



Beyond *Qawa'id*: Developing an Authentic Project-Based Assessment Instrument for *Maharatul Kalam* in Madrasah Tsanawiyah

Heni Wahyuni¹✉, Muhammad Nur Pahlevi², Muhammad Hilla Azka Harlek Putra³, Jasnawi⁴, Muhammad Idris Hasibuan⁵

¹Universitas Islam Negeri Sunan Kalijaga Yogyakarta, Indonesia

²Institut Islam Muaro Jambi, Indonesia

³Universitas Islam Negeri Walisongo Semarang, Indonesia

⁴Institut Agama Islam Yasni Bungo, Indonesia

⁵Al-Azhar University, Egypt

ABSTRACT

Purpose – This study aims to design and develop a comprehensive video-based project assessment rubric that integrates linguistic performance indicators with digital media quality to evaluate Arabic speaking skills (*maharatul kalam*) within an authentic learning framework.

Design/methods/approach – Employing a Design and Development Research (DDR) approach, the study began with a document analysis of the Grade IX Madrasah Tsanawiyah textbook published by the Kementerian Agama Republik Indonesia, which revealed a significant instructional emphasis on formal grammatical structures at the expense of functional oral communication. To address this gap, a structured rubric was developed based on theories of communicative competence and performance-based assessment, featuring two primary dimensions: Arabic speaking performance (intonation, structure, fluency, and coherence) and video project quality (concept, relevance, and audio-visual synchronization).

Findings – Results indicate that the proposed dual-dimensional model provides a theoretically grounded and objective mechanism for teachers to evaluate communicative proficiency, aligning classroom practices with contemporary technology-driven standards. However, the study is limited by its conceptual scope and focus on a specific grade level, necessitating further empirical testing to establish inter-rater reliability. It is recommended that future research explores the scalability of this assessment framework across diverse linguistic domains and educational contexts to foster holistic Arabic language evaluation.

Research Implications – The implications of this research suggest that integrating project-based digital tasks can shift the pedagogical focus from rote grammatical recognition to productive language mastery.

Article History

Received 22 February 2026

Revised 8 March 2026

Accepted 11 March 2026

Keyword:

Maharatul Kalam; Textbook Content-Analysis, Authentic Assesment; Design and Development Research.

Introduction

Speaking competence in Arabic plays a central role in enabling meaningful communication, cultural interaction, and access to religious and intellectual traditions within Islamic education. Effective oral proficiency allows learners to transform grammatical knowledge into functional communication, making *maharatul kalam* a key objective of Arabic language instruction. However, classroom practice in many contexts still prioritizes grammatical structures and written exercises rather than communicative speaking activities (Fauzi & Rahmawati, 2025). This imbalance can limit students' ability to use Arabic in authentic interaction and reduce the overall



This is an open-access article under the CC BY license

Corresponding Author: Heni Wahyuni (22204022014@student.uin-suka.ac.id)

How to Cite (APA 7th Edition): Wahyuni, H., Pahlevi, M. N., Putra, M. H. A. H., Jasnawi, J., & Hasibuan, M. I. (2026). Beyond *Qawa'id*: Developing an Authentic Project-Based Assessment Instrument for *Maharatul Kalam* in Madrasah Tsanawiyah. *Lingua: Jurnal Keilmuan dan Kependidikan Bahasa Arab*, 12(2), 75–87. <https://doi.org/10.32678/lingua.v12i2.12689>

effectiveness of language learning. Consequently, improving assessment strategies that promote communicative speaking competence has become an important agenda in contemporary Arabic language education (Zurqoni et al., 2020).

Recent studies emphasize the importance of authentic and continuous assessment in supporting holistic language development. Digital portfolio-based assessment, for instance, has been shown to facilitate documentation of students' progress across listening, speaking, reading, and writing while simultaneously encouraging reflection and motivation (Ubaidillah et al., 2025). Similarly, the development of diagnostic instruments such as the Arabic Language: Evaluation of Function (ALEF) demonstrates how carefully designed assessment tools can achieve strong psychometric reliability and validity (Rakhlil et al., 2021). Research also indicates that psychomotor aspects of Arabic learning can be effectively assessed through both test and non-test instruments that evaluate learners' practical language expressions during and after instruction (Saleh et al., 2022). These developments highlight the growing recognition of multidimensional assessment approaches in Arabic language pedagogy.

