Journal of Arabic Language and Learning in Focus
e-ISSN: 2987-3347
Volume 2, Issue. 2, 2025, pp. 58-71
https://jurnal.isimupacitan.ac.id/index.php/alif.

Analysis of the Alignment of Content Presentation and Assessment Techniques in Grade VII MTs Arabic Language Textbooks

Dennis Mutiara*[1], Ita Rodiah[2], Ahmad Fauzi[3], Hassan Mohammed Al-Shamrani[4]

[1,2,3] Sunan Kalijaga State Islamic University Yogyakarta, Indonesia.
[4]King Saud University, Saudi Arabia.
e-mail:
dnnsmutiara@gmail.com
ita.rodiah@uin-suka.ac.id
ahmad.fauzi@uin-suka.ac.id
Shamran@ksu.edu.sa

Submitted: 12-09-2029 Revised : 13-10-2026 Accepted: 31-10-2026

ABSTRACT. This study examines the alignment between material presentation and assessment techniques in Bahasa Arab MTs Kelas VII published by Indonesia's Ministry of Religious Affairs in 2020. The analysis addresses instructional sequencing, language-skill coverage, cognitive demands, and correspondence between learning activities and assessment. Findings indicate coherent progression of vocabulary, expressions, grammar, and four language skills, with strongest alignment in vocabulary and grammar. However, assessment predominantly measures foundational linguistic competence rather than independent communication. Speaking relies on controlled responses, reading emphasizes literal comprehension, and writing prioritizes completion, rearrangement, and sentence-level accuracy. Listening activities are included, but their implementation cannot be established through textbook analysis alone. Opportunities for spontaneous interaction, interpretation, functional writing, and higher-order language use remain limited. The textbook's structural coherence therefore provides a useful foundation but requires stronger communicative assessment. Revisions should incorporate scaffolded open-ended tasks, transparent performance criteria, formative feedback, and accessible audio resources to strengthen instructional assessment alignment.

Keywords: *Arabic Textbook Evaluation; Instructional Alignment; Material Presentation; Language Assessment; Communicative Competence; Madrasah Tsanawiyah.*

https://doi.org/10.52640/alif.339
How to Cite Mutiara et al.(2025). The Effect of Two Stay Two Stray Cooperative Learning Model on Mathematics Communication Ability. *Journal of Arabic Language and Learning in Focus,* Volume 2(2), 58-71.

INTRODUCTION

Textbooks occupy a strategic position in language education because they organize the linguistic input, learning activities, and assessment opportunities through which curricular expectations are translated into instructional practice. Their pedagogical value, however, cannot be inferred from the quantity of content or the inclusion of prescribed topics alone. A more substantive evaluation examines how textbook components work together to support the intended learning outcomes. Canale (2021) conceptualizes language textbooks through the interconnected dimensions of representation, interaction, and learning, highlighting the need to examine them as pedagogical artifacts rather than neutral repositories of linguistic information. Complementing this perspective, Huang et al. (2023) propose a comprehensive textbook comparison framework encompassing structure, content, expectations, and language. Together,

these perspectives provide a basis for investigating textbook quality through the relationships among its components rather than through isolated judgments of content coverage.

One important dimension of these relationships concerns the presentation of learning materials. In this study, presentation refers not merely to visual arrangement but to the organization, sequencing, contextualization, and scaffolding of content and activities. These features determine the learning opportunities made available before students are expected to demonstrate achievement. Evidence from language textbook research illustrates why such features require systematic scrutiny. Rahmat and Coxhead (2021), for example, identify substantial vocabulary demands in an Indonesian government-published EFL textbook series, demonstrating the importance of examining linguistic load rather than assuming that officially designated grade levels ensure accessibility. Similarly, Rahimi and Sayyadi (2019) report relatively high cognitive load in textbook-based listening activities. Although these studies concern English learning, they identify analytical considerations relevant to Arabic textbook evaluation: the demands of instructional materials must be examined in relation to the support provided to learners.

For beginning learners of Arabic, the relationship between content demands and instructional support warrants particular attention. Textbook evaluation needs to consider whether vocabulary, grammatical structures, and receptive and productive activities form a coherent progression toward the competencies identified in each unit. Direct evidence from Arabic education indicates that correspondence with a proficiency framework does not necessarily entail an adequate balance of learning opportunities. Risager (2021) find that Arabic textbook materials for Islamic junior high schools generally correspond to the A1–A2 levels of the Common European Framework of Reference for Languages, while their presentation remains dominated by text comprehension. Their recommendation for more contextualized materials and greater opportunities to practise language skills suggests that evaluation should distinguish the proficiency level of textbook content from the pedagogical adequacy of its presentation.

Contextual relevance also extends beyond linguistic difficulty to the ways textbooks represent social experience and cultural meaning. Risager (2021) examines language textbooks as "windows to the world," drawing attention to the perspectives through which learners encounter other people and societies. Within the Indonesian context, Isnaini et al. (2019)analyze multicultural values communicated through textbook visuals, while Samsudin et al. (2021) investigate cultural values represented in a Korean-as-a-foreign-language textbook for Indonesian learners. These studies do not establish the quality of Arabic textbooks, but they demonstrate that learning materials require analysis beyond their verbal content. For the present study, this broader perspective supports examining whether texts, illustrations, examples, and task situations provide meaningful contexts for the language use subsequently required in assessment.

