Verbs to mention, explain, and distinguish, which focus on LOTS. The results of the initial analysis did not identify question levels at C5 as evaluation and at C6 as creating. This significant disparity in assessment practices underpins the current study, which aims to develop assessments focused on HOTS. Previous research on HOTS in Arabic language learning has focused on identifying HOTS items within one Arabic skill. However, those studies have not integrated HOTS into both tarkib and Qira'ah. This gap highlights a disconnect between curricular demands for HOTS-based assessment and the limited availability of standardized Arabic HOTS assessment.

The present study addresses this gap by developing and validating an Arabic assessment instrument for tarkib and Qira'ah that is explicitly based on HOTS. The assessment is constructed by deriving indicators from Bloom's higher-order levels; analyzing, evaluating, and creating, and also items aligned with these indicators, and empirically evaluating item quality using validity and reliability test. The final product is practically applicable in classroom Arabic assessment contexts.

RESEARCH METHODOLOGY

This research used a field research design with a Research and Development (R&D) approach. Sugiyono explained that R&D was used to develop a specific product and test its effectiveness (Baihaqi et al., 2025). This study employed the ADDIE development model proposed by Dick and Carey, comprising five stages: analysis, design, development, implementation, and evaluation (Faizah et al., 2024). Data were collected using a qualitative approach (Sarosa, 2021), including documentation from eleventh-grade Arabic textbooks and

questionnaires. The respondents included three teachers and sixty students. The ADDIE model was applied to obtain valid and practical test items.

The analysis stage involves examining subject matter aligned with the essential competencies. At the design stage is a planning step in preparing HOTS-based Arabic questions. Meanwhile, the development stage involves validating the test question before the trial. The feasibility assessment of the questions utilized two categories of validators: content and language. The implementation stage is carried out on eleventh-grade students at Madrasah Aliyah Negeri (MAN) 4 Agam. All of the questions were analyzed to assess validity, difficulty, and discrimination. The Evaluation process includes test validity, test normality, test reliability, and the t-test. In the final stage of evaluation, it produces a feasible and valid test for evaluating Arabic learning. Thus, the products resulting from this development research are HOTS-based Arabic assessment questions for tarkib and qiraah, suitable for the eleventh grade of MAN 4 Agam. The final product was a set of HOTS-based Arabic assessment for tarkib and Qira'ah ready to be used in classroom.

RESULT AND DISCUSSION

Sari & Amrullah (2020) wrote an Arabic textbook, which was published by the Directorate of KSKK Madrasah as part of the implementation of KMA Number 183 of 2019 concerning the PAI and Arabic curriculum in Madrasah. This book presents essays and multiple-choice questions as a means of learning evaluation. It consists of six chapters and includes two exam questions: odd-semester and even-semester exams. Each sub-chapter contains exercises designed to support student's language skills: listening (Istima'), reading (Qira'ah), writing (Kitabah), grammar (Tarkib), and speaking (Kalam). Finally, each chapter concludes with a summary.

Analysis

In Tasawwuq material, questions for Tarkib and Qira'ah are presented in essay form. The question analysis is as follows:

Table 1. Question Analysis on Chapter *Tasawwuq*

Skill	Form of Question	KKO	Level	Number of Questions
Tarkib	عيّن العدد والمعدود فيما يلي!	Understanding	C2	5 Questions
	املأ الفراغات التالية بالعدد والمعدود المناسبين!	Understanding	C2	5 Questions
	هات جملا كما في المثال!	Understanding	C2	8 Questions
Qira'ah	اقرأ الجمل التالية ثم قل (صحيح) إذا كانت الجملة صحيحة أو (خطأ) إذا كانت خاطئة ثم صحّح الخطأ	Analyzing	C4	5 Questions
	هل تباع في المجمّع التجاريّ الحاجات اليومية؟	Remembering	C1	5 Questions
	لماذا يختار الناس أن يشتروا حاجاتهم من المجمّع التّجاريّ؟	Understanding	C2	
	هل في المجمّع التّجاريّ تخفيضات كثيرة؟	Remembering	C1	
	أذكر خصائص المجمّع التجاريّ!	Remembering	C1	
	ما الفرق بين السوق التقليدي والمجمّع التّجاريّ؟	Analyzing	C4	