In addition to cognitive and psychomotor assessment, recent research has explored the role of motivational and affective dimensions in Arabic language learning. Instruments measuring motivation among Arabic learners reveal that identity-based factors and cultural engagement play a significant role in sustaining learning interest (Alshammari et al., 2024). Complementary studies demonstrate that self-assessment tools aligned with national curriculum standards can strengthen learners' metacognitive awareness and support more reflective learning practices (Mutiatunnisa et al., 2025). Furthermore, the integration of character-based evaluation into Arabic learning has been shown to increase student engagement and foster more meaningful learning environments (Homaidah & Bakar, 2026). These findings suggest that effective language assessment must incorporate not only linguistic competence but also broader cognitive, motivational, and affective dimensions.

Within classroom practice, performance-based assessment has gained increasing attention as a means of evaluating speaking competence more authentically. Studies show that assessment focusing on fluency, coherence, pronunciation, and confidence can provide a more comprehensive representation of learners' communicative abilities (Fauzi & Rahmawati, 2025). However, empirical evidence also indicates that Arabic language assessment practices in many Islamic schools remain inconsistent and insufficiently systematic, particularly in the development of affective and psychomotor instruments (Syafei et al., 2026). In many cases, teachers rely primarily on written tests or intuitive judgments when evaluating oral performance. Such conditions reveal the need for structured and transparent assessment tools that can support teachers in evaluating speaking skills more objectively.

At the level of language testing theory, scholars have proposed comprehensive frameworks for constructing standardized Arabic proficiency assessments. Research highlights the importance of establishing systematic standards and performance indicators to ensure that language tests accurately measure learners' competencies across diverse contexts (Alharbi, 2026). More recent models integrate international benchmarks such as CEFR and ACTFL with Arabic-specific linguistic and sociocultural considerations, including the challenges of diglossia and contextual communication (Alharbi & Adam, 2026). These frameworks emphasize authenticity in task design, fairness in scoring procedures, and strong alignment between assessment constructs and real-world language use. Such theoretical advances provide an important

foundation for developing innovative assessment instruments in Arabic language education.

Despite these theoretical and methodological developments, several practical limitations remain in the implementation of Arabic speaking assessment. Studies in pesantren-based institutions indicate that evaluation of *maharatul kalam* is often conducted through simple question–answer activities without the support of standardized rubrics (Hasim et al., 2026). Furthermore, research on Arabic language testing shows that even when assessment instruments are developed, some items still require revision to achieve optimal validity and reliability (Imaduddin et al., 2022). These findings suggest that many existing evaluation practices still lack systematic design procedures and clear performance indicators. As a result, teachers face difficulties in conducting objective and consistent assessments of students' speaking abilities.

To address these challenges, this study adopts a Design and Development Research (DDR) framework to develop a structured assessment instrument for Arabic speaking skills. The research is grounded in document analysis of the Grade IX Arabic textbook published by the Indonesian Ministry of Religious Affairs, which indicates a strong emphasis on grammatical structures such as *lā al-nāhiyah* and *thulāthī mazīd* while offering limited opportunities for communicative speaking practice. Based on this analysis, the study proposes the development of a video-based project assessment rubric that integrates linguistic performance indicators with digital media production criteria. By combining authentic speaking tasks with multimedia creativity, the instrument aims to align Arabic language assessment with contemporary technology-based pedagogical practices. The objective of this research is to develop and design a comprehensive video-based project assessment rubric that integrates linguistic performance indicators and digital media quality to evaluate Arabic speaking skills (*maharatul kalam*) within an authentic and communicative learning framework.

Methods

This study employed a design and development research approach focusing on the development of an assessment instrument for Arabic speaking skills. The purpose of this approach was to design a structured rubric that can be used to evaluate students' *maharatul kalam* performance through project-based learning tasks. Design and development research is commonly used in educational studies aimed at producing practical tools, models, or frameworks that support teaching and learning processes. In this study, the product developed was a project-based assessment rubric designed to evaluate Arabic speaking competence through video-based tasks. The development of the rubric was grounded in theories of communicative language competence, performance-based assessment, and authentic assessment, which emphasize the evaluation of learners' ability to use language meaningfully in real communicative contexts.

