



## An Evaluation of the Assessment System in the 7th-Grade MTS Arabic Textbook Based on the Principles of Henry Douglas Brown

Arini Lutfi Solihah<sup>1\*</sup>, Nasokah<sup>2</sup>, Ashief El Qorny<sup>2</sup>

Program Studi Pendidikan Bahasa Arab, Fakultas Ilmu Tarbiyah dan Keguruan,  
Universitas Sains AlQur'an Jawa Tengah di Wonosobo

\*Penulis Korespondensi: [arinilutfisolihah17@gmail.com](mailto:arinilutfisolihah17@gmail.com)

**Abstract.** *This study aims to evaluate the assessment system in the Arabic textbook for 7th-grade MTs, Semester 2, published by the Ministry of Religious Affairs, based on the principles of evaluation according to H. Douglas Brown. The evaluation principles used include practicality, reliability, validity, authenticity, and washback. This study employs a qualitative approach using a descriptive research design through library research. The primary data source in this study is the Arabic textbook for Grade 7 MTs, Semester 2, while secondary data was obtained from various relevant literature. Data collection was conducted through documentary analysis, while data analysis utilized content analysis. The results indicate that the evaluation system in the textbook is quite good and systematic. Assessments are presented in both formative and summative formats and include a variety of question types, such as multiple-choice, short-answer, matching, translation, sentence construction, true-false, and fill-in-the-blank. The assessment also covers the four language skills: listening, speaking, reading, and writing. Based on Brown's principles, the assessment in this textbook meets the criteria of practicality, validity, and authenticity. The reliability aspect shows good potential but remains dependent on the teacher's consistency in administration and scoring. Meanwhile, the washback aspect has the potential to positively impact learning, although this still requires validation through field research. Thus, the Arabic textbook for Grade 7 MTs, Semester 2, is suitable for use as teaching material and a learning assessment tool; however, it still requires refinement in the scoring guidelines, teaching methods, and the strengthening of communicative exercises.*

**Keywords:** *Evaluation, assessment system, Arabic textbook, H. Douglas Brown, MTs.*

**Abstrak.** Penelitian ini bertujuan untuk mengevaluasi sistem penilaian dalam buku ajar Bahasa Arab Kelas VII MTs Semester 2 terbitan Kementerian Agama ditinjau dari prinsip-prinsip evaluasi menurut H. Douglas Brown. Prinsip evaluasi yang digunakan meliputi kepraktisan, reliabilitas, validitas, autentisitas, dan washback. Penelitian ini menggunakan pendekatan kualitatif dengan jenis penelitian deskriptif melalui kajian kepustakaan (*library research*). Sumber data primer dalam penelitian ini adalah buku ajar Bahasa Arab Kelas VII MTs Semester 2, sedangkan data sekunder diperoleh dari berbagai literatur yang relevan. Teknik pengumpulan data dilakukan melalui studi dokumentasi, sedangkan teknik analisis data menggunakan analisis isi (*content analysis*). Hasil penelitian menunjukkan bahwa sistem evaluasi dalam buku ajar tersebut cukup baik dan sistematis. Evaluasi disajikan dalam bentuk formatif dan sumatif, serta memuat variasi bentuk soal, seperti pilihan ganda, uraian singkat, menjodohkan, menerjemahkan, menyusun kalimat, benar-salah, dan mengisi kalimat rumpang. Evaluasi juga mencakup empat keterampilan berbahasa, yaitu menyimak, berbicara, membaca, dan menulis. Ditinjau dari prinsip Brown, evaluasi dalam buku ajar ini telah memenuhi aspek kepraktisan, validitas, dan autentisitas. Aspek reliabilitas menunjukkan potensi yang baik, tetapi masih bergantung pada konsistensi guru dalam pelaksanaan dan penskoran. Sementara itu, aspek washback berpotensi memberikan dampak positif terhadap pembelajaran, meskipun masih memerlukan pembuktian melalui penelitian lapangan. Dengan demikian, buku ajar Bahasa Arab Kelas VII MTs Semester 2 layak digunakan sebagai bahan ajar dan sarana evaluasi pembelajaran, namun masih memerlukan penyempurnaan pada pedoman penskoran, metode pembelajaran, dan penguatan latihan komunikatif.

**Kata kunci:** Evaluasi; Sistem penilaian; Buku ajar; bahasa arab; H. Douglas Brown; MTs.

## **1. LATAR BELAKANG**

Learning is an active process of acquiring skills, knowledge, and positive attitudes to bring about positive personal change. One key component in improving the quality of the learning process and outcomes is the use of textbooks. Textbooks serve as the primary learning resource for students and as a guide for teachers in designing and implementing instruction aligned with the curriculum (Deepublish, 2022). Textbooks serve as a reference for students in independent learning and assist teachers in organizing systematic instruction. A good textbook is systematically and structurally organized, making it easier for students to understand the material on their own. Ali Mudlofir states that a textbook is material systematically organized in both written and non-written forms.

Meanwhile, Bacon defines a textbook as a book designed for classroom use, written by experts in their field, and equipped with appropriate learning resources (Deepublish, 2022). Thus, a good textbook includes at least three main sections: the introduction, the main body, and the conclusion. In addition to presenting material, textbooks also provide assessment tools to measure students' learning outcomes. Assessment is the stage that follows measurement, while evaluation is the decision-making process based on the results of that assessment (Faiz et al., 2022).