There are 28 items consisting of Tarkib and Qira'ah with a percentage of each Cognitive Domain. There were no HOTS questions at the C5 and C6 levels. In the Qira'ah skill section, six questions were identified at the C4 level, while in the Tarkib section, none were found. Most of the questions presented still fall under Lower Order Thinking Skills (LOTS), accounting for 79%, while HOTS questions made up only 21%. Therefore, it can be concluded that the questions related to Tarkib and Qira'ah in the Arabic textbook for Grade XI Madrasah Aliyah are not HOTS-based. The results of the percentage analysis are described in the table below:

Table 2. Cognitive Level Percentage

Cognitive Level	C1	C2	C3	C4	C5	C6	Total
Number of Questions	3	19	0	6	0	0	28
Percentage	11%	68%	0%	21%	0%	0%	100%

The percentage results in the table above show that the number of LOTS-based questions is higher than that of HOTS-based questions. The table below explains the difference between LOTS and HOTS-based questions.

Table 3. Percentage of HOTS and LOTS

Cognitive	**Level**	**Total**
HOTS	C1 - C3	21%
LOTS	C4 - C6	79%

Based on the table above presents the distribution of cognitive levels in an Arabic language assessment, categorizing items into Higher Order Thinking Skills (HOTS) and Lower Order Thinking Skills (LOTS) by the Bloom's taxonomy. HOTS items, corresponding to levels C4 (analyzing), C5 (evaluating), and C6 (creating), constitute 79% of the total, while LOTS items at C1 (remembering), C2 (understanding), and C3 (applying) account for the remaining 21%. Therefore, the result describe that it is the Arabic assessment especially for Tarkib and Qira'ah need for the development of the test.

The distribution of questions shows that the assessment of Tasawwuq material still focuses on developing LOTS rather than HOTS. With 79% of questions at levels C1 and C2, students are primarily encouraged to remember and understand information, rather than to analyze, evaluate, or create, as required in HOTS-based learning. This observation is consistent with previous studies by Hilmi et al., (2022) demonstrating that many language assessment instruments, including those for foreign languages, are still dominated by LOTS and do not adequately support the development of students' critical and creative thinking skills (Cojorn & Sonsupap, 2024). Consequently, this analysis underscores the need to develop and revise Tarkib and Qira'ah questions to achieve a more balanced representation of C4–C6 levels. That assessments align with the 2013 Curriculum and the 21st-century learning paradigm, both of which emphasize HOTS as a primary focus of Arabic language instruction in Madrasah Aliyah.

Design

The first stage in designing questions is analyzing the Basic Competencies and Competency Achievement Indicators. Educators must examine the "Kata Kerja Operasional" (KKO) or Operational Verbs used in the textbook, as each cognitive level within the KKO

framework represents a different level of complexity. Therefore, educators need to identify indicators that align appropriately with the student's grade level.

HOTS-based cognitive level analysis is two items on the "Kompetensi Dasar" (KD) or Basic Competencies. However, it is possible if the Basic Competencies that is not in the C4 to C6 category is formulated into HOTS-based questions. At the LOTS level, Competency Achievement Indicators are developed to be HOTS-based. All can be adjusted to the needs of students and educators' learning materials analysis.

The second stage is to formulate a question grid. There are four stages in formulating HOTS-based question grids: 1) selection of Basic Competencies that will be used as HOTS questions, 2) determination of material based on Basic Competencies, 3) formulation of question indicators, and 5) formulation of cognitive levels (Wicaksono, 2021). The third stage in making questions is based on the grids that have been formulated. That means that the question design can be aligned with the formulation of the matrix question. Using Operational Verbs with matrix question will make educators easier to formulate question instructions based on HOTS-based cognitive levels. The development of Arabic assessment questions in the Tasawwuq Chapter was divided into more varied essay questions with different cognitive levels: four questions at level C4, four at level C5, and two at level C6. The fourth stage is about scoring. The Tarkib and Qira'ah questions have ten questions, with details of five questions for Tarkib and five questions for Qira'ah. The questions are formulated as essays and multiple-choice questions. Each cognitive level has a different score, so a rubric was formulated to minimize subjective assessment.

The rubric is designed based on the questions formulated. Scoring is done with specific formulas that are detailed in the rubric (Wicaksono, 2021). The multiple-choice rubric presents two cognitive levels, while other assessments are based on the essay rubric. The assessment will not be subjective if the rubric is formulated. Each answer will be valuable if the answer meets the criteria that have been formulated in the assessment rubric.