Assessment constitutes another central dimension of textbook quality because it specifies the performances through which learning is made observable. Evaluating assessment therefore requires more than counting exercises or identifying their surface formats. It also entails examining the language skills, knowledge demands, and cognitive processes elicited by the tasks. In a systematic review of 41 empirical studies, Soe (2024) finds that Bloom's Taxonomy and its revised version are frequently used to analyze critical thinking in English language textbooks, while higher-order thinking receives less attention than lower-order thinking. The review also identifies limited attention to the reliability and validity of coding procedures. These findings support an assessment analysis that combines explicit task-classification criteria with transparent methodological procedures.

Nevertheless, the presence of higher-order thinking tasks should not be treated as a sufficient indicator of assessment quality. An alignment-oriented analysis must also ask whether learners have received appropriate opportunities to develop the knowledge and skills that those

tasks require. Sumardi andGuci(2023) examine the challenges and prospects of higher-order-thinking-based language assessment literacy, situating assessment design within the broader demands of language teaching. Building on this concern, the present study distinguishes task complexity from instructional coherence. A demanding assessment task may be poorly aligned if the preceding materials do not prepare learners for its requirements, whereas a relatively simple task may serve an appropriate function within an explicitly scaffolded learning sequence. The analytical priority is therefore the correspondence between learning opportunities and the evidence of achievement required.

The relationship between material presentation and assessment can consequently be examined as an internal instructional sequence. The relevant question is not simply whether a textbook includes vocabulary, grammar, four language skills, and evaluation exercises, but whether these elements are connected in a pedagogically defensible manner. For example, a unit that states a productive-language objective needs to be examined for the models, guided practice, and performance opportunities provided before the corresponding assessment. This relational approach is consistent with Canale (2021) emphasis on interaction and learning and Huang et al. (2023) multidimensional treatment of textbook design. It shifts evaluation from the presence of components to the coherence of the learning pathway those components construct.

In the Indonesian madrasah context, recent Arabic textbook research provides a relevant empirical basis for this investigation. Beyond the proficiency and presentation issues identified by Khambali et al. (2023), Amalia (2020) evaluate an Arabic textbook for junior high school using the criteria of the Education Standards, Curriculum, and Assessment Agency. Their analysis addresses content, presentation, language, and graphic design and identifies indicators requiring revision, additional material, or innovation. Taken together, these studies suggest that textbook evaluation benefits from separating distinct quality dimensions while also investigating their relationships. They further establish the relevance of examining how curricular expectations are operationalized through the material and activities provided to students.

The object of the present study is Bahasa Arab MTs Kelas VII, the 2020 Ministry of Religious Affairs textbook identified in this study as written by Faruq Baharudin and edited by Muh. Wahib Dariyadi. The investigation focuses specifically on the relationship between material presentation and assessment techniques within its learning units. Rather than treating the textbook's official publication status as evidence of pedagogical adequacy, the study examines the instructional opportunities contained in the text itself. Particular attention is directed to the progression from presented input and modelled examples to practice activities and assessment demands. This focus allows the analysis to address a more specific question than whether the textbook is generally suitable: whether students are systematically prepared for the performances they are subsequently expected to demonstrate.

The contribution of the study is therefore framed as an analytical extension rather than an unsupported claim that the selected textbook has never been investigated. Existing research has examined Arabic textbook proficiency alignment and multidimensional feasibility, while broader language textbook research has addressed vocabulary demands, cognitive load, cultural representation, and critical thinking. These lines of inquiry provide important foundations, but their findings do not, by themselves, establish the internal alignment between presentation and assessment in the selected Grade VII textbook. A focused analysis is consequently needed to determine where the textbook's instructional sequence is coherent, where assessment demands exceed the preparation provided, and where stated competencies receive insufficient opportunities for practice or demonstration.

Accordingly, this study aims to analyze the alignment between material presentation and assessment techniques in Bahasa Arab MTs Kelas VII, published by the Ministry of Religious Affairs in 2020. It examines three interconnected aspects: the organization and scaffolding of

learning materials; the language-skill coverage, task formats, and cognitive demands of assessment; and the correspondence between those assessment demands and the preceding learning opportunities. Through these dimensions, the study seeks to identify strengths, inconsistencies, and underrepresented competencies within the textbook. Its intended contribution is an explicit account of internal pedagogical coherence that can inform Arabic textbook evaluation and provide targeted recommendations for improving instructional materials for Madrasah Tsanawiyah learners.

METHOD

This study adopted a qualitative descriptive design to examine the alignment between content presentation and assessment techniques in the Grade VII Madrasah Tsanawiyah (MTs) Arabic language textbook published by the Ministry of Religious Affairs of the Republic of Indonesia in 2020. This design was selected to provide a systematic, context-sensitive description of the textbook's pedagogical features while maintaining a close connection between the findings and the documentary evidence (Doyle et al., 2020). The primary data source was Bahasa Arab MTs Kelas VII, written by Faruq Baharudin and edited by Muh. Wahib Dariyadi. Document analysis constituted the principal method of inquiry, enabling the systematic examination of the textbook as a pre-existing educational document rather than as evidence of classroom implementation (Morgan, 2022).

