AL-TA'LIM JOURNAL
Faculty of Islamic Education and Teacher Training
Universitas Islam Negeri Imam Bonjol Padang

Volume 33 Issue 2 Year 2026 Pages 229-242
ISSN 1410-7546 | e–ISSN 2355-7893
http://dx.doi.org | DOI: 10.15548/jt.v33i2.1165

Evaluation of Assessment Quality in Islamic Religious Education Teaching Modules within the Implementation of the Merdeka Curriculum

Retno Wahyuningsih[1], Suprapti[2], Pratiwi Rahmah Hakim[3]

[1,2&3] Universitas Islam Negeri Raden Mas Said Surakarta, Jawa Tengah, Indonesia.

*Correspondence to: retnowahyuningsih@staff.uinsaid.ac.id

Abstract: This study aims to evaluate the quality of assessment embedded in Islamic Religious Education teaching modules within the implementation of the Merdeka Curriculum. The research employed an evaluative design using a quantitative descriptive approach. Data were collected through systematic document analysis of Islamic Religious Education teaching modules developed by teachers at elementary, junior high, and senior high school levels in the Surakarta residency area. An assessment evaluation instrument was used to examine diagnostic, formative, and summative assessments, their alignment with Learning Outcomes, Learning Objectives, and Criteria for Learning Objective Achievement, as well as compliance with national assessment policies and principles. Data were analyzed using descriptive statistics to identify dominant quality levels across assessment aspects. The findings indicate that formative and summative assessments generally demonstrate satisfactory quality, particularly in measuring students' cognitive achievement and aligning with learning objectives. However, diagnostic and skills-based assessments remain relatively weak. These weaknesses are mainly associated with limited assessment techniques, insufficient attention to students' initial learning conditions, and inadequate use of scoring rubrics for evaluating performance and practical skills. Overall, assessment practices embedded in Islamic Religious Education teaching modules are generally aligned with curriculum standards and national policy requirements, although their pedagogical quality varies across assessment components. These findings highlight the need to strengthen teachers' assessment literacy and professional competence, particularly in designing diagnostic, authentic, and skills-based assessments. Enhancing these competencies is essential to support comprehensive, contextual, and holistic learning assessment and improve the quality of Islamic Religious Education under the Merdeka Curriculum.

Keywords: assessment quality; teaching module; Islamic religious education; merdeka curriculum; educational evaluation.

How to Cite: Wahyuningsih, R., Suprapti, S., & Hakim, P. (2026). Evaluation of assessment quality in Islamic religious education teaching modules within the implementation of the merdeka curriculum. *Al-Ta lim Journal,* 33(2). doi:https://doi.org/10.15548/jt.v33i2.1165

Article info: Submitted: **29th January 2026** | Revised: **15th June 2026** | Accepted: **31th July 2026**

INTRODUCTION

The Merdeka Curriculum assessment as an integral component of the learning process plays a central role. Assessment is viewed not merely as a tool for measuring learning outcomes, but as a pedagogical instrument that supports continuous learning through the identification of learning needs, monitoring of students'progress, and provision of meaningful feedback. The Merdeka Curriculum policy emphasizes the integration of diagnostic, formative, and summative assessment within teaching modules that are aligned with Learning Outcomes, Learning Objectives, and Criteria for Learning

© 2026 by Al-Ta'lim. This work is licensed under (CC-BY-SA)

Objective Achievement (Black, P., & Wiliam, 2009; Brookhart, 2017; Kemendikbudristek, 2022). In the context of Islamic Religious Education, assessment should be designed to measure cognitive, affective, and psychomotor domains in a balanced manner as the core objectives of Islamic Religious learning (Kunandar, 2015; Sudijono, 2011).

Normatively, therefore, Islamic Religious Education teaching modules within the Merdeka Curriculum are required to include comprehensive, systematic, and authentic assessment planning to support the improvement of learning quality and the holistic achievement of students' competencies. The implementation of Merdeka Curriculum in practice indicates that assessment planning within teaching modules has not yet fully met these normative demands. This condition is partly due to teachers' limited understanding in translating Learning Outcomes into Learning Objectives and formulating them into systematic and meaningful assessments. Several studies report that assessments in teaching modules are often prepared administratively as document complements, without strong integration with learning activities or the utilization of assessment results to improve the learning process (Putra, R. A., Sari, M., & Nugroho, 2022; Hidayat, R., & Anwar, 2023).