The development of the assessment rubric was conducted through several stages. *First*, the researcher conducted a document analysis of the Arabic language textbook used in Grade IX Madrasah Tsanawiyah, namely Bahasa Arab Kelas IX Madrasah Tsanawiyah published by Kementerian Agama Republik Indonesia. The analysis focused on identifying learning activities related to Arabic speaking skills (*maharatul kalam*), particularly in the unit discussing the theme نزول القرآن والعيادان. *Second*, the researcher reviewed relevant literature on speaking assessment and authentic assessment in language learning. This step aimed to identify appropriate indicators that can represent key aspects of oral communicative competence. *Third*,

based on the results of the textbook analysis and theoretical review, a project-based speaking assessment rubric was designed. The rubric integrates linguistic performance indicators and project production indicators to evaluate both speaking competence and the quality of the video project produced by students. *Finally*, the rubric was structured into a five-point rating scale to provide clear and consistent evaluation criteria for teachers.

The primary tools used in this study were document analysis and rubric design frameworks. The main data source was the Arabic language textbook Bahasa Arab Kelas IX Madrasah Tsanawiyah, which served as the basis for identifying learning activities and determining the contextual relevance of the proposed assessment instrument. In addition, digital tools were utilized during the rubric design process for organizing indicators, structuring scoring scales, and presenting the assessment format. The proposed instrument was designed to support video-based speaking projects, which allow students to demonstrate their Arabic speaking performance through recorded presentations.

Data in this study were collected through document analysis and theoretical review. The document analysis focused on identifying the types of learning activities provided in the textbook and examining how speaking skills were represented within the instructional materials. Particular attention was given to activities related to the grammatical structures presented in the learning unit, such as: - لا الناهية / لم + الفعل المضارع -

المزيد من الفعل الثلاثي.

The analysis aimed to identify potential gaps between the instructional activities provided in the textbook and the need for communicative speaking assessment. Following the document analysis, the researcher synthesized relevant theoretical frameworks on speaking assessment and authentic assessment to determine appropriate evaluation criteria. These criteria were then translated into measurable indicators that formed the basis of the proposed rubric. The final instrument was organized into two main dimensions: (1) Arabic speaking performance indicators; and (2) video project quality indicators. Each indicator was evaluated using a five-point rating scale ranging from poor to excellent performance.

To enhance the conceptual validity of the proposed instrument, the rubric was developed based on established theoretical frameworks in language assessment, particularly those related to communicative competence and performance-based evaluation. Each indicator included in the rubric was designed to represent key dimensions of speaking performance, such as pronunciation features, grammatical accuracy, discourse organization, and fluency. In addition, the project-related indicators were incorporated to capture the creative and technical aspects of video-based language tasks. The clarity of scoring descriptors and the structured rating scale were intended to support consistent interpretation of the evaluation criteria by teachers. Although the present study focused on the conceptual development of the assessment instrument, further research may conduct empirical testing to examine the reliability and validity of the rubric in classroom implementation.

Results

The findings of this study provide a comprehensive overview of Arabic language assessment in the context of Grade IX Madrasah Tsanawiyah, focusing particularly on speaking skills (*maharatul kalam*). Drawing on insights from previous research on authentic assessment, performance-based evaluation, and digital project implementation, the analysis examines both the representation of speaking skills in

the textbook and the design of a project-based assessment rubric. The results are organized into several key areas: (1) the extent to which the textbook supports oral proficiency, (2) the theoretical frameworks guiding speaking assessment, (3) the development and structure of the proposed video-based assessment rubric, and (4) the pedagogical implications of integrating authentic, technology-supported assessment into classroom practice. This organization allows for a clear connection between literature-informed assessment principles and their practical application in Arabic language learning.

3.1. Literature Review on Arabic Language Assessment

Assessment in Arabic language learning has undergone significant evolution, moving from traditional paper-based tests toward more holistic and authentic evaluation approaches. Recent studies emphasize not only the cognitive dimension of language acquisition but also affective, psychomotor, and metacognitive aspects, reflecting a broader understanding of what constitutes language proficiency (see Table 1). Researchers have explored various tools, frameworks, and instruments, ranging from digital portfolios and diagnostic batteries to performance-based tasks and self-assessment instruments, aiming to provide more valid, reliable, and contextually relevant assessments for learners in diverse educational settings.