The analysis focused on topic sequencing, stated learning objectives, the presentation of language skills and linguistic components, learning activities, supporting exercises, and assessment tasks. Particular attention was given to the relationships among *mahārahistimā'* (listening), *mahārahkalām* (speaking), *mahārahqirā'ah* (reading), and *mahārahkitābah* (writing), together with vocabulary and grammatical content. Data were collected through repeated reading, identification, classification, and documentation of textual and visual elements relevant to these analytical dimensions. The examination of instructional explanations, activities, and task requirements was also informed by language-textbook evaluation research that distinguishes the presence of instructional content from the opportunities provided for learners to practise the targeted competence.

The data were analysed using qualitative content analysis, operationalised through three interrelated stages: data reduction, categorisation, and interpretation. During data reduction, sections relevant to material presentation and assessment were identified and retained for detailed examination. During categorisation, the selected material was organised according to content organisation, language skills, linguistic components, learning activities, assessment forms, and the correspondence between learning objectives and assessment requirements. This process followed the principle that content analysis should move beyond the collection of excerpts towards the systematic organisation of textual and visual evidence into analytically meaningful categories Kleinheksel et al. (2020). During interpretation, the categories were examined using constructive alignment as an analytical lens, with attention to the consistency among intended learning outcomes, the learning opportunities provided, and the performances required by assessment tasks Ruge et al. (2019).

Accordingly, alignment was examined not merely through shared topics but also through the correspondence between the language skills developed in the materials and those elicited by assessment activities. Analytical credibility was strengthened through repeated examination of the textbook and the checking of interpretations against their original textual contexts and relevant theoretical literature. Throughout this process, explicit textbook features were distinguished from the researchers' interpretations of their pedagogical implications, maintaining a transparent relationship between description, abstraction, and interpretation Lindgren et al. (2020). The findings were presented descriptively to identify strengths, weaknesses, and patterns of alignment between content presentation and assessment techniques. The conclusions were restricted to the

textbook's documented instructional and assessment provisions and did not extend to its effectiveness in classroom practice or its effects on student achievement.

RESULT AND DISCUSSION

Results

Organization and Presentation of Arabic Language Materials

The analysis of *Bahasa Arab MTs Kelas VII*, written by Faruq Baharudin and edited by Muh. Wahib Dariyadi, published by the Ministry of Religious Affairs of the Republic of Indonesia in 2020, indicates that the textbook presents Arabic learning materials in a relatively systematic and structured sequence. The materials are organized into thematic units that introduce students to vocabulary, expressions, grammatical elements, and language skills within contexts relevant to students' daily lives. Each unit generally moves from the introduction of vocabulary and expressions toward practice activities and assessment exercises. This structure provides a recognizable learning progression and enables students to encounter new linguistic forms before being asked to apply them in exercises. The presentation therefore demonstrates an attempt to connect linguistic knowledge with contextual Arabic language use.

The content presentation also demonstrates integration between vocabulary, language expressions, and grammatical structures. Vocabulary is generally introduced as an important foundation before students engage with sentences and communicative exercises. Grammatical material is presented in relation to the themes being studied rather than being completely separated from contextual learning. This organization can help students understand grammatical structures through examples rather than through abstract explanations alone. However, the depth of contextualization varies across the learning units, with some activities emphasizing recognition and reproduction of language forms more strongly than authentic communication. Consequently, the textbook provides a substantial linguistic foundation, although opportunities for students to independently produce meaningful Arabic communication could be expanded.

The presentation of the four language skills *mahārahistimā'* (listening), *mahārahkalām* (speaking), *mahārahqirā'ah* (reading), and *mahārahkitābah* (writing) shows an effort to develop language competence comprehensively. Reading and writing activities are relatively visible through texts, sentence-completion exercises, vocabulary tasks, and written responses. Listening and speaking activities are also incorporated, particularly through dialogues, question-and-answer exercises, and oral language practice. Nevertheless, the intensity and complexity of activities across the four skills are not entirely balanced. Some exercises remain oriented toward controlled practice, meaning that students are expected to reproduce predetermined language patterns rather than construct extended and spontaneous expressions.

The textbook further demonstrates a progression from relatively simple linguistic activities toward more complex forms of practice. Students are first introduced to vocabulary and expressions and subsequently directed to use these elements in sentences or short communicative activities. This sequence indicates an instructional pattern of introduction, practice, and application. Such progression is particularly relevant for Grade VII students because they may possess different levels of prior Arabic knowledge. The systematic organization can therefore reduce cognitive demands by allowing learners to acquire basic linguistic components before undertaking more complex tasks. However, the progression toward higher-order language production remains less prominent than activities involving identification, matching, repetition, translation, and completion.

Presentation of Learning Activities and Language Competencies

The results indicate that the textbook provides various forms of learning activities designed

to reinforce the materials presented in each unit. Activities include vocabulary recognition, sentence construction, dialogue practice, reading comprehension, writing exercises, and grammatical application. These activities generally correspond to the learning topics introduced earlier in the unit. The correspondence is particularly apparent when students are asked to apply newly introduced vocabulary in sentences or dialogues. Thus, the textbook does not present learning materials independently from practice but attempts to establish continuity between explanation and exercise.