Evaluation is conducted systematically and continuously to collect, analyze, and interpret data regarding the learning process and outcomes. In the context of school education, teachers serve not only as motivators but also as planners and implementers of systematic and effective learning. The learning process encompasses planning, implementation, evaluation, and follow-up. Through evaluation, teachers can determine students' level of understanding and identify appropriate follow-up steps (Susanto, 2023). The success of evaluation is heavily influenced by the assessment tools used. One commonly used assessment tool is a test, particularly in the form of objective questions such as multiple-choice items (Muslimah & Widiyanti, 2023).

Therefore, evaluation instruments must be carefully designed to accurately reflect students' abilities. To achieve high-quality evaluation, certain principles must be followed in their development. Henry Douglas Brown identified five principles of evaluation: practicality, reliability, validity, authenticity, and washback (Solihin & Yusuf, 2024). These principles serve as benchmarks for assessing the quality of evaluation in learning. However, in practice, it is not yet known to what extent these principles have been applied in the 7th-grade Arabic textbook published by the Ministry of Religious Affairs and used at MTs N 2 Wonosobo.

Previous research indicates that there remains a lack of clarity regarding the application of evaluation and development research methods (Wandana, 2023). Based on this, this study focuses on evaluating the assessment system in the 7th-grade Arabic textbook at MTs N 2 Wonosobo using Henry Douglas Brown's principles of evaluation. This study aims to determine the extent to which the assessment system in the textbook meets the criteria for good evaluation and supports the achievement of learning objectives

## **2. KAJIAN TEORITIS**

### **1. Arabic Language Textbooks and Learning Assessment**

Textbooks are a key component of the learning process, serving as a learning resource for students and a guide for teachers in conducting lessons. A good textbook is systematically organized, aligned with the curriculum, and includes exercises or assessments to measure the achievement of learning objectives. In Arabic language

learning, textbooks must include the four language skills—listening, speaking, reading, and writing so that students can both understand and actively use the language.

Learning assessment is a systematic process to determine students' level of achievement regarding learning objectives. Assessment involves tests, measurements, and grading. Tests are used as tools to obtain data; measurements produce numerical data; assessments interpret measurement results; while evaluation is used to make decisions based on assessment results (Faiz et al., 2022). In this article, the 7th-grade MTs Arabic textbook for Semester 2 includes formative and summative evaluations and covers the four language skills.

## 2. Principles of Assessment According to H. Douglas Brown

H. Douglas Brown identifies five principles of language assessment: practicality, reliability, validity, authenticity, and washback (Brown, 2004). Practicality refers to the ease of administering the assessment in terms of time, cost, administration, and scoring. Reliability concerns the consistency of assessment results. Validity indicates the extent to which the evaluation truly measures the ability it is intended to measure.

Furthermore, authenticity concerns the alignment of evaluation tasks with real-world, meaningful language use contexts. Washback, meanwhile, refers to the impact of the assessment on the learning process, for both teachers and students (Brown, 2004; Brown & Abeywickrama, 2010). These five principles serve as the basis for evaluating the quality of the assessment system in the 7th-grade MTs Arabic textbook for Semester 2.

## 3. Relevant Previous Research

Several previous studies have shown that evaluation plays a crucial role in determining the quality of learning. Faiz et al. explain the importance of understanding tests, measurement, assessment, and evaluation in education (Faiz et al., 2022). Susanto emphasizes that evaluation tools must be designed appropriately to objectively measure students' abilities (Susanto, 2023). Meanwhile, Muslimah and Widiyanti highlight that analyzing Arabic language learning achievement tests is essential to assess the quality of the evaluation instruments used (Muslimah & Widiyanti, 2023).

Based on these theories and prior research, this study aims to analyze the assessment system in the 7th-grade MTs Arabic textbook for Semester 2 using H. Douglas Brown's evaluation principles. Implicitly, this study is grounded in the assumption that the evaluation system in the textbook needs to be examined to determine its alignment with sound evaluation principles and its support for achieving learning objectives.

### 3. METODE PENELITIAN

This study employs a qualitative approach using a descriptive research design through literature review (library research). A literature review is a method that involves examining various written sources such as books, journals, articles, and

other documents relevant to the research topic. This method aims to obtain theoretical data and information to support the research analysis (Rita Kumala Sari, 2021; Milya Sari & Asmendri, 2020). Additionally, a literature review enables researchers to understand concepts, theories, and the results of previous studies, thereby providing a strong foundation for conducting the research (Aqil, 2020).

The purpose of this study is to describe the assessment system in the 7th-grade MTs Arabic textbook and to analyze its alignment with the principles of evaluation proposed by Henry Douglas Brown. The choice of the library research method was based on the primary data source, which consists of written documents, namely the 7th-grade MTs Arabic textbook published by the Ministry of Religious Affairs. This study falls under the category of qualitative research. Qualitative research is a type of research that produces descriptive data in the form of written or spoken words from the observed objects and does not use statistical procedures (Safarudin & Zulfamanna, 2023). This approach is used to understand phenomena in depth based on narrative data (Wahidmurni, 2017).

Therefore, the primary data in this study consists of textbook content that is analyzed in depth. Additionally, this study is descriptive in nature, meaning it aims to systematically and factually describe the phenomenon under investigation. Descriptive research is used to explore and clarify a social phenomenon or phenomenon that occurs (Syahrizal & Jailani, 2023). In this context, the study focuses on describing the assessment system in the 7th-grade MTs Arabic textbook.

The data sources in this study consist of two types: primary data sources and secondary data sources. Primary data is the main source used in this study, specifically the 7th-grade MTs Arabic textbook published by the Ministry of Religious Affairs. Meanwhile, secondary data consists of supporting data obtained from various sources such as books, journals, and scientific articles relevant to the research topic (Haifa et al., 2025).