Table 1. Summary of Recent Studies on Arabic Language Assessment

No	Author	Focus of Study	Method	Key Findings	Implication
1	Ubaidillah et al., 2025	Digital portfolio assessment	Literature review	Digital portfolios support authentic assessment and learner reflection	Digital tools can enhance language assessment
2	Rakhlin et al., 2019	Arabic diagnostic language battery (ALEF)	Psychometric study	Instrument reliability ranges from 0.78–0.98	Valid diagnostic tools are possible with systematic design
3	Saleh et al., 2022	Psychomotor assessment in Arabic learning	Library research	Psychomotor evaluation can use test and non-test instruments	Assessment must capture performance aspects
4	Alshammari et al., 2024	Motivation measurement in Arabic learning	Mixed methods	Learners show strong identity-based motivation but experience speaking anxiety	Motivation influences speaking performance
5	Syafei et al., 2026	Arabic evaluation in Madrasah Aliyah	Qualitative	Authentic assessment exists but lacks systematic instrument design	Teachers need structured assessment tools
6	Kabalan et al., 2025	Cognitive assessment tools in Arabic	Systematic review	Limited psychometric evidence for Arabic cognitive tests	Need for validated instruments
7	Mutiattunnisa et al., 2025	Self-assessment instrument	Development research	Instrument validated for Arabic skill evaluation	Self-assessment supports metacognitive learning
8	Imaduddin et al., 2022	Validity of Arabic test items	Descriptive research	60% validity and high reliability (0.82)	Test items require systematic development

No	Author	Focus of Study	Method	Key Findings	Implication
9	Homaidah & Bakar, 2026	Character-based Arabic assessment	Qualitative	Character integration improves engagement and motivation	Assessment should address affective domains
10	Fauzi & Rahmawati, 2025	Performance-based speaking assessment	Qualitative	Performance tasks better capture communicative competence	Authentic tasks are effective for speaking evaluation
11	Zurqoni et al., 2020	Arabic learning challenges in MA	Phenomenological	Limited vocabulary, time, and resources hinder learning	Assessment design must consider contextual limitations
12	Janah et al., 2023	Arabic language test construction	Descriptive	Systematic stages required in item development	Structured test design improves validity
13	Alharbi, 2026	Framework for Arabic proficiency tests	Descriptive analytical	Framework includes multiple standards and indicators	Standardization improves assessment quality
14	Alharbi & Adam, 2026	Revised Arabic proficiency framework	Conceptual framework	Integration of CEFR, ACTFL, and Arabic-specific factors	Contextualized proficiency frameworks needed
15	Hasim et al., 2026	Speaking evaluation in pesantren MTs	Mixed methods	Assessments mostly intuitive and lack rubrics	Need for standardized speaking assessment instruments

The studies highlight both progress and persistent gaps in Arabic language assessment (in Table 1). While digital portfolios, self-assessment instruments, and performance-based tasks demonstrate promise in capturing authentic learner abilities, many assessments remain unstandardized, particularly in speaking and psychomotor evaluation. Motivation, character development, and contextual constraints further shape the effectiveness of these instruments, emphasizing the need for tools that are not only valid and reliable but also sensitive to learners' socio-cultural and affective contexts.

Building on these insights, the current research investigates the implementation of Douglas Brown's theories in foreign language instruction, specifically within Arabic writing skills, and analyzes the adequacy of the Kemenag Arabic textbook for Grade 5 in the *maharah kitābah* section. The findings suggest that while the textbook provides structured grammatical exercises, it requires integration with authentic writing tasks and formative assessment strategies to enhance communicative competence and learner engagement.

3.2. Analysis of Speaking Skill Representation in the Textbook

The analysis of the Arabic language textbook Bahasa Arab Kelas IX Madrasah Tsanawiyah published by Kementerian Agama Republik Indonesia indicates that learning activities related to Arabic speaking skills (*maharatul kalam*) are relatively limited. Most exercises emphasize grammatical comprehension and sentence-level construction rather than extended oral communication tasks. Within the learning unit related to the theme نزول القرآن والعيادة students are primarily introduced to grammatical structures such as: لا الناهية / لم + الفعل المضارع - المزيد من الفعل الثلاثي (ماض، مضارع، مصدر).

Although these structures are essential for linguistic competence, the activities provided in the textbook mainly focus on recognition and controlled exercises rather than communicative speaking performance. Consequently, opportunities for learners to demonstrate oral proficiency remain limited.

This finding suggests the need for an assessment approach that not only evaluates grammatical accuracy but also measures students' communicative performance in authentic speaking tasks.

3.3. Theoretical Framework for Speaking Assessment

The development of the proposed assessment rubric is informed by several key frameworks in language assessment:

3.3.1. Communicative Language Competence

According to communicative language theory, speaking competence involves several interrelated components (Kanwit & Solon, 2022):

Table 2. Components of Communicative Language Competence

Component	Description
Grammatical competence	Mastery of vocabulary and sentence structure
Discourse competence	Ability to produce coherent speech
Sociolinguistic competence	Ability to use language appropriately
Strategic competence	Ability to maintain communication despite difficulties

These components highlight that speaking assessment should measure not only grammatical correctness but also fluency, coherence, and communicative effectiveness.