Table 4. The Design of Question Based HOTS

Basic Competencies	Essential Materials	Question Indicator	Item	Cognitive Level	Dimension of Knowledge
Understand the social functions, text structures, and linguistic elements (sounds, words, and meanings) of texts related to themes: التسوق (في السوق التقليدي، في السوبرماركيت involving speech acts describing the shape, scale,	*Adad* (1-10 *muannats* and *Mudzakkar*)	**Selecting** the *'adad* that belongs to the *mudzakkar* category	Multiple Choice	C4	Factual
	Adad and *Ma'dud* (*muannats* and *Mudzakkar*)	**Analyzing** sentences that contain *'adad* and *ma'dud* in a sentence	Essay	C4	Conseptual
	Adad mudzakkar and *mu'anntas* number 11-20 and 21-99	**Comparing** two *'adad* whose *ma'dud mansub* in numbers 11-99	Essay	C5	Conseptual

Basic Competencies	Essential Materials	Question Indicator	Item	Cognitive Level	Dimension of Knowledge
properties, and characteristics of an object by paying attention to the form, meaning, and function of the grammatical structure عدد ألف ومليون ومليار وبليون	*Adad mudzakkar* and *mu'anntas* number 100-900 and 901-999 or 101-110	**Categorize** *ma'dud* in the form of *mansub* and *jer* in *'adad* 11-99 and 100-1000	Multiple Choice	C6	Factual
		Summarizing the rules of *'adad* and *ma'dud*	Essay	C5	Factual
Analyzing ideas from Arabic texts related to the theme: التسوق) في السوق التقليدي، في السووبرماركيت) by paying attention to the form, meaning, and function of the grammatical structure عدد ألف ومليون ومليار وبليون	النص عن: المجمع التجاري	**Analyzing** the main idea in the third paragraph of the text المجمع التجاريّ	Essay	C4	Conseptual
		Proving with data that modern markets are better than traditional markets	Essay	C5	Factual
		Comparing two components such as the difference between traditional and modern markets	Essay	C4	Conseptual
		Rearrange sentence fragments into a paragraph	Essay	C6	Procedural
		Making the conclusion of المجمع التجاري	Essay	C5	Metacognitive

Developing an Arabic assessment based on HOTS, that is focused on analyzing, evaluating, and creating. In the concept Taksonomi Bloom for knowledge competency, there are three aspects as conceptual, procedural, and metakognitive (Kartini et al., 2022). Studies in Arabic education reveal that when tasks like Qira'ah incorporate these elements such as inferring main ideas, judging an author's position, contrasting arguments, or suggesting new titles and conclusions, students achieve greater depth in comprehension and more active critical involvement with the material. Further research on HOTS-driven Arabic instruction confirms that these approaches yield markedly better learning results and enhanced critical thinking skills over conventional lower-level questions focused on rote memory and surface-level grasp (Verawati et al., 2022).

Development

The design of HOTS-based questions in Arabic for grade XI of Madrasah Aliyah is by formulating question indicators in the form of: 1) Selecting, 2) Analyze, 3) Compare, 4) Categorize, 5) Conclude, 6) Prove, 7) Reorganize, and 8) Conclude.

Table 5. The Question Items

Cognitive Level	Materials	Questions
C4	Adad (1 until 10 *muannats* and *Mudzakkar*)	**Selecting** *'adad* that is in the *mudzakkar* category
C4		**Analyzing** sentences that contain *'adad* and *ma'dud* in a sentence
C5	*Adad mudzakkar* and *mu'anntas* number 11 – 20 and 21 – 99	**Comparing** two *'adad* whose *ma'dud mansub* in numbers 11 - 99
C6	*Adad mudzakkar* and *mu'anntas* number 100 – 900 and 901 – 999 or 101 – 110	**Categorize** *ma'dud* in the form of *mansub* and *jer* in *'adad* 11 - 99 and 100 - 1000
C4	النص عن: المجمع التجاري	**Analyzing** the main idea in the third paragraph of the text المجمع التجاريّ
C5		**Proving** with data that modern markets are better than traditional markets
C4		**Comparing** two components such as the difference between traditional and modern markets
C6		**Rearrange** sentence fragments into a paragraph
C5		**Making** the conclusion of المجمع التجاري

The table documents the distribution of Higher Order Thinking Skills (HOTS) items across cognitive levels C4 (analyzing), C5 (evaluating), and C6 (creating) in an Arabic language assessment, detailing source materials and corresponding question categories. It lists nine entries, with C4 appearing most frequently (four instances), followed by C5 (two instances) and C6 (one instance), alongside additional descriptive tasks aligned to these levels. This structure illustrates the alignment between instructional materials and HOTS-oriented questions in Arabic. This design supports HOTS dominance (79% as per prior data), encouraging students to engage in critical thinking in Arabic rather than just memorization. This approach fosters the development of advanced analytical and creative skills among learners of Arabic.