In relation to *mahārahkalām*, the textbook provides activities that encourage students to practice short conversations and responses. These activities are useful for developing basic oral interaction because students are exposed to functional expressions and sentence patterns that can be reproduced in classroom communication. However, many speaking activities remain highly structured, with students following predetermined dialogue patterns or responding to specific prompts. The results suggest that the textbook facilitates controlled speaking practice more effectively than spontaneous speaking. Activities requiring students to negotiate meaning, express personal opinions, describe unfamiliar situations, or create their own dialogues appear comparatively limited.

For *mahārahqirā'ah*, the textbook presents reading passages accompanied by questions or comprehension exercises. The questions generally direct students to identify information contained in the text and demonstrate their understanding. This arrangement creates a clear connection between reading material and assessment. Nevertheless, reading comprehension tends to emphasize literal understanding, such as identifying words, phrases, or explicit information. More analytical activities, such as interpreting implicit meaning, evaluating information, or connecting the text with students' personal experiences, are less dominant. The results therefore indicate that the textbook is more strongly oriented toward foundational reading comprehension than advanced interpretive reading.

The development of *mahārahkitābah* is reflected in exercises requiring students to write words, complete sentences, arrange linguistic elements, and produce short written responses. These activities provide gradual practice in applying vocabulary and grammatical structures. However, writing tasks tend to emphasize accuracy of linguistic form rather than extended written communication. Students are more frequently required to complete or reproduce language structures than to develop independent paragraphs or communicative texts. This indicates that the textbook provides an appropriate foundation for beginning Arabic writing but could provide more opportunities for creative and functional writing.

The listening component is generally connected to vocabulary, expressions, or dialogues presented in the learning units. Listening activities allow students to recognize Arabic sounds, words, and expressions and subsequently respond to questions or instructions. The presence of listening practice contributes to the integration of receptive and productive language skills. However, the effectiveness of listening activities depends substantially on the availability and quality of accompanying audio materials and classroom implementation. Therefore, the textbook itself provides the instructional framework, but the achievement of listening competence may require additional media and teacher-supported activities.

Assessment Techniques in the Textbook

The analysis shows that the textbook employs several assessment techniques to measure students' understanding of the material. These include objective exercises, short-answer questions, completion tasks, matching activities, translation, sentence construction, dialogue practice, reading comprehension questions, and writing assignments. The variety of exercises demonstrates an effort to assess more than one dimension of Arabic language competence. Assessment activities are also frequently positioned after the presentation and practice of particular materials, allowing students to demonstrate their understanding of recently introduced

concepts.

The strongest alignment is observed between vocabulary and grammatical materials and the assessment techniques used to evaluate them. When students are introduced to specific vocabulary items, subsequent exercises commonly require recognition, recall, matching, or use of those vocabulary items in sentences. Similarly, grammatical concepts are generally followed by exercises requiring students to identify or apply the relevant structures. This correspondence indicates that the assessment activities are not randomly placed but are generally derived from the learning content presented in the preceding sections.

Assessment of reading competence is also relatively well aligned with the presentation of reading texts. Students are asked to answer questions based on the information provided in the texts, identify meanings, and demonstrate comprehension. This creates a direct relationship between the instructional material and its assessment. However, the assessment tends to measure students' ability to retrieve explicit information rather than their ability to critically interpret or evaluate a text. Thus, the assessment system is effective for measuring basic comprehension but provides fewer opportunities to assess higher levels of reading competence.

The assessment of speaking competence is present through dialogue and oral-response activities. These tasks enable teachers to observe students' ability to pronounce words, use expressions, and construct simple utterances. Nevertheless, the assessment techniques are generally less standardized than written exercises. The textbook provides opportunities for oral practice, but explicit criteria or rubrics for assessing pronunciation, fluency, accuracy, vocabulary use, interaction, and communicative effectiveness are not consistently prominent. Consequently, the measurement of speaking competence may depend considerably on teacher judgment and classroom implementation.

Writing assessment similarly focuses primarily on linguistic accuracy, sentence construction, and the application of vocabulary and grammar. Such assessment is useful for determining whether students have mastered basic written language structures. However, it provides relatively limited evidence of students' ability to organize ideas into meaningful and coherent Arabic texts. More authentic writing assessment, such as composing short messages, descriptions, personal narratives, or functional texts, would provide stronger evidence of communicative writing competence.

Alignment between Content Presentation and Assessment Techniques

The central finding of the analysis is that there is a generally positive alignment between the presentation of learning materials and the assessment techniques in the Grade VII MTs Arabic textbook. Most assessment activities correspond directly to the vocabulary, expressions, grammatical structures, and language skills presented within each unit. This alignment creates instructional coherence because students are generally assessed on knowledge and skills that they have previously encountered during learning. The sequence of material presentation followed by practice and assessment also provides a relatively clear learning pathway for students and teachers.