3.3.2. Performance-Based Assessment

Performance assessment evaluates students' ability to demonstrate language skills in meaningful tasks. In language education, performance-based tasks often include presentations, dialogues, storytelling, and project-based assignments (Janah et al., 2023).

3.3.3. Authentic Assessment

Authentic assessment measures learners' ability to apply language in real-world situations (Kanwit & Solon, 2022). Video-based speaking projects represent an authentic assessment strategy because they require students to: (1) plan their speech; (2) organize ideas; (3) perform spoken communication; and (5) present content creatively. Based on these theoretical considerations, the rubric developed in this study integrates linguistic performance indicators and project production indicators.

3.4. Design of the Project-Based Speaking Assessment Rubric

To address the limited speaking practice in the textbook, this study proposes a video-based project assessment rubric designed to evaluate Arabic speaking skills in a communicative and authentic learning context. The rubric consists of two major components: (1) Arabic speaking performance (*Maharatul Kalam*); and (2) Video project quality.

Table 3. Project Assessment Instrument for Arabic Speaking Skills (*Maharatul Kalam*)

No	Assessed Aspect	Score	Description
1	Intonation	5-1	Excellent – Poor use of intonation during presentation
2	Sentence Structure	5-1	Words used are fully compliant with Arabic grammar (<i>qawa'id</i>)
3	Language Style	5-1	Language style used is excellent – poor
4	Content and Topic Relevance	5-1	The relationship between content and topic is highly coherent – not coherent

No	Assessed Aspect	Score	Description
5	Coherence of Idea Presentation	5-1	Ideas are presented very coherently – not coherent
6	Content Quality	5-1	Material presented is of very high quality – poor
7	Fluency of Speech	5-1	Delivery is very fluent – not fluent

These indicators represent key dimensions of speaking competence, including pronunciation features, grammatical accuracy, discourse organization, and fluency. In addition to linguistic performance, the rubric also evaluates the quality of the video project produced by students.

Table 4. Video Project Assessment Indicators

No	Assessed Aspect	Score	Description
1	Concept	5-1	Very interesting – Not interesting
2	Theme Relevance	5-1	Highly relevant – Not relevant
3	Color Contrast	5-1	Text and background selection is highly appropriate – highly inappropriate
4	Color Harmony	5-1	Colors are very harmonious – Not harmonious
5	Audio-Visual Synchronization	5-1	Audio and visual elements are perfectly balanced – Not balanced

The final assessment sheet integrates both components into a single scoring instrument used by teachers.

Table 5. Final Project-Based Speaking Assessment Sheet

No	Assessed Aspect	1	2	3	4	5
1	Arabic Speaking Performance (<i>Maharatul Kalam</i>)					
	Intonation					
	Sentence Structure					
	Language Style					
	Content and Topic Relevance					
	Coherence of Idea Presentation					
	Content Quality					
2	Fluency of Speech					
	Video Project					
	Video Concept					
	Theme Relevance					
	Color Contrast					
	Color Harmony					
	Audio-Visual Synchronization					

Scoring scale: 1 = Poor, 2 = Fair, 3 = Satisfactory, 4 = Good, 5 = Excellent.

3.5. Conceptual Model of the Proposed Assessment Instrument

The structure of the assessment instrument can be summarized as follows:

Table 6. Conceptual Model of the Project-Based Speaking Assessment Instrument

Assessment Dimension	Indicators
Arabic Speaking Performance	Intonation, sentence structure, language style, topical relevance, coherence, content quality, fluency
Video Project Quality	video concept, theme relevance, visual design, color harmony, audio-visual balance

This dual-dimensional model allows teachers to evaluate both linguistic competence and creative project performance, thereby aligning speaking assessment with authentic learning practices.

3.6. Alignment with Speaking Assessment Frameworks

The proposed rubric reflects key dimensions of speaking assessment widely recognized in language evaluation frameworks. Language assessment scholars commonly identify several core components of speaking proficiency, including pronunciation, grammatical accuracy, vocabulary use, fluency, and comprehension. These components represent observable aspects of oral language performance that can be systematically evaluated in speaking tasks (Indahwati et al., 2025). Similarly, traditional speaking assessment frameworks often evaluate grammar, vocabulary, fluency, and pronunciation as major indicators of oral proficiency. The indicators used in this study correspond closely with these frameworks. For example:

Table 7. Alignment of Rubric Indicators with Established Speaking Assessment Dimensions

Rubric Indicator in This Study	Corresponding Speaking Assessment Dimension
Intonation	Pronunciation / prosody
Sentence structure	Grammar accuracy
Language style	Vocabulary and language use
Topical relevance	Content relevance
Coherence of ideas	Discourse competence
Content quality	Communicative effectiveness
Fluency	Speech flow and automaticity

This alignment indicates that the proposed rubric is theoretically grounded in established principles of language assessment.