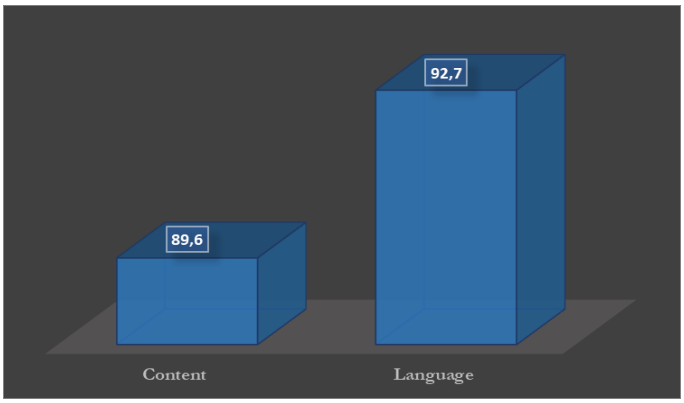

Figure 1. The Percentage of Validity Test

The results of the evaluation expert validation concluded that the average percentage of overall aspect assessment was 93.98%. This shows that the HOTS question instrument perfectly interprets material and language aspects. The instrument design shows that the HOTS questions used a variety of advanced cognitive processes, such as analyzing, comparing, categorizing, concluding, proving, and reorganizing information in the context of the *'adad* material and the text المجمع التجاريّ. The questions are distributed across cognitive levels C4, C5, and C6, matching indicators like selecting, analyzing, comparing, categorizing, concluding, proving, reorganizing, and summarizing. This match shows that learning objectives, instructional materials, and assessment formats are well integrated. As a result, students are encouraged not only to recall arithmetic rules (*adad*) and text content, but also to think critically and creatively about the material. Therefore, served as a model for developing higher-order thinking assessments in Arabic language classes for 11th-grade students at Madrasah Aliyah.

Implementation

At this stage, product trials have been designed and validated by validators. The implementation of HOTS-based Arabic language assessment in qira'ah and tarkib skills was conducted through a limited-scale trial at Madrasah Aliyah Negeri 4 Agam. For the purposes of the trial, purposive sampling was used to select 30 eleventh-grade students. During the implementation stage in the eleventh grade at MAN 4 Agam, HOTS questions consisting of essay and multiple-choice formats, were administered to each student individually to objectively assess their higher-order thinking abilities. The teacher explained the question format, time allocation, and answer sheet procedures, distributed the question packets, and established a focused classroom environment to ensure students could concentrate.

Evaluation

From the nine questions that were developed, the validity results were obtained by the table below:

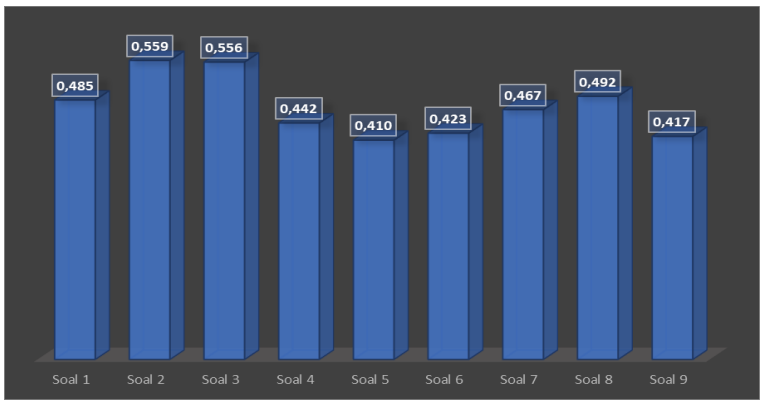

Figure 2. The Result of Validity Item Test

To determine the effectiveness of the development of Arabic HOTS questions, the researchers conducted a product test at MAN 4 Agam with two tests: pre-test and post-test.