The alignment is strongest at the level of lower- and middle-order learning objectives, particularly remembering, recognizing, understanding, and applying linguistic forms. Students are frequently asked to recall vocabulary, identify grammatical structures, complete sentences, translate expressions, and apply patterns in controlled contexts. These assessment techniques appropriately measure foundational Arabic language knowledge. The results suggest that the textbook successfully establishes consistency between what is taught and what is assessed at the basic competency level.

However, the alignment becomes less comprehensive when viewed from the perspective of communicative competence and higher-order language production. Although the textbook

includes speaking, writing, reading, and listening activities, assessment techniques do not always require students to use Arabic in authentic and open-ended situations. There is a noticeable tendency toward structured exercises with predetermined answers. Consequently, the textbook's assessment system is effective in measuring linguistic knowledge and controlled language use but less effective in measuring spontaneous communication, creativity, critical interpretation, and independent language production.

Table 1 The findings can therefore be summarized in the following table:

Aspect	Content Presentation	Assessment Technique	Alignment
Vocabulary	Vocabulary related to unit themes	Matching, recall, completion, sentence use	High
Grammar	Grammatical structures supported by examples	Identification, completion, transformation, application	High
Listening	Dialogues and listening-oriented activities	Oral response and comprehension tasks	Moderate–High
Speaking	Dialogues and guided conversations	Repetition, question-answer, dialogue practice	Moderate
Reading	Thematic reading texts	Comprehension questions and information identification	High
Writing	Word and sentence construction	Written responses and sentence-production tasks	Moderate–High
Communicative competence	Functional expressions and dialogues	Mostly controlled communication tasks	Moderate
Higher-order skills	Limited analytical and creative activities	Limited open-ended assessment	Moderate–Low

Overall, the results demonstrate that the textbook has a strong structural relationship between instructional content and assessment, particularly for vocabulary, grammar, and basic receptive language skills. The main limitation lies not in the absence of assessment but in the dominant use of controlled assessment formats. More authentic performance-based assessments, project tasks, portfolios, open-ended speaking activities, and contextual writing tasks could strengthen the relationship between the textbook's learning objectives and actual communicative competence. The textbook consequently provides a solid foundation for Grade VII Arabic learning, but its assessment component could be further developed to capture broader dimensions of Arabic language proficiency.

Discussion

The findings indicate that Bahasa Arab MTs Kelas VII, published by the Indonesian Ministry of Religious Affairs in 2020, demonstrates a generally coherent relationship among material presentation, learning activities, and assessment. Vocabulary, expressions, grammatical structures, and language skills are organized through a progression from introduction to controlled practice and subsequent application. This organization provides a plausible instructional foundation, although sequential presentation alone does not establish whether students retain and transfer their learning. Kim and Webb (2022) meta-analysis demonstrates that the effectiveness of repeated practice depends partly on its distribution over time, particularly when learning is assessed through delayed tests. Accordingly, the textbook's structural coherence should be understood as an enabling condition for learning rather than evidence of sustained language development. Its instructional value could be strengthened by systematically revisiting previously introduced language across later units and different communicative contexts.

The integration of vocabulary, expressions, and grammatical structures constitutes a further strength because it creates opportunities to connect linguistic forms with thematic meanings. Nevertheless, thematic organization does not necessarily ensure sufficiently frequent

or varied encounters with target language. Uchihara et al. (2019) identified a positive relationship between repetition and incidental vocabulary learning, while also demonstrating that this relationship varies according to learner characteristics and instructional conditions. Complementing this finding, Webb et al. (2020) showed that gains from intentional vocabulary-learning activities differ considerably across tasks and decline on delayed measures. These findings suggest that the textbook's vocabulary provision should be evaluated not only in terms of thematic relevance and initial coverage, but also through opportunities for retrieval, contextual reuse, and productive deployment. Vocabulary exercises therefore need to extend beyond matching and completion toward meaningful sentence construction and communication across units.

The visible representation of mahārah*istimā', mahārahkalām, mahārahqirā'ah,* and *mahārahkitābah* suggests that the textbook aims to support a multidimensional conception of language competence. However, the inclusion of four skill labels should not be equated with their functional integration. Integration requires learners to use information obtained through one skill to accomplish a meaningful purpose through another. meta-analysis of incidental vocabulary learning through meaning-focused input offers a useful basis for considering how reading and listening can contribute to lexical development, without assuming that input exposure guarantees mastery. In the present textbook, stronger integration could involve listening to a description, identifying relevant information in a short text, discussing its meaning, and producing a brief written response. Such sequences would establish clearer relationships among receptive input, linguistic development, and purposeful output.

The predominance of controlled speaking exercises represents an important limitation in this progression. Dialogues, repetition, guided questions, and predetermined responses provide manageable opportunities for initial production, but they offer limited evidence of learners' capacity to formulate meanings independently meta-analysis reports positive overall outcomes for task-based language teaching, supporting the inclusion of activities organized around meaningful task completion rather than the reproduction of predetermined forms alone. A relevant Indonesian comparison is provided by, whose study identified differences in speaking opportunities and assessment practices across school contexts, including reliance on vocabulary memorization and workbook-based activities. Although their research concerned English rather than Arabic, it helps illuminate the distinction between practising oral forms and creating opportunities for communicative participation. The present textbook would therefore benefit from carefully scaffolded information-gap activities, interviews, and role plays that retain familiar vocabulary while allowing variation in responses.