The data collection technique used in this study is a documentary study. The documentary technique involves collecting and analyzing various written documents related to the research object. This technique aims to obtain valid and systematic data to address the research problem (Romdona et al., 2025; Ardiansyah et al., 2023). In this study, the researcher identified sections in the textbook that contain assessment systems, such as practice questions, assignments, and other forms of evaluation.

The data analysis technique used is content analysis. Content analysis is a research technique used to draw valid and replicable conclusions from a text based on the context of its use (Susanto et al., 2022). Qualitative data analysis generally aims to organize data into categories, patterns, and units of description so that it can be interpreted systematically (Qomaruddin & Sa'diyah, 2024; Jailani & Saksitha, 2024). The data analysis steps in this study include:

1. determining the focus of the analysis, namely the assessment system in the 7th-grade MTs Arabic textbook;
2. collecting and identifying data in the form of textbook sections containing assessment components such as practice questions, assignments, and exams;
3. classifying the data based on Henry Douglas Brown's principles of evaluation, namely practicality, reliability, validity, authenticity, and washback;
4. analyzing and interpreting the data by comparing the assessment system in the textbook against these principles; and
5. draw conclusions regarding the level of alignment between the assessment

system in the textbook and sound evaluation principles.

6. Thus, through this method, it is hoped that a clear and comprehensive picture of the quality of the assessment system in the 7th-grade MTs Arabic textbook can be obtained, as viewed through Henry Douglas Brown's evaluation principles.

#### **4. HASIL DAN PEMBAHASAN**

##### **A. Textbook Profile**

The textbook analyzed in this study is the Arabic Language Textbook for 7th Grade MTs, Second Semester. This textbook is one of the instructional materials used to support the Arabic language learning process at the Madrasah Tsanawiyah level, specifically for 7th grade in the second semester. This book was compiled by the Arabic Language MGMP Compilation Team of the Ministry of Religious Affairs of Wonosobo Regency for the 2022–2025 period and published on November 1, 2023. The existence of this textbook demonstrates a systematic effort by the compilers to provide teaching materials that meet the needs of Arabic language learning in madrasahs.

Substantively, the development of this textbook is based on the current curriculum policy, namely the Decision of the Director General of Islamic Education No. 3211 of 2022 regarding Learning Outcomes for Islamic Religious Education and Arabic Language in the Merdeka Curriculum at Madrasahs. Thus, this book serves not only as a source of learning materials but also as a form of educational policy implementation that aligns the direction of learning with the principles of the Merdeka Curriculum (Director General of Islamic Education, 2022)

Within the context of the Merdeka Curriculum, this textbook includes learning outcomes that cover the four language skills: listening, speaking, reading, and writing. These four skills form a crucial foundation for Arabic language learning, as language proficiency involves not only comprehension but also the active and communicative use of the language. Additionally, this book integrates elements of the Pancasila Student Profile, Learning Objectives, and the Learning Objective Sequence as part of the instructional planning. This integration demonstrates that the textbook is designed not only to develop language competencies but also to foster the holistic character development of students.

The material in this textbook is organized into three main chapters: al-Darsu al-Awwalu: al-‘Unwān (The Title), al-Darsu al-Thānī: al-Baytu (The House), and al-Darsu al-Thālīth: Min Yaumiyyāti al-Urah (From the Family Diary). This structure indicates that the material is organized progressively according to the developmental level of seventh-grade students. A tiered approach to organizing material is essential in language learning because students require structured stages to understand linguistic elements in a continuous manner.

This textbook is also designed to help students understand the material in line with the learning objectives and to encourage them to be active and creative in seeking and discovering the knowledge being studied. Thus, students are not merely positioned as recipients of information, but also as active participants in the learning process. This is in line with the spirit of the Merdeka Curriculum, which emphasizes student-centered learning

In addition, this textbook emphasizes that Arabic can be used as a means of communication, both actively through speech and passively through writing, in accordance with the students' developmental level. This statement indicates that Arabic language learning in this textbook is focused on the practical function of language as a means of communication. However, the learning approach or method used in the textbook

has not been explained in detail. This is an important point to note because the learning method plays a significant role in determining the success of classroom instruction.

Overall, the Arabic textbook for 7th-grade MTs, Semester 2, can be understood as a teaching resource developed with a clear curriculum orientation, containing structured learning outcomes, and designed to support the gradual development of students' Arabic language competencies. This book plays a strategic role in helping teachers conduct instruction in accordance with the requirements of the Merdeka Curriculum, although the aspects of teaching approaches and methods still need to be explained in greater detail.

### **B. Description of Assessment in the Textbook**

The evaluation system in the Arabic textbook for 7th-grade MTs, Semester 2, is presented in various forms of assessment distributed throughout each learning activity. This evaluation not only serves to measure student learning outcomes but is also an integral part of the ongoing, step-by-step learning process. Generally, the evaluations in this textbook can be classified into formative and summative evaluations.

Formative assessment is an evaluation conducted at the beginning of or during the learning process. The purpose of formative assessment is to provide information or feedback to teachers and students to improve the learning process, identify learning needs, and track students' progress over time. In this textbook, formative assessment is evident in every learning activity that includes several exercises. These exercises are designed in various formats to support continuous mastery of the material.

In addition to formative assessments, this textbook also includes summative assessments. Summative assessments are evaluations conducted to determine the overall achievement of learning objectives. This assessment is typically administered at the end of the learning process, at the end of a unit, or at the end of a chapter. In this textbook, summative assessments are found in the final assessment section at the end of each chapter. These assessments serve to measure students' mastery of the material covered in its entirety.