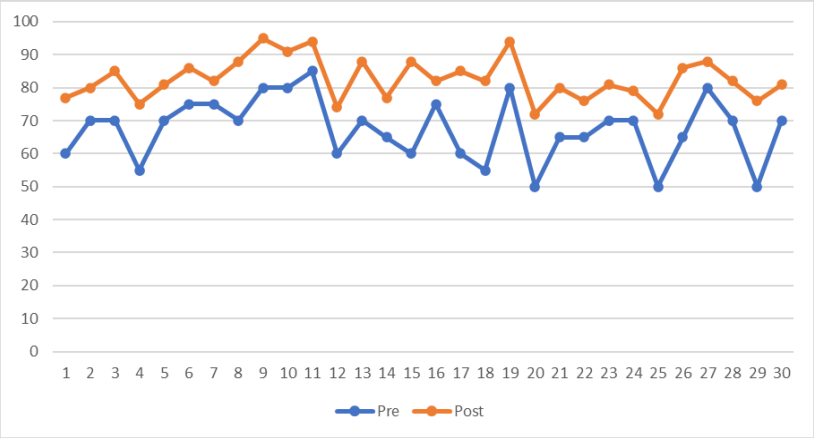

Figure 3. The Result of Pre-test and Post-test

The results of the pre-test before being treated and the post-test after being treated increased on average from 67,33 to 82,57. The average pretest score was 60–75, while the average posttest score was 75–90, indicating improvement in learning ability after the treatment. It can be concluded that HOTS questions improve student learning outcomes after being treated in the experimental class.

Table 6. The Result of Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Pre-Test - Post-Test	-15.233	6.168	1.126	-17.537	-12.930	-13.527	29	.000

The results of the Paired Samples t-Test indicate a statistically significant difference between pre-test and post-test scores ($n = 30$, $df = 29$, $t = -13.527$, $p < 0.001$). The mean post-test score increased by 15.233 points compared to the pre-test, with a standard deviation of 6.168 and a 95% confidence interval for the mean difference ranging from -17.537 to -12.930. These results provide robust empirical evidence that Arabic language assessment utilizing higher-order thinking skills in Qira'ah and Tarkib maharah enhances students' critical thinking competencies.

These findings align with previous studies showing that integrating HOTS into Arabic language learning and assessment remains limited to a small number of questions but has a strong cognitive impact. Yasmara et al., (2023) found that of the 16 types of qira'ah questions in the Ministry of Religious Affairs' book, only six were truly HOTS-based, and all of them focused on determining paragraph titles, answering questions based on the text, and constructing sentences from pictures. This pattern shows that the dimensions of analyzing, evaluating, and creating have begun to be accommodated, but have not yet become dominant in the design of existing qira'ah exercises. Meanwhile, Mustofa et al., (2022) proved that when HOTS principles were explicitly applied in Arabic language learning, there was a significant increase in students' critical thinking skills; this was evident from the difference in the average scores of the experimental class (68.13) and the control class (54.69) with a significance value of $0.00 < 0.05$.

Ilmiani & Delima (2021) emphasized that HOTS-based qira'ah learning needs to be designed at three levels are analysis, evaluation, and creation. While Haniefa (2022) emphasizes that HOTS assessment of the four Arabic language skills must align with the dimensions of creative, evaluative, and analytical thinking processes described in the competency achievement indicators. Thus, the results of this study reinforce the consensus that developing HOTS-based questions and learning activities not only raise the cognitive level of the items but also significantly improve students' critical thinking skills and advanced language proficiency in Arabic.

CONCLUSION

The findings indicate that current Arabic assessment instruments for the Tasawwuq chapter in the eleventh-grade Madrasah Aliyah textbook predominantly consist of lower-order thinking skills (LOTS) items, with 79% of questions at C1–C2, 21% at C4, and none at C5–C6. This distribution reveals a significant misalignment with 21st-century curricular requirements for higher-order thinking skills (HOTS) in learning. To address this issue, a HOTS-oriented assessment for tarkib and qira'ah was systematically developed using the ADDIE model. The resulting instrument achieved a high level of expert validation (93.98%) in both content and language, and featured a balanced representation of C4, C5, and C6 items targeting analysis, evaluation, and creation. Implementation of this instrument at MAN 4 Agam led to a statistically significant improvement in student outcomes, with mean scores rising from 67.33 (pretest) to 82.57 (posttest), and a paired-samples t-test yielded $t = -13.527$, $p < 0.001$. These results demonstrate that HOTS-based Arabic assessment can effectively enhance students' higher-order thinking in qira'ah and tarkib.