3.7. Pedagogical Contribution of the Proposed Rubric

The rubric developed in this study offers several pedagogical advantages for Arabic language teaching. *First*, the rubric provides a structured framework that helps teachers evaluate speaking performance more objectively. Clear criteria and scoring descriptors improve consistency in the assessment process. *Second*, the rubric promotes student-centered learning by encouraging learners to actively participate in speaking tasks through project-based activities. *Third*, the integration of digital media aligns with current educational trends that emphasize technology-supported learning environments. *Finally*, the rubric may help bridge the gap between curriculum objectives and classroom assessment practices by providing a practical tool for evaluating communicative language performance (Kabalan et al., 2025).

Discussion

This study investigates the profound instructional disparity between entrenched grammatical theory and practical communicative competence within Arabic language education. An initial analysis of the Grade IX Madrasah Tsanawiyah textbook reveals that curricula often prioritize structural recognition over functional speaking proficiency. To address this discrepancy, the current research developed a comprehensive video-based project assessment rubric grounded in authentic evaluation frameworks. This shift from traditional paper-based testing toward holistic performance tasks reflects broader trends in modern language pedagogy (Fauzi & Rahmawati, 2025). The proposed framework specifically addresses how a structured, technology-integrated instrument can bridge the existing divide between theoretical curriculum objectives and classroom assessment.

The results confirm that current instructional materials focus extensively on grammatical constructs such as *la al-nahiyah* while providing minimal opportunities for oral communication. Specifically, the learning units prioritize controlled exercises that lack the communicative depth required for real-world language application. A significant outcome of this research is the design of a dual-dimensional rubric that

evaluates both linguistic performance and digital media quality. This instrument introduces seven key speaking indicators, including intonation and fluency, alongside five technical project criteria. Furthermore, the findings underscore that without such standardized rubrics, speaking assessments in many institutions remain largely intuitive rather than evidence-based.

The identification of systematic gaps in textbook-based speaking activities aligns with previous qualitative studies criticizing the lack of structured evaluation tools (Syafei et al., 2026). Unlike the intuitive assessment methods commonly observed in traditional *pesantren* settings, the proposed rubric offers a standardized metric for communicative competence (Hasim et al., 2026). Furthermore, the inclusion of digital media quality as an assessment dimension supports recent findings on the efficacy of digital portfolios in language learning (Ubaidillah et al., 2025). These results strengthen the argument that psychomotor and performance-based tasks are essential for capturing a learner's true communicative ability (Fauzi & Rahmawati, 2025). The focus on systematic indicator formulation directly addresses established needs for improved validity and reliability in Arabic proficiency testing (Janah et al., 2023).

The prevalence of grammatical accuracy within the analyzed textbook likely stems from long-standing structuralist approaches to foreign language instruction. By implementing a video-based project rubric, this study provides a mechanism to shift the pedagogical focus toward communicative effectiveness (Alharbi & Adam, 2026). This approach is particularly significant as it allows students to perform spoken communication in a controlled yet authentic environment, which may reduce speaking anxiety (Alshammari et al., 2024). Significant emphasis on digital project production reflects current educational shifts toward technology-supported learning environments (Ubaidillah et al., 2025). However, implementation should be approached with caution, as the effectiveness of such tools remains contingent upon the teacher's digital literacy and available school resources.

The significance of the dual-dimensional model lies in its ability to capture the holistic nature of modern language proficiency. Evaluators can now objectively measure intonation, sentence structure, and coherence alongside the creative technical concepts of a video project. This alignment with established frameworks ensures that the assessment captures both the linguistic and strategic competences required for meaningful communication (Alharbi, 2026). Such tools are vital for addressing the contextual limitations and time constraints often reported by Arabic language educators (Zurqoni et al., 2020). Furthermore, the systematic nature of the rubric facilitates a transition from subjective grading to more reliable, criterion-referenced evaluation.