The reading component demonstrates clear correspondence between the texts presented and the questions used to assess comprehension. This relationship supports instructional transparency because assessment draws on material that students have encountered. Nevertheless, the prominence of explicit information retrieval limits the range of comprehension processes elicited. Villarroel et al. (2020) argue that conventional written tests can be redesigned through realism, cognitive challenge, and evaluative judgement, rather than remaining predominantly decontextualized measures of lower-order performance. Applied to the present findings, this perspective suggests complementing literal questions with accessible inference, comparison, justification, and contextual decision-making tasks. Importantly, greater cognitive demand need not require substantially more advanced Arabic: students can compare two short descriptions, explain a character's choice using familiar expressions, or select an appropriate response from information distributed across a text.

The writing activities similarly demonstrate progression from word-level responses toward sentence construction and short written production. This sequence provides structured opportunities to consolidate linguistic resources, but the predominance of completion, rearrangement, and reproduction restricts what the resulting assessment can reveal about

independent meaning-making. Moradian et al. (2021), drawing on a sociocultural perspective, examine dialogic interaction in EFL writing assessment, offering a relevant framework for considering assessment as an interactive learning process rather than solely a judgement of finished products. Lee (2020) likewise questions the straightforward transfer of corrective-feedback research into authentic writing classrooms. Together, these perspectives support extending the textbook's writing provision through short communicative products, guided feedback, and revision. Personal profiles, brief descriptions, and simple messages would enable students to organize meanings while continuing to receive support for linguistic accuracy.

Feedback design is particularly relevant to this proposed extension. Expanding writing tasks should not simply increase the number of errors teachers are expected to correct. Mao and Lee (2020) synthesis identifies feedback scope as a substantive issue in written corrective-feedback research, indicating that decisions about what to address require deliberate instructional consideration. For the present textbook, feedback could therefore prioritize a small set of criteria appropriate to each task, such as successful communication of required information, organization, target vocabulary, and selected grammatical features. This would make the relationship between learning objectives, student production, and teacher response more explicit. It would also prevent functional writing from being reduced to an expanded grammar exercise in which communicative success remains secondary to exhaustive error identification.

The listening component requires a comparable distinction between the presence of activities and the quality of the learning opportunities they provide. The textbook establishes a framework for listening practice, but analysis of its printed content cannot independently establish the accessibility of audio materials, the nature of classroom delivery, or students' listening development. In'nami et al. (2023) found a relationship between metacognitive awareness and second-language listening comprehension, while also showing that metacognitive awareness explains only part of the variation in listening performance. This finding supports a cautious interpretation: effective listening development cannot be attributed to a single instructional feature. The textbook could be supplemented with accessible recordings and staged activities that guide learners to anticipate content, identify key information, check understanding, and connect what they hear with a subsequent speaking or writing task. These are proposed improvements rather than outcomes demonstrated by the present document analysis.

The strongest instructional–assessment correspondence appears in vocabulary and grammar, where matching, identification, completion, transformation, translation, and sentence-construction tasks directly reflect previously introduced content. Nevertheless, correspondence between taught forms and test items represents only one dimension of assessment quality. A broader interpretation requires attention to whether assessment supports students in understanding quality, monitoring performance, and improving subsequent work. Panadero et al. (2019) explain how evaluative judgement can connect formative assessment with self- and co-regulated learning, while Andrade (2019) review highlights the formative role of student self-assessment. These perspectives suggest that the textbook's existing exercises could become more instructionally productive when accompanied by explicit success criteria, brief reflection prompts, and opportunities to revise responses. Assessment would then support both the demonstration of current knowledge and the regulation of further learning.

The moderate alignment identified in speaking assessment further illustrates why task format and scoring criteria should be considered together. Repetition and predetermined dialogues reveal some aspects of pronunciation and linguistic control, but provide restricted opportunities to observe independent expression and interaction. Performance tasks would broaden the evidence available, although their inclusion would require transparent assessment criteria. Panadero and Jonsson (2020) emphasize that rubrics need careful design and implementation rather than being treated as universally effective instruments. Panadero et al. (2023) meta-analysis additionally reports a positive, moderate effect of rubric use on academic

performance, with smaller effects on self-regulation and self-efficacy. For this textbook, a concise analytic speaking rubric could distinguish task fulfilment, comprehensibility, interaction, and control of target language. Such criteria would clarify what successful performance entails without requiring beginners to demonstrate advanced proficiency.

The limited representation of higher-order assessment should likewise be interpreted in relation to learners' language resources rather than treated as a requirement to replace foundational practice entirely. Remembering, understanding, and applying remain relevant to introductory language learning; the concern is whether students also receive appropriately supported opportunities to interpret, justify, compare, and create. Sumardi andGuci(2023) investigation of HOTS-based language assessment literacy highlights the importance of teachers' capacity to translate higher-order objectives into assessment practice. Consequently, textbook revision should connect intended cognitive processes with feasible linguistic demands and clear scoring criteria. For Grade VII learners, this might involve selecting and explaining an appropriate expression, comparing short profiles, or creating a brief response to a familiar situation. These tasks would extend cognitive engagement while maintaining continuity with the vocabulary and structures already taught.