The inclusion of formative and summative assessments demonstrates that the assessment system in this textbook is not solely focused on final outcomes but also takes into account the students' learning process. Formative assessment serves as a tool for diagnosing and improving learning, while summative assessment serves to evaluate overall learning achievement. With these two forms of assessment, this textbook can be said to facilitate learning that is more focused, measurable, and sustainable

The types of assessments included in this textbook are also quite diverse. These assessment formats include multiple-choice questions, short-answer questions, matching exercises, sentence completion based on examples, vocabulary sequencing, translation, true-false questions, and fill-in-the-blank exercises. This diversity in assessment formats demonstrates the textbook authors' effort to evaluate students' abilities across various aspects

Multiple-choice questions are used to assess students' ability to understand concepts and identify the correct answers. This form of assessment helps evaluate students' basic understanding of the material they have studied. Meanwhile, short-answer questions are designed to assess students' ability to express their understanding in writing. Through this format, students are trained to answer questions independently rather than simply selecting from provided options

The matching assessment format is designed to train students' accuracy in linking concepts, vocabulary, or appropriate answer pairs. The sentence-construction assessment based on examples is used to evaluate students' ability to apply the sentence patterns they

have learned. The vocabulary sequencing assessment aims to train students to understand the correct word order in Arabic. The translation assessment is used to measure students' ability to understand meaning and to translate from Arabic to another language or vice versa. Meanwhile, the sentence completion exercise serves to measure students' understanding of vocabulary, language structure, and sentence context.

The diversity of these assessment formats demonstrates that the evaluations in this textbook do not focus solely on a single aspect of proficiency. Multiple-choice and true-false questions emphasize conceptual understanding, while short essays, sentence construction, and translation tasks focus on the ability to actively express understanding. Meanwhile, matching exercises, vocabulary sequencing, and sentence completion serve to train accuracy, vocabulary mastery, and understanding of language structure. Thus, the variety of assessment methods in this textbook can be considered sufficiently relevant to support comprehensive Arabic language learning.

The assessments in this textbook also cover the four language skills: listening (الإستماع), speaking (الكلام), reading (القراءة), and writing (الكتابة). The listening skill is designed to train students to understand sounds, speech, or information conveyed orally. The speaking skill aims to develop students' ability to express ideas and information orally using Arabic. The reading skill serves to improve the ability to understand written texts or information, while the writing skill is designed to train students to construct sentences and communicate ideas in writing.

These four skills are crucial aspects of language learning because they are interrelated in shaping students' language proficiency. An assessment that covers all these skills demonstrates that this textbook does not focus solely on receptive skills, such as listening and reading, but also addresses productive skills, such as speaking and writing. Thus, the evaluation system in this textbook can be said to support the more comprehensive development of students' Arabic language skills.

In terms of placement, the assessments in this textbook are integrated into each learning activity. Overall, this textbook contains learning activities spread across three chapters. These activities include listening, vocabulary, dialogue, composition, as well as reading and writing. The inclusion of assessments in each activity indicates that this textbook not only emphasizes final assessment but also focuses on the ongoing assessment of the learning process.

The assessments provided after each learning activity are broken down into several exercises or training exercises. The number of exercises in each activity is not always the same. Some activities consist of two exercises, three exercises, or even four exercises. This variation in the number of exercises indicates that the book's authors have tailored the assessment format to the characteristics of the material, the level of difficulty, and the learning objectives to be achieved. Consequently, students are not limited to a single type of question but gain a more diverse learning experience.

For example, in Chapter I of Activity 1: Listening, there are three exercises focused on listening skills. One form of assessment in this section is that students are asked to listen to vocabulary words spoken by the teacher. This form of assessment indicates that evaluation is not only directed at written comprehension but also at the ability to perceive, recognize, and understand the sounds of the Arabic language orally. Next, in Activity 2: Vocabulary, there are two exercises: multiple-choice and matching vocabulary. The variety of these assessment methods indicates that the textbook incorporates various evaluation techniques to measure comprehension of meaning, mastery of vocabulary, and accuracy in matching words with their meanings.

Thus, the assessment system in this textbook can be considered quite systematic, as it includes assessments for each learning activity, exercises for each activity, and a final assessment for each chapter. This structure demonstrates that the assessment in the textbook is designed to support continuous, measurable, and comprehensive Arabic language learning.

### **C. Analysis of the Assessment System Based on H. Douglas Brown's Principles**

The analysis of the assessment system in the 7th-grade MTs Arabic textbook for the second semester was conducted using the five principles of assessment proposed by H. Douglas Brown: practicality, reliability, validity, authenticity, and washback. These five principles were used to assess the quality of the evaluation to ensure that the assessment instruments are truly capable of accurately measuring students' abilities and supporting the learning process (Brown, 2004).

#### **1. Practicality**

Based on the principle of practicality, a good assessment should not require significant costs, should be able to be conducted within a reasonable timeframe, be easy to administer, and have clear and efficient scoring procedures (Brown, 2004). Based on the analysis, the assessments in the 7th-grade MTs Arabic textbook for the second semester have met the principle of practicality.

This is evident from the forms of assessment presented in each learning activity and exercise. Assessment types such as multiple-choice, matching, translation, true-false, sentence construction, and fill-in-the-blank are simple and easy to implement in the classroom. These assessments do not require significant costs, can be completed within the standard class time allocation, and are easy for teachers to manage.