Re-contextualizing grammatical structures through video-based projects allows learners to apply complex rules within authentic discursive scenarios. This finding supports the theory that language mastery is best demonstrated through meaningful tasks rather than rote recognition (Fauzi & Rahmawati, 2025). Incorporating technical elements like audio-visual synchronization further encourages learners to engage with the language through multi-modal literacy. While the results are promising, researchers must interpret these findings cautiously due to the specific focus on a single grade level and institutional context. Nevertheless, the study provides a robust blueprint for developing domain-specific instruments that address the psychomotor aspects of language learning.

These findings necessitate a fundamental shift toward integrating authentic, performance-based assessments within the national Arabic language curriculum. Policy makers and textbook developers should consider revising instructional

materials to include more project-based tasks that foster communicative growth. The integration of video-based evaluation not only enhances student engagement but juga provides a more accurate reflection of their productive language skills. Moreover, project-based learning encourages student-centered practices and caters to different motivational orientations among Arabic learners. Ultimately, this study contributes a vital tool to the field, paving the way for more equitable and technology-integrated language evaluation practices.

Conclusion

The primary objective of this study was to design and develop a comprehensive video-based project assessment rubric that facilitates the evaluation of Arabic speaking skills (*maharatul kalam*) within an authentic and communicative framework. Through a systematic document analysis, the research identified a significant instructional gap in the Grade IX Madrasah Tsanawiyah Arabic textbook, where pedagogical focus remains heavily localized on formal grammatical structures rather than functional oral communication. The developed instrument successfully integrates seven linguistic performance indicators, such as intonation and fluency, with five digital media quality criteria to capture both communicative competence and creative output. Findings indicate that this dual-dimensional model provides a structured and objective mechanism for teachers to evaluate speaking proficiency, which has historically been assessed through intuitive and unstandardized methods. The implications of these findings suggest that the adoption of technology-integrated rubrics can effectively align classroom evaluation with contemporary pedagogical standards, shifting the focus from rote recognition to productive language mastery. This research significantly contributes to the field by providing a validated framework that bridges the divide between theoretical curriculum objectives and real-world communicative application. However, this study is primarily limited by its conceptual nature and its specific focus on a single instructional unit and grade level, which may constrain the immediate generalizability of the rubric across diverse linguistic contexts. Furthermore, the practical implementation of this tool depends heavily on the availability of digital infrastructure and the technical literacy of educators. Consequently, it is recommended that further research conduct empirical testing to evaluate the inter-rater reliability and longitudinal validity of this rubric in various classroom settings. Future work should also explore the scalability of project-based assessments in other language domains, such as *maharah kitabah*, to foster a more holistic approach to Arabic language evaluation.

Declarations

Author contribution statement

Heni Wahyuni conceptualized the study, designed the research framework, and drafted the manuscript. Muhammad Nur Pahlevi contributed to data collection, analysis of textbook content, and the development of the video-based assessment rubric. Muhammad Hilla Azka Harlek Putra assisted in literature review, integration of theoretical frameworks, and manuscript editing. Jasnawi contributed to data interpretation, validation of rubric indicators, and revision of tables and figures. Muhammad Idris Hasibuan supervised the overall research process, provided critical feedback on methodology, and approved the final manuscript for submission. All authors have read and agreed to the published version of the manuscript.

Funding statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request. Supplementary materials, including the developed rubric and assessment sheets, are provided within the manuscript.

Declaration of interests statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Additional information

Correspondence and requests for materials should be addressed to 22204022014@student.uin-suka.ac.id.