The implications for authentic assessment should also remain context-sensitive. Sokhanvar et al. (2021) identify benefits of authentic assessment for learning experiences and employability-related skills in higher education, while Vlachopoulos and Makri (2024) emphasize both its potential and the implementation challenges associated with resources, training, and stakeholder participation. These findings do not demonstrate equivalent effects among Grade VII Arabic learners, but they provide a defensible rationale for adapting authenticity to learners' immediate communicative environments. In the present context, authenticity could involve introducing oneself, asking for familiar information, describing classroom objects, or composing a short message. The priority is not to imitate complex professional tasks, but to create credible purposes, audiences, and outcomes for using the language being learned.

Overall, the textbook's principal contribution lies in its structural coherence and provision of foundational linguistic practice, whereas its main developmental need concerns the transition from controlled language manipulation to purposeful communication and broader assessment evidence. Revision should therefore preserve the existing progression while strengthening connections among input, interaction, production, feedback, and assessment criteria. More open-ended tasks, short performance assessments, guided revision, learner reflection, and accessible listening resources would address this need without discarding the textbook's current instructional strengths. Finally, because the present study examines textbook content, its conclusions concern the learning opportunities represented in the materials, not demonstrated effects on students' communicative proficiency. Classroom observation, analysis of student work, and performance-based assessment would be needed to determine how those opportunities are implemented and with what outcomes.

CONCLUSION

This study identifies a generally coherent relationship between material presentation and assessment techniques in Bahasa Arab MTs Kelas VII, published by the Ministry of Religious Affairs in 2020. Vocabulary, expressions, grammatical structures, and language skills are presented through a gradual progression from introduction to practice and application. The strongest correspondence occurs in vocabulary and grammar, where assessment largely measures linguistic content that learners have previously encountered. This organization provides a structured foundation for introductory Arabic instruction and represents an important pedagogical strength of the textbook.

Nevertheless, alignment is uneven across the competencies assessed. The close

relationship between instructional content and controlled exercises does not consistently extend to independent communicative performance. Speaking assessment offers limited opportunities for spontaneous interaction, reading tasks predominantly assess literal comprehension, and writing activities emphasize the reproduction and manipulation of linguistic forms rather than the independent organization of meaning. Likewise, the presence of listening activities establishes an intended instructional component but does not, by itself, demonstrate the availability of suitable audio resources or the quality of classroom implementation. The central limitation is therefore not the absence of instructional components, but the restricted progression from foundational practice toward purposeful, varied, and cognitively engaging language use.

The study contributes an account of internal pedagogical coherence that distinguishes component coverage from the quality of the relationships among objectives, materials, practice, and assessment. Textbook improvement should preserve systematic sequencing while expanding appropriately scaffolded communicative tasks, interpretive reading, functional writing, learner reflection, and performance-based assessment. Clear success criteria and opportunities for feedback and revision would further strengthen the connection between learning activities and evidence of achievement. Such improvements should remain proportionate to Grade VII learners' linguistic resources, ensuring that increased communicative and cognitive demands are supported by adequate preparation.

Finally, these conclusions are limited to the instructional opportunities represented in the textbook. They do not establish how teachers implement the materials or how effectively students develop language proficiency through their use. Further research incorporating classroom observation, teacher and student perspectives, student work, and language-performance evidence is needed to examine how the identified alignment operates in practice.

REFERENCE

Amalia, A. (2020). Ice Breaking dalamPembelajaran Bahasa Arab. *Shaut al Arabiyyah*, *8*(1), 75. https://doi.org/10.24252/saa.v8i1.11551

Andrade, H. L. (2019). A Critical Review of Research on Student Self-Assessment. *Frontiers in Education*, *4*, 87. https://doi.org/10.3389/feduc.2019.00087

Canale, G. (2021). The language textbook: Representation, interaction and learning. *Language, Culture and Curriculum*. https://doi.org/10.1080/07908318.2020.1797080

Doyle, L., McCabe, C., Keogh, B., Brady, A., & McCann, M. (2020). An overview of the qualitative descriptive design within nursing research. *Journal of Research in Nursing*, *25*(5), 443–455. https://doi.org/10.1177/1744987119880234

Huang, A. Y. Q., Lu, O. H. T., & Yang, S. J. H. (2023). Effects of artificial Intelligence–Enabled personalized recommendations on learners' learning engagement, motivation, and outcomes in a flipped classroom. *Computers & Education*, *194*, 104684. https://doi.org/10.1016/j.compedu.2022.104684

In'nami, Y., Cheung, M. W. -L., Koizumi, R., & Wallace, M. P. (2023). Examining Second Language Listening and Metacognitive Awareness: A Meta-Analytic Structural Equation Modeling Approach. *Language Learning*, *73*(3), 759–798. https://doi.org/10.1111/lang.12548

Isnaini, Setyono, &Ariyanto. (2019). A visual semiotic analysis of multicultural values in an Indonesian English textbook. *Indonesian Journal of Applied Linguistics*. https://doi.org/10.17509/ijal.v8i3.15253