Furthermore, the question formats used do not require the use of special media or devices. Teachers can conduct assessments directly in the classroom using the textbook as the primary resource. Thus, from a practical standpoint, the assessments in this textbook are suitable for use in seventh-grade MTs Arabic language instruction.

#### **2. Reliability**

Reliability relates to the consistency of assessment results. An evaluation can be considered reliable if it produces stable and trustworthy data when used under relatively similar conditions (Brown, 2004). In this textbook, the aspect of reliability is essentially evident through the systematic and repeated presentation of exercises in each learning activity. However, the implementation of reliability in the classroom remains highly dependent on the teacher's role.

To achieve good reliability, teachers need to pay attention to several factors, including student readiness, consistency in scoring, uniformity in evaluation procedures, and the clarity of assessment instruments. Teachers must also ensure that the material being assessed has been taught beforehand, that students understand the instructions, and that the classroom environment supports the evaluation process.

Thus, the reliability of the assessments in this textbook is determined not only by the types of questions provided but also by how the teacher conducts the assessments. If the teacher can maintain consistency in implementation, uniformity in procedures, and accuracy in scoring, the assessment results will be more objective and can serve as a basis for evaluating students' abilities.

#### **3. Validity**

Validity indicates the extent to which a test actually measures what it is intended to measure. In the context of learning, a valid assessment must align with the learning objectives, the material being taught, and the competencies to be achieved (Brown, 2004). Based on the analysis results, the assessments in the 7th-grade MTs Arabic textbook for Semester 2 can be said to meet the validity criteria.

This is evident from the alignment between the assessment items and the material covered in each chapter. The assessments are not designed outside the learning context but are directly related to the content and skills being developed. In other words, the evaluation items in this textbook represent the competencies to be measured.

For example, in Chapter II on the topic of *الْبَيْتُ*, the evaluation presented in Activity 2: Dialogue demonstrates alignment between the material and the assessment format. In that section, students are asked to read a dialogue text with a classmate in front of the class, answer questions based on the dialogue, and imitate the provided examples. This form of assessment aligns with the learning objectives because it measures the ability to understand dialogue, speaking skills, and comprehension of the material.

Based on this example, the assessments in this textbook can be considered to have content validity. The assessments presented not only test memorization but also measure skills relevant to the learning material. Therefore, the assessments in this textbook can be used as a reference for evaluating students' learning achievements.

#### 4. Authenticity

Authenticity in assessment refers to the extent to which the tasks or questions provided reflect language use in meaningful contexts that are relevant to students' lives (Brown, 2004). In the Arabic textbook for 7th-grade MTs, Semester 2, the aspect of authenticity is evident in the use of materials, vocabulary, and exercises that are appropriate for the students' developmental level.

The language used in this textbook can be said to be quite appropriate for the average ability of junior high school/MTs students. The material is presented simply, progressively, and in an easy-to-understand manner. One example is the presentation of vocabulary (*المفردات*) accompanied by supporting images. These images help students understand the meaning of the vocabulary, remember the words, and connect the words with the objects in question.

Additionally, several assessment formats encourage students to use the language actively. For instance, students are asked to imitate examples, construct sentences using vocabulary words correctly, translate, and complete sentence gaps. Such assessment formats not only measure recall ability but also train comprehension, precision, and the ability to apply Arabic language elements accurately.

Thus, the assessments in this textbook can be considered authentic, as some of the exercises focus on language use in meaningful contexts. Nevertheless, authenticity could be further enhanced by incorporating more communicative forms of assessment, such as simple conversations, question-and-answer practice, or tasks based on everyday situations.

#### 5. Washback

Washback refers to the impact of assessment on the learning process,

affecting both how teachers teach and how students learn. Effective assessment should have a positive impact on learning, encouraging students to learn more effectively and helping teachers refine their teaching strategies (Brown, 2004; Brown & Abeywickrama, 2010).

Regarding the washback aspect, the Arabic textbook for 7th-grade MTs, Semester 2, cannot yet be empirically proven to have had a positive impact on students, as this requires direct observation of the implementation of assessment in the classroom. However, theoretically, this textbook demonstrates the potential for positive washback.

This potential is evident in the presentation of assessments within each learning activity, which are broken down into various exercises. Such an assessment pattern indicates that learning is not solely focused on measuring final outcomes but also on the gradual process of practice. With a variety of exercises, teachers can provide feedback to students, while students can assess their level of understanding of the material being studied.

In addition, the variety of assessment formats can also help students develop better learning strategies. Students not only practice answering multiple-choice questions but also construct sentences, translate, match vocabulary, and fill in the blanks. Thus, the assessments in this textbook have the potential to encourage students to be more active in understanding and using the Arabic language.

Nevertheless, the positive washback from this textbook still needs to be proven through field research. Direct observation of the learning process is necessary to determine whether the assessments in the textbook truly influence students' learning motivation, learning strategies, and development of Arabic language skills.

#### **D. Strengths and Weaknesses of the Assessment in the Textbook**

Based on the analysis, the assessment in the 7th-grade MTs Arabic textbook for the second semester has several strengths. First, this textbook has been compiled in accordance with the Merdeka Curriculum. This is evident from the correlation between learning outcomes, learning objectives, the sequence of learning objectives, content, and assessments. This alignment indicates that the textbook can assist teachers in conducting instruction in accordance with the direction of the current curriculum policy (Director General of Islamic Education, 2022).

Second, the assessments in this book take various forms. These include multiple-choice, matching, translation, true-false, sentence construction, and fill-in-the-blank exercises. This diversity of question types allows teachers to assess students' abilities across various aspects, including vocabulary comprehension, language structure, and the ability to apply the material in sentence form.