The limitations of this study include limited instrument testing conducted at a single institution and a one-group pretest–posttest design, which limits the generalizability of the learning outcome improvement. Instrument development is still focused on two skills. The limitations of this study can be used as suggestions for further research.

REFERENCES

Arrias, J. C., Alvarado, D., & Calderón, M. (2019). *Menilik Konsep Kemampuan Berpikir Tingkat Tinggi (Higher Older Thingking Skills) Dalam Pembelajran Matematika*. *6*(2), 5–10.

Baihaqi, A. I., Syammary, N. A., Masrifah, K., Murtadho, N., & Nurhidayati, N. (2025). Development of HOTS-Based Arabic Language Module to Improve Students' Critical Thinking Skills. *Arabi : Journal of Arabic Studies*, *10*(1), 39–52. https://doi.org/10.24865/ajas.v10i1.908

Cojorn, K., & Sonsupap, K. (2024). A collaborative professional development and its impact on teachers' ability to foster higher order thinking. *Journal of Education and Learning (EduLearn)*, *18*(2), 561–569. https://doi.org/10.11591/edulearn.v18i2.21182

Faizah, F., A'yuni, K., & Al-Hakim, M. F. (2024). Development of The Alibaba Media Game Assisted with The Make a Match Model in Arabic Language Learning To Improve The Vocabulary Mastery of Class IV Primary School Students. *Al-Mudarris*, *7*(1). https://doi.org/10.32478/kpve1947

Fanani, Moh. Z., Ariyana, Y., Pudjiastuti, A., Bestary, R., Zamroni, & Ainin, M. (2018). Strategi Pengembangan Soal HOTS Pada Kurikulum 2013. *Kementerian Pendidikan Dan Kebudayaan*, *4*(4), 1–87. https://doi.org/10.30762/ed.v2i1.582

Farros, A. A. (2024). Analysis of Arabic Language Questions Based on Higher Order Thinking Skills (HOTS) in the Arabic Language Book for Class XI Ministry of Religion. *Al-Muyassar: Journal of Arabic Education*, *3*(1), 52–69. https://doi.org/10.31000/al-muyassar.v3i1.10273

Fayyoumi, K. A. R. M. A., Hanandeh, A. M., & Ayasrah, S. (2024). Analyzing Active Tasks in Jordanian Primary Arabic Textbooks: Teaching Strategies and Skill Development Implications. *International Journal of Instruction*, *17*(2), 497–518.

Gozali, I., Lie, A., Tamah, S. M., & Jemadi, F. (2021). HOTS questioning ability and HOTS perception of language teachers in Indonesia. *HOTS Questioning Ability and HOTS Perception of Language Teachers in Indonesia.*, *11*(1), 60–71. https://doi.org/10.17509/ijal.v11i1.34583

Haniefa, R. (2022). Implementasi Model Penilaian Hots (Higher Order Thinking Skills) Pada Penilaian Empat Keterampilan Berbahasa Arab. *Ta'limi | Journal of Arabic Education and Arabic Studies*, *1*(1), 49–71. https://doi.org/10.53038/tlmi.v1i1.11

Hilmi, I., Fadlila, N., Ramadanti, E., Retnawati, H., & Arliani, E. (2022). Development of Higher Order Thinking Skills Test Based on Revised Bloom Taxonomy. *JTAM (Jurnal Teori Dan Aplikasi Matematika)*, *6*(2), 341–353.

Ilmiani, A. M., & Delima, D. (2021). Innovation in Learning Arabic Reading Skills using Higher Order Thinking Skills. *Al-Ta'rib : Jurnal Ilmiah Program Studi Pendidikan Bahasa Arab IAIN Palangka Raya*, *9*(1), 99–110. https://doi.org/10.23971/altarib.v9i1.2603

Isnaeni, W., Rudyatmi, E., Ridlo, S., Ingesti, S., & Adiani, L. R. (2021). Improving students' communication skills and critical thinking ability with ICT-oriented problem-based learning and the assessment instruments with HOTS criteria on the immune system material. *Journal of Physics: Conference Series*, *1918*(5), 052048. https://doi.org/10.1088/1742-6596/1918/5/052048