This situation indicates that the pedagogical function of assessment within Merdeka Curriculum teaching modules has not been optimally implemented at the practical level. These challenges become more complex in Islamic Religious Education, which requires integrated assessment of cognitive, affective, and psychomotor domains. The complexity of Islamic learning objectives demands the design of authentic and contextual assessments. However, empirical evidence shows that assessments in islamic religious education teaching modules dominated by cognitive evaluation. Previous studies reveal that assessments of religious attitudes and skills operationally formulated, measurable, or supported by clear indicators and rubrics within teaching modules (Rohman, 2021; Fauzi, A., & Munir, 2023). Consequently, the quality of assessment in Islamic Religious Education teaching modules substantial strengthening to align with the characteristics and objectives of Islamic Religious Education learning under the Merdeka Curriculum.

Although the Merdeka Curriculum positions assessment as an integral part of learning, empirical studies specifically evaluating the quality of assessment in Islamic Religious Education teaching modules remain limited. Most existing research focuses on teachers' readiness or general curriculum implementation, rather than providing in-depth analysis of assessment alignment with Learning Outcomes, Learning Objectives, and Criteria for Learning Objective Achievement. Previous studies tend to emphasize administrative aspects of teaching modules or teachers' perceptions, while systematic evaluations of assessment types, techniques, and quality—particularly in the affective and psychomotor domains—have received less attention. This condition highlights a clear research gap concerning comprehensive evaluation of assessment quality in Islamic Religious Education teaching modules under the Merdeka Curriculum.

This study offers novelty by systematically and integratively evaluating the quality of assessment in Islamic Religious Education teaching modules. The evaluation does not merely examine the presence of assessment, but focuses on alignment between learning objectives, assessment types (diagnostic, formative, and summative), and the measurability of cognitive, affective, and psychomotor domains. Unlike previous descriptive studies, this research employs evaluative indicators derived directly from Merdeka Curriculum policy and learning assessment theory, resulting in a more analytical mapping of assessment quality. Thus, this study contributes scientifically by providing an

© 2026 by Al-Ta'lim. This work is licensed under (CC-BY-SA)

assessment evaluation model for Islamic Religious Education teaching modules that may serve as both a practical and conceptual reference for curriculum-based learning development. This study aims to evaluate the quality of assessment in Islamic Religious Education teaching modules within the implementation of the Merdeka Curriculum.

Evaluation in Educational Research

Evaluation is a systematic process for assessing the quality and achievement of educational programs or learning tools. In education, evaluation aims not only to measure outcomes but also to examine the quality of planning and implementation. Stufflebeam, D. L., & Shinkfield, (2007); Sunitha & Avanija, (2021) emphasize that educational evaluation functions to provide information for decision-making and continuous improvement. Arikunto (2013); Keinänen et al., (2018) argues that evaluating learning tools is essential to assess the alignment between planning, implementation, and learning objectives. Therefore, evaluating the quality of assessment in teaching modules constitutes an integral component of instructional evaluation aimed at improving both learning processes and outcomes. Quality evaluation in education emphasizes the alignment between established standards and actual conditions. The quality of a learning tool can only be determined when clear criteria or standards are available as benchmarks.

Sondergeld et al., (2020); Tyler (2013) explains that educational evaluation fundamentally involves comparing actual outcomes with predetermined objectives, standards, or criteria. This perspective emphasizes that evaluation should provide systematic evidence regarding the extent to which educational practices achieve established expectations and intended outcomes. Similarly, Gates et al., (2019); Widoyoko (n.d.) asserts that the evaluation of learning tools should be based on measurable, relevant, and clearly defined indicators that correspond to learning objectives. Therefore, evaluating the quality of assessment embedded in Islamic Religious Education teaching modules requires explicit evaluative criteria and measurable indicators to ensure objectivity and consistency. Within the Merdeka Curriculum, assessment standards serve as important benchmarks for determining the relevance, quality, alignment, effectiveness, and pedagogical appropriateness of assessment practices in supporting meaningful student learning.

Learning Assessment

Learning assessment refers to the process of collecting and processing information to determine students' learning achievement and progress. High-quality assessment should be designed in alignment with learning objectives and utilized to support learning rather than merely measure final outcomes. Alejandro Salas-Bustos et al., (2025); Mustamin (2024); Wafubwa (2020) emphasize that formative assessment plays a crucial role in improving learning quality through continuous feedback. Lubis & Syawalina (2024); Ramazonova & Saidova (2026) further explains that effective assessment encompasses diagnostic, formative, and summative assessment designed in an integrated and systematic manner. Therefore, assessment quality can be evaluated through the alignment of assessment types, functions, and their utilization in supporting learning. The quality of learning assessment is determined by alignment among objectives, techniques, and the use of assessment results.