Third, the assessments in this textbook cover the four language skills: listening, speaking, reading, and writing. The inclusion of these four skills indicates that the assessment does not focus solely on a single linguistic aspect but aims to develop students' Arabic language proficiency comprehensively. This aligns with language learning principles that emphasize the integration of receptive and productive skills.

Fourth, the assessments in this textbook are practical. The question formats used are easy to implement, do not require significant costs, and can be applied in classroom instruction. This aligns with the principle of practicality in assessment as

articulated by Brown, who states that a good test must be efficient to administer, easy to implement, and appropriate for the learning context (Brown, 2004).

Although it has several strengths, the assessment in this textbook also has a number of weaknesses. First, the teaching methods used in the book have not been explained in detail. The lack of explanation regarding the approach or teaching methods means that teachers must interpret appropriate strategies on their own to implement the material and assessments in the classroom. In fact, teaching methods are a crucial element that can influence the success of using the textbook.

Second, the reliability of the evaluation process remains highly dependent on the teacher. Although the evaluation formats in the textbook have been systematically designed, assessment results may vary if the teacher lacks clear scoring guidelines or fails to apply evaluation procedures consistently. Therefore, the teacher's role is essential in maintaining objectivity, uniformity, and consistency in assessment.

Third, the washback effect has not yet been empirically proven. The assessments in this textbook do have the potential to positively impact the learning process, but this effect cannot be confirmed without field research. Therefore, direct observation is still needed to determine whether the assessments in this textbook truly impact students' learning motivation, learning methods, and development of Arabic language skills.

Based on the above discussion, it can be concluded that the assessments in the Arabic textbook for 7th-grade MTs, Semester 2, have strengths in terms of alignment with the curriculum, variety of assessment formats, coverage of the four language skills, and practicality of use. However, this book still has weaknesses in terms of explanations of teaching methods, reliance on teacher implementation for reliability, and the lack of empirical evidence regarding washback effects. Thus, this textbook can be considered quite good as an Arabic teaching resource, but it still requires refinement so that its evaluation system is more optimal and can support learning more effectively.

## **5. KESIMPULAN DAN SARAN**

Based on the research findings and discussion, it can be concluded that the Arabic textbook for 7th grade MTs, Semester 2, published by the Ministry of Religious Affairs, is of fairly good quality both as a teaching resource and as a tool for assessing learning. This is evident from the alignment of the book's content with the learning outcomes of the Merdeka Curriculum, the systematic presentation of the material, and the coverage of the four language skills: listening, speaking, reading, and writing. Furthermore, the types of assessments presented in this textbook are quite varied, such as multiple-choice, matching, translation, true-false, sentence construction, and fill-in-the-blank exercises. This variety indicates that the assessments do not focus solely on a single aspect of ability but also aim to measure students' comprehension, vocabulary mastery, language structure, and language skills more comprehensively.

Based on H. Douglas Brown's principles of evaluation, the assessment system in this textbook generally meets several criteria for good evaluation. In terms of practicality, the assessments are easy to administer, do not require significant costs, and can be completed within a reasonable amount of class time. In terms of validity, the assessment formats align with the content and competencies intended to be measured. In terms of authenticity, the language used and the exercise formats are sufficiently appropriate for the proficiency level of seventh-grade MTs students. Meanwhile, regarding reliability, this textbook shows good potential, but its

implementation still depends on the teacher's consistency in providing instructions, managing the evaluation process, and conducting scoring. In terms of washback, the assessments in this textbook have the potential to positively impact Arabic language learning because the exercises are presented progressively in each learning activity. However, this impact cannot yet be empirically confirmed as it still requires validation through field research.

Based on these conclusions, this textbook is suitable for use as a teaching resource and assessment tool for seventh-grade Arabic language instruction at MTs. However, some improvements are still needed to optimize the quality of assessment. The textbook's authors are encouraged to provide clearer scoring guidelines, particularly for subjective tasks such as translation, sentence construction, and speaking exercises. Arabic teachers are also expected to use this textbook creatively by adapting teaching and assessment strategies based on student conditions, as well as maintaining consistency in the implementation of assessments to ensure more objective evaluation results. Additionally, communicative practice should be strengthened through simple dialogues, question-and-answer sessions, speaking practice, or tasks based on everyday situations so that students become more accustomed to using Arabic in real-world contexts. Further research is recommended to conduct field studies to determine the actual impact of the assessments in the textbook on students' learning motivation, learning strategies, and development of Arabic language skills.

#### **DAFTAR REFERENSI**

- Aqil. (2020). Studi Kepustakaan Mengenai Kualitas Pelayanan Terhadap Kepuasan Pasien Rawat Jalan Di Rumah Sakit. *Jurnal Ilmiah Pamenang* 2, no. 2: 1–6. <https://doi.org/10.53599/jip.v2i2.58>.
- Ardiansyah, dkk. (2023). Teknik Pengumpulan Data Dan Instrumen Penelitian Ilmiah Pendidikan Pada Pendekatan Kualitatif dan Kuantitatif. *Jurnal IHSAN: Jurnal Pendidikan Islam* 1, no. 2: 1–9. <https://doi.org/10.61104/ihsan.v1i2.57>.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. New York: Pearson Education.
- Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices*. New York: Pearson Education.
- Deepublish. "Buku Ajar: Manfaat, Contoh, Jenis dan Cara Menulis," Situs Resmi Deepublis, 31 Januari 2022. <https://penerbitdeepublish.com/pengertian-buku-ajar/> 12.47 WIB (Diakses 27 November 2025).
- Direktur Jenderal Pendidikan Islam. (2022). *Keputusan Direktur Jenderal Pendidikan Islam Nomor 3211 Tahun 2022 tentang Capaian Pembelajaran Pendidikan Agama Islam dan Bahasa Arab dalam Kurikulum Merdeka pada Madrasah*. Jakarta: Kementerian Agama Republik Indonesia.
- Faiz, dkk. (2024). Memahami Makna Tes, Pengukuran (Measurement), Penilaian (Assessment), Dan Evaluasi (Evaluation) Dalam Pendidikan. *Jurnal Education and development* 10, no. 3: 492–95. <https://doi.org/10.31004/obsesi.v5i2.972>.