Kartini, N. E., Nurdin, E. S., Hakam, K. A., & Syihabuddin, S. (2022). Telaah Revisi Teori Domain Kognitif Taksonomi Bloom dan Keterkaitannya dalam Kurikulum Pendidikan Agama Islam. *Jurnal Basicedu*, *6*(4), 7292–7302. https://doi.org/10.31004/basicedu.v6i4.3478

Muradi, A., Mubarak, F., Darmawaty, R., & Hakim, A. R. (2020). Higher Order Thinking Skills Dalam Kompetensi Dasar Bahasa Arab. *Arabi : Journal of Arabic Studies*, *5*(2), 177. https://doi.org/10.24865/ajas.v5i2.293

Mustofa, S., Desrani, A., & Ritonga, A. W. (2022). HOTS in Arabic Learning: A Study of The Implementation of HOTS on Students' Critical Thinking Ability. *Al-Ta'rib : Jurnal Ilmiah Program Studi Pendidikan Bahasa Arab IAIN Palangka Raya*, *10*(2), 133–144. https://doi.org/10.23971/altarib.v10i2.4088

Pisriwati, S. A., Siswanto, D. H., Hardi, Y., & Alghiffari, E. K. (2024). Question Making Training with LOTS, MOTS, and HOTS Cognitive Levels for High School Teachers. *Journal of Social and Community Development*, *1*(01), 9–18. https://doi.org/10.56741/jscd.v1i01.666

Ragab, A., Kaid, A., & Sayed, A. K. (2024). Enhancing Higher Order Thinking Skills (HOTS) in Education: Strategies and Outcomes. *TOFEDU: The Future of Education Journal*, *3*(5), 1488–1499. https://doi.org/10.61445/tofedu.v3i5.267

Riswanda, J. (2018). Pengembangan Soal Berbasis Higher Order Thinking Skill (Hots) Serta Implementasinya di SMA Negeri 8 Palembang. *Jurnal Penelitian Pendidikan Biologi*, *2*(1), 49–58.

Sari, R. R., & Amrullah, H. (2020). *Bahasa Arab MA Kelas XI*. Direktorat KSKK Madrasah.

Sarosa, S. (2021). *Analisis Data Penelitian Kualitatif*. PT Kanisius.

Setyorini, R. (2025). The Dominance of LOTS In Summative Assessment: The Challenge of Improving HOTS in Indonesian Language Learning. *EDUCATIONE*, 33–42. https://doi.org/10.59397/edu.v3i1.44

Tuela, A. I., & Palar, Y. N. (2022). Analysis of Higher Order Thinking Skills (HOTS) Based on Bloom Taxonomy in Comprehensive Examination Questions. *Jurnal Kependidikan : Jurnal Hasil Penelitian dan Kajian Kepustakaan di Bidang Pendidikan, Pengajaran, dan Pembelajaran*, *8*(4), 957–971. https://doi.org/10.33394/jk.v8i4.5885

Tyas, E. H., & Naibaho, L. (2021). HOTS Learning Model Improves the Quality of Education. *International Journal of Research-GRANTHAALAYAH*, *9*(1), 176–182. https://doi.org/10.29121/granthaalayah.v9.i1.2021.3100

Verawati, Febriani, Muflihah, Hasanah, Susanti, & Fitriani. (2022). HOTS Analysis of Task Instructions in Bahasa Arab Madrasah Aliyah Textbook Published by The Ministry of Religious Affairs. *Edukatif : Jurnal Ilmu Pendidikan*, *4*(1).

Wahyuni, A. S., Alfatha, S., Nuraini, A., Yusba, F. A., & Rahmansyah, E. (2023). Communication skill (4C) of the 21st Century's Learning In Arabic Language Research Journal At Indonesia. *Annual International Conference in Education and Islamic Studies 2022*.

Wicaksono, A. R. (2021). Pengembangan Soal Berbasis HOTS Mata Pelajaran PAI di SMK 17 Seyegan. *Bintang : Jurnal Pendidikan Dan Sains*, *3*(1), 94–112.

Yasmara, R., Al Husna, L., Wargadinata, W., & Nurhadi, N. (2023). Exploration HOTS (Higher Order Thinking Skills) in Arabic Textbook. *Tanwir Arabiyyah: Arabic as Foreign Language Journal*, *3*(1), 97–110. https://doi.org/10.31869/afl.v3i2.4795