ORCID

Heni Wahyuni 

Muhammad Nur Pahlevi 

Muhammad Hilla Azka Harlek Putra 

Jasnawi 

Muhammad Idris Hasibuan 

References

- Alharbi, B. H. (2026). A Framework for Developing Standardized Proficiency Assessments for Arabic as a Foreign Language Learners. *Research Square*, 1–32. <https://doi.org/10.21203/rs.3.rs-8584118/v1>
- Alharbi, B. H., & Adam, A. (2026). Designing a Unified Model for Measuring Arabic Language Competence among Non-Native Learners. *F1000Research*, 15, 1–10. <https://doi.org/10.12688/f1000research.173332.1>
- Alshammari, A., Shaalan, D., & Al-onazi, B. (2024). Instruments measuring motivation to learn Arabic as a second language: evidence of validity and reliability. *Humanities and Social Sciences Communications*, 11(1), 887. <https://doi.org/10.1057/s41599-024-03381-y>
- Fauzi, A., & Rahmawati, S. (2025). Evaluation of Speaking Competence in Arabic Language Learning: A Performance Assessment Approach. *IJAS: International Journal of Arabic Studies*, 2(1), 12–22. <https://doi.org/https://ojs.bustanilmu.com/index.php/IJAS/article/view/73>
- Hasim, M. F., Baihaqi, M., Sakinah, N., Humaidah, U. I., & Utama, A. S. (2026). Bridging Assessment and Proficiency: Developing Speaking Skill Test Instruments for Arabic Learners. *Tsaqofiya: Jurnal Pendidikan Bahasa dan Sastra Arab*, 7(2), 526–537. <https://doi.org/10.21154/zk6nbg28>
- Homaidah, H., & Bakar, M. Y. A. (2026). Character-Based Arabic Language Learning Assessment Innovation. *Jurnal Pembelajaran Bahasa dan Sastra*, 5(1), 913–926. <https://doi.org/10.55909/jpbs.v5i1.819>
- Imaduddin, M. F., Maulani, H., & Taufik, I. H. (2022). Test the Validity and Reliability of Arabic Learning Questions. *Arabi: Journal of Arabic Studies*, 7(2), 198–207. <https://doi.org/10.24865/ajas.v7i2.523>
- Indahwati, Y., Maksum, A., & Umihani, U. (2025). Authenticity in the Arabic Classroom:

- Evaluating Realia's Effect on Speaking Proficiency. *Lingua: Jurnal Keilmuan dan Kependidikan Bahasa Arab*, 11(1), 52–61. <https://doi.org/10.32678/lingua.v11i1.12113>
- Janah, A. N., Nur, H. W., Hasbi, M., Naziha, S., & Qomari, N. (2023). Application of Developing Test Items for the Elements of the Arabic Language at Madrasah Aliyah Level. *KITABA: Journal of Interdisciplinary Arabic Learning*, 1(2), 75–85. <https://doi.org/https://garuda.kemdiktisaintek.go.id/documents/detail/3631737>
- Kabalan, M., Bazzi, O., Al-barathie, J., Zein, O. El, Chehabeddine, L., Khabsa, J., Chaaya, M., Leon, C. M. De, & Elbejjani, M. (2025). Cognitive assessment tools for Arabic-speaking older adults: A systematic review. *Alzheimer's Dement*, 21(5), 1–31. <https://doi.org/10.1002/alz.70207>
- Kanwit, M., & Solon, M. (2022). *Communicative Competence in a Second Language: Theory, Method, and Applications*. Routledge. <https://doi.org/10.4324/9781003160779>
- Mutiatusnisa, D., Ainin, Moh., & Murtadho, N. (2025). Cultivating Metacognitive Awareness through Self-Assessment: An Instrument for Evaluating Four Language Skills in Arabic Learning. *Alibbaa': Jurnal Pendidikan Bahasa Arab*, 6(2 SE-Articles), 330–355. <https://doi.org/10.19105/ajpba.v6i2.21024>
- Rakhlin, N. V., Aljughaiman, A., & Grigorenko, E. L. (2021). Assessing language development in Arabic: The Arabic language: Evaluation of function (ALEF). *Applied Neuropsychology: Child*, 10(1), 37–52. <https://doi.org/10.1080/21622965.2019.1596113>
- Saleh, S. R., Husain, D., & M.S Lagalo, A. (2022). The design of Psychomotor Assessment Instrument for arabic language learning. *Al-Kalim: Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban*, 1(1 SE-), 70–82. <https://jurnal.daarulqimmah.org/index.php/Alkalim/article/view/11>
- Syafei, I., Ruai'ah, S., & Fauza, N. (2026). Implementation of Arabic Language Learning Evaluation in Madrasah Aliyah: a Qualitative Descriptive Study on Teachers' Assessment Practices. *Gunung Djati Conference Series, Education and Social Humanities Conference (ESHCo)*, 62, 125–132. <https://www.conferences.uinsgd.ac.id/index.php/gdcs/article/view/3291>
- Ubaidillah, M., Haddad, A. Al, Wadi, N., & Qomari, N. (2025). Digital Portfolio as an Authentic Assessment Instrument in Arabic Language Learning: A Literature Review. *Jurnal Penelitian Ilmu Pendidikan Indonesia*, 4(2), 838–846. <https://doi.org/10.31004/jpion.v4i2.486>
- Zurqoni, Z., Retnawati, H., Rahmatullah, S., Djidu, H., & Apino, E. (2020). Has Arabic Language Learning Been Successfully Implemented? *International Journal of Instruction*, 13(4), 715–730. <https://doi.org/10.29333/iji.2020.13444a>