- Haifa, dkk. (2020). Metode Qirā'ah Dalam Pembelajaran Keterampilan Reseptif Berbahasa Arab Untuk Pendidikan Tingkat Menengah. *Lisanuna*, 10(1), 32–44.
- Jailani, & Saksitha. (2024). Tehnik Analisis Data Kuantitatif dan Kualitatif Dalam Penelitian Ilmiah. *Journal Genta Mulia* 15 (t.t.).
- Muslimah, & Widiyanti. (2023). Analisis Daya Beda Tes Hasil Belajar Bahasa Arab Siswa SMA Mamba'ul Hikmah Paron Ngawi. *AL-MU'ARRIB: Journal Of Arabic Education* 3, no. 2: 13–23. <https://doi.org/10.32923/al-muarrib.v3i2.3594>.
- Qamaruddin, & Sa'diyah. (2024). Kajian Teoritis tentang Teknik Analisis Data dalam Penelitian Kualitatif: Perspektif Spradley, Miles dan Huberman. *Journal of Management, Accounting, and Administration* 1, no. 2: 77–84. <https://doi.org/10.52620/jomaa.v1i2.93>.
- Restalia, dkk. (2025). Evaluasi Efektivitas Asesmen Berbasis Karakter Kolaboratif Dalam Pembelajaran Kooperatif Siswa Kelas 2 SD. *TEACHING : Jurnal Inovasi Keguruan dan Ilmu Pendidikan* 5, no. 2.
- Safarudin, & Zulfamanna. (2023). Problematika Pembelajaran Bahasa Arab. *Jurnal Sathar* 1, no. 1: 34–41. <https://doi.org/10.59548/js.v1i1.41>.
- Mohamad Rian Liputo, Endah Fazirah Hakim, Nur Ainiyah, & Juliani A. Ma'atjala. (2023). Respon Peserta Didik Terhadap Penggunaan Aplikasi Wordwall dalam Pembelajaran Bahasa Arab. *Al-Kilmah*, 2(2), 72–83. <https://doi.org/10.58194/alkilmah.v2i2.1852>.
- Sari, Milya., & Asmendri. (2023). Penelitian Kepustakaan (Library Research) dalam Penelitian Pendidikan IPA. *Natural Science* 6, no. 1 (2020): 41–53. <https://doi.org/10.15548/nsc.v6i1.1555>.
- Solihin, & Yusuf. (2021). *Online Language Test And Assessment: Indonesian Elementary School Teachers' Perception*.
- Susanto. (2021). Pengembangan Alat Dan Teknik Evaluasi Tes Dalam Pendidikan." *Jurnal Tarbiyah Jamiat Kheir* 1, no. 1.
- Susanto, dkk. (2022). Content Analysis Method: Analisis Prinsip Isi dan Prinsip Kualitas Laporan Keberlanjutan. *Ekonomis: Journal of Economics and Business* 6, no. 2: 619. <https://doi.org/10.33087/ekonomis.v6i2.650>.
- Syahrizal, & Jaiani. (2023). Jenis-Jenis Penelitian Dalam Penelitian Kuantitatif dan Kualitatif. *Jurnal QOSIM Jurnal Pendidikan Sosial & Humaniora* 1, no. 1: 13–23. <https://doi.org/10.61104/jq.v1i1.49>.
- Tim Penyusun MGMP Bahasa Arab Kementerian Agama Kabupaten Wonosobo. (2023). *Buku ajar Bahasa Arab kelas VII MTs semester 2*. Wonosobo: Kementerian Agama Kabupaten Wonosobo.
- Turohmah, F., Mayori, E., & Sari, R. Y. (2020). Media pembelajaran Word Wall dalam meningkatkan kemampuan mengingat kosa kata bahasa Arab. *Jurnal Pendidikan Luar Sekolah*, 14(2), 013–019. <https://doi.org/10.32832/jpls.v13i2.2634>.
- Wulandari, P., Masruroh, M., Arianti, C. A. L. E., & Nazilah, A. (2024). Pengaruh Media Pembelajaran Wordwall Terhadap Keterampilan Membaca Permulaan Siswa Sekolah Dasar. *Jurnal Multidisiplin Ilmu Akademik*, 1(3), 791–803.
- Artikel Jurnal (satu, dua, atau lebih dari dua penulis)**
- Aqil. (2020). Studi Kepustakaan Mengenai Kualitas Pelayanan Terhadap Kepuasan Pasien Rawat Jalan Di Rumah Sakit. *Jurnal Ilmiah Pamenang* 2, no. 2: 1–6. <https://doi.org/10.53599/jip.v2i2.58>.