Khambali, K., Asrori, I., Rosyidi, A. W., &Nurhasanah, N. (2023). Analysis of Teaching Materials for Arabic Textbooks for Islamic Junior High School Based on CEFR Standards. *Arabiyat :Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban*, *10*(1), 88–105. https://doi.org/10.15408/a.v10i1.30439

Kim, S. K., & Webb, S. (2022). The Effects of Spaced Practice on Second Language Learning: A Meta-Analysis. *Language Learning*, *72*(1), 269–319. https://doi.org/10.1111/lang.12479

Kleinheksel, A. J., Rockich-Winston, N., Tawfik, H., & Wyatt, T. R. (2020). Demystifying Content Analysis. *American Journal of Pharmaceutical Education*, *84*(1), 7113. https://doi.org/10.5688/ajpe7113

Lee, I. (2020). Utility of focused/comprehensive written corrective feedback research for authentic L2 writing classrooms. *Journal of Second Language Writing*, *49*, 100734. https://doi.org/10.1016/j.jslw.2020.100734

Lindgren, B.-M., Lundman, B., &Graneheim, U. H. (2020). Abstraction and interpretation during the qualitative content analysis process. *International Journal of Nursing Studies*, *108*, 103632. https://doi.org/10.1016/j.ijnurstu.2020.103632

Moradian, M. R., Miri, M., & Alamdar, P. (2021). The role of dialogic interaction in EFL writing assessment: A sociocultural perspective. *Indonesian Journal of Applied Linguistics*, *10*(3). https://doi.org/10.17509/ijal.v10i3.31754

Morgan, H. (2022). Conducting a Qualitative Document Analysis. *The Qualitative Report*. https://doi.org/10.46743/2160-3715/2022.5044

Panadero, E., Broadbent, J., Boud, D., & Lodge, J. M. (2019). Using formative assessment to influence self- and co-regulated learning: The role of evaluative judgement. *European Journal of Psychology of Education*, *34*(3), 535–557. https://doi.org/10.1007/s10212-018-0407-8

Panadero, E., & Jonsson, A. (2020). A critical review of the arguments against the use of rubrics. *Educational Research Review*, *30*, 100329. https://doi.org/10.1016/j.edurev.2020.100329

Panadero, E., Jonsson, A., Pinedo, L., & Fernández-Castilla, B. (2023). Effects of Rubrics on Academic Performance, Self-Regulated Learning, and self-Efficacy: A Meta-analytic Review. *Educational Psychology Review*, *35*(4), 113. https://doi.org/10.1007/s10648-023-09823-4

Rahimi & Sayyadi. (2019). The cognitive load of listening activities of a cognitive-based listening instruction. *Indonesian Journal of Applied Linguistics*. https://doi.org/10.17509/ijal.v9i2.20236

Rahmat & Coxhead. (2021). Investigating vocabulary coverage and load in an Indonesian EFL textbook series. *Indonesian Journal of Applied Linguistics*. https://doi.org/10.17509/ijal.v10i3.31768

Risager, K. (2021). Language textbooks: Windows to the world. *Language, Culture and Curriculum*. https://doi.org/10.1080/07908318.2020.1797767

Ruge, G., Tokede, O., &Tivendale, L. (2019). Implementing constructive alignment in higher education – cross-institutional perspectives from Australia. *Higher Education Research & Development*, *38*(4), 833–848. https://doi.org/10.1080/07294360.2019.1586842

Samsudin, Ansas, &Triarisanti. (2021). The representation of cultural values in Korean as a foreign language (KFL) textbook. *Indonesian Journal of Applied Linguistics*. https://doi.org/10.17509/ijal.v10i3.31749

Soe. (2024). A Systematic Literature Review of Textbook Evaluation Studies. *Theory and Practice of Second Language Acquisition*.

Sokhanvar, Z., Salehi, K., &Sokhanvar, F. (2021). Advantages of authentic assessment for improving the learning experience and employability skills of higher education students: A systematic literature review. *Studies in Educational Evaluation*, *70*, 101030. https://doi.org/10.1016/j.stueduc.2021.101030

Sumardi, S., &Guci, R. I. (2023). HOTS-based language assessment literacy: Challenges and prospects in English language teaching. *Indonesian Journal of Applied Linguistics*, *12*(3), 831–840. https://doi.org/10.17509/ijal.v12i3.44261

Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The Effects of Repetition on Incidental Vocabulary Learning: A Meta-Analysis of Correlational Studies. *Language Learning*, *69*(3),

559–599. https://doi.org/10.1111/lang.12343

Villarroel, V., Boud, D., Bloxham, S., Bruna, D., & Bruna, C. (2020). Using principles of authentic assessment to redesign written examinations and tests. *Innovations in Education and Teaching International*, *57*(1), 38–49. https://doi.org/10.1080/14703297.2018.1564882

Vlachopoulos, D., & Makri, A. (2024). A systematic literature review on authentic assessment in higher education: Best practices for the development of 21st century skills, and policy considerations. *Studies in Educational Evaluation*, *83*, 101425. https://doi.org/10.1016/j.stueduc.2024.101425

Webb, S., Yanagisawa, A., & Uchihara, T. (2020). How Effective Are Intentional Vocabulary-Learning Activities? A Meta-Analysis. *The Modern Language Journal*, *104*(4), 715–738. https://doi.org/10.1111/modl.12671