- Ardiansyah, dkk. (2023). Teknik Pengumpulan Data Dan Instrumen Penelitian Ilmiah Pendidikan Pada Pendekatan Kualitatif dan Kuantitatif. *Jurnal IHSAN: Jurnal Pendidikan Islam* 1, no. 2: 1–9. <https://doi.org/10.61104/ihsan.v1i2.57>.
- Faiz, dkk. (2024). Memahami Makna Tes, Pengukuran (Measurement), Penilaian (Assessment), Dan Evaluasi (Evaluation) Dalam Pendidikan. *Jurnal Education and development* 10, no. 3: 492–95. <https://doi.org/10.31004/obsesi.v5i2.972>.
- Haifa, dkk. (2020). Metode Qirā'ah Dalam Pembelajaran Keterampilan Reseptif Berbahasa Arab Untuk Pendidikan Tingkat Menengah. *Lisanuna*, 10(1), 32–44.
- Jailani, & Saksitha. (2024). Tehnik Analisis Data Kuantitatif dan Kualitatif Dalam Penelitian Ilmiah. *Journal Genta Mulia* 15 (t.t.).
- Muslimah, & Widiyanti. (2023). Analisis Daya Beda Tes Hasil Belajar Bahasa Arab Siswa SMA Mamba'ul Hikmah Paron Ngawi. *AL-MU'ARRIB: Journal Of Arabic Education* 3, no. 2: 13–23. <https://doi.org/10.32923/al-muarrib.v3i2.3594>.
- Mohamad Rian Liputo, Endah Fazirah Hakim, Nur Ainiyah, & Juliani A. Ma'atjala. (2023). Respon Peserta Didik Terhadap Penggunaan Aplikasi Wordwall dalam Pembelajaran Bahasa Arab. *Al-Kilmah*, 2(2), 72–83. <https://doi.org/10.58194/alkilmah.v2i2.1852>.
- Qamaruddin, & Sa'diyah. (2024). Kajian Teoritis tentang Teknik Analisis Data dalam Penelitian Kualitatif: Perspektif Spradley, Miles dan Huberman. *Journal of Management, Accounting, and Administration* 1, no. 2: 77–84. <https://doi.org/10.52620/jomaa.v1i2.93>.
- Restalia, dkk. (2025). Evaluasi Efektivitas Asesmen Berbasis Karakter Kolaboratif Dalam Pembelajaran Kooperatif Siswa Kelas 2 SD. *TEACHING : Jurnal Inovasi Keguruan dan Ilmu Pendidikan* 5, no. 2.
- Safarudin, & Zulfamanna. (2023). Problematika Pembelajaran Bahasa Arab. *Jurnal Sathar* 1, no. 1: 34–41. <https://doi.org/10.59548/js.v1i1.41>.
- Sari, Milya., & Asmendri. (2023). Penelitian Kepustakaan (Library Research) dalam Penelitian Pendidikan IPA. *Natural Science* 6, no. 1 (2020): 41–53. <https://doi.org/10.15548/nsc.v6i1.1555>.
- Susanto. (2021). Pengembangan Alat Dan Teknik Evaluasi Tes Dalam Pendidikan.” *Jurnal Tarbiyah Jamiat Kheir* 1, no. 1.
- Susanto, dkk. (2022). Content Analysis Method: Analisis Prinsip Isi dan Prinsip Kualitas Laporan Keberlanjutan. *Ekonomis: Journal of Economics and Business* 6, no. 2: 619. <https://doi.org/10.33087/ekonomis.v6i2.650>.
- Syahrizal, & Jaiani. (2023). Jenis-Jenis Penelitian Dalam Penelitian Kuantitatif dan Kualitatif. *Jurnal QOSIM Jurnal Pendidikan Sosial & Humaniora* 1, no. 1: 13–23. <https://doi.org/10.61104/jq.v1i1.49>.
- Turohmah, F., Mayori, E., & Sari, R. Y. (2020). Media pembelajaran Word Wall dalam meningkatkan kemampuan mengingat kosa kata bahasa Arab. *Jurnal Pendidikan Luar Sekolah*, 14(2), 013–019. <https://doi.org/10.32832/jpls.v13i2.2634>.
- Wulandari, P., Masruroh, M., Arianti, C. A. L. E., & Nazilah, A. (2024). Pengaruh Media Pembelajaran Wordwall Terhadap Keterampilan Membaca Permulaan Siswa Sekolah Dasar. *Jurnal Multidisiplin Ilmu Akademik*, 1(3), 791–803.

### Textbook

- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. New York: Pearson Education.

*An Evaluation of the Assessment System in the 7th-Grade  
MTS Arabic Textbook Based on the Principles of Henry Douglas Brown*

- Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices*. New York: Pearson Education.
- Tim Penyusun MGMP Bahasa Arab Kementerian Agama Kabupaten Wonosobo. (2023). *Buku ajar Bahasa Arab kelas VII MTs semester 2*. Wonosobo: Kementerian Agama Kabupaten Wonosobo.

**Internet Sources by Organization Name**

- Deepublish. "Buku Ajar: Manfaat, Contoh, Jenis dan Cara Menulis," Situs Resmi Deepublis, 31 Januari 2022. <https://penerbitdeepublish.com/pengertian-buku-ajar/> 12.47 WIB (Diakses 27 November 2025).

**Working Paper**

- Solihin, & Yusuf. (2021). *Online Language Test And Assessment: Indonesian Elementary School Teachers' Perception*.

**Laporan Instansi/Lembaga/Organisasi/Perusahaan**

- Direktur Jenderal Pendidikan Islam. (2022). *Keputusan Direktur Jenderal Pendidikan Islam Nomor 3211 Tahun 2022 tentang Capaian Pembelajaran Pendidikan Agama Islam dan Bahasa Arab dalam Kurikulum Merdeka pada Madrasah*. Jakarta: Kementerian Agama Republik Indonesia.