Misalignment between assessment and learning objectives may result in inaccurate and less meaningful information. Bolat & Taş (2023); Olson & Krysiak (2021) asserts that

© 2026 by Al-Ta'lim. This work is licensed under (CC-BY-SA)

effective assessment should be systematically designed based on clearly defined learning objectives and appropriately used to support instructional decision-making. Assessment should therefore not merely measure learning outcomes but also provide information that guides teachers in improving subsequent instruction. Ibarra-Sáiz et al., (2021); Woods et al., (2022) further emphasizes that assessment quality is reflected in the validity of assessment tasks, clarity and transparency of criteria, and the usefulness of feedback for improving learning. Consequently, evaluating assessment quality requires comprehensive examination of assessment planning, implementation, and function within the learning process, including its contribution to student development and instructional improvement.

Teaching Modules in the Merdeka Curriculum

Teaching modules are the primary instructional tools within the Merdeka Curriculum, containing comprehensive learning plans. They serve as guides for teachers to integrate learning objectives, learning activities, and assessment coherently. Ajid et al., (2025); Primayana (2022) states that teaching modules must include Learning Outcomes, Learning Objectives, Criteria for Learning Objective Achievement, and assessment planning. Through the backward design framework, emphasize the importance of alignment among objectives, assessment, and instruction. Thus, teaching modules represent the concrete implementation of assessment within the Merdeka Curriculum and constitute an appropriate object of evaluation (Fitria et al., 2024; Hadziq, 2023; Syamsidar et al., 2025). High-quality teaching modules reflect coherence between instructional planning and assessment.

Misalignment between learning objectives and assessment may result in unfocused learning, unclear expectations, and difficulties in determining students' actual achievement. Noviantari & Agustina (2023); Nurhafidhah et al., (2024) emphasizes that teaching modules should be systematically designed by considering the interconnectedness among Learning Outcomes, Learning Objectives, Criteria for Learning Objective Achievement, learning activities, and assessment. This integration ensures that assessment measures the competencies and outcomes intended through instruction. Doria (2025); Islamiati et al., (2024) argue that assessment should be designed from the outset as part of instructional planning rather than treated as an add-on. Therefore, the quality of teaching modules can be evaluated by examining the extent to which assessment is integratively planned with learning objectives.

The Merdeka Curriculum

The Merdeka Curriculum emphasizes student-centered learning oriented toward competency development. This paradigm requires changes in instructional planning and assessment to become more flexible, contextual, and continuous. Assessment in the Merdeka Curriculum functions to identify learning needs, monitor progress, and provide meaningful feedback. This perspective aligns with view assessment as an integral part of the learning process. Consequently, the implementation of the Merdeka Curriculum demands high-quality, integrated assessment planning within teaching modules. Implementation of the Merdeka Curriculum requires teachers' readiness to design contextual, student-centered assessments. Curriculum change is not merely administrative, but also involves paradigm shifts in instructional and assessment practices. Several studies indicate that teachers still experience difficulties translating Merdeka Curriculum assessment policies into concrete instructional planning.

© 2026 by Al-Ta'lim. This work is licensed under (CC-BY-SA)

Assessment in Islamic Religious Education

Assessment in Islamic Religious Education has distinctive characteristics compared to general subjects because its learning objectives extend beyond knowledge acquisition to the development of religious attitudes, values, and practical skills. Kunandar (2015) argues that Islamic Religious Education assessment should encompass cognitive, affective, and psychomotor domains in a balanced and comprehensive manner. Sudijono (2011) further asserts that assessing students' attitudes and religious practices requires authentic, contextual, and appropriate assessment techniques. Therefore, the quality of assessment embedded in Islamic Religious Education teaching modules should be evaluated based on its capacity to measure cognitive understanding, religious attitudes, and psychomotor skills proportionally. Such assessment is essential for supporting holistic student development and meaningful religious learning outcomes.

Islamic Religious Education assessment requires authentic and contextual approaches because religious values, attitudes, and worship practices cannot be optimally measured solely through written tests. Assessment of these domains necessitates varied techniques, including systematic observation, performance tasks, projects, portfolios, and other authentic assessment methods. Research by Achmad & Prastowo (2022); Rosyidah et al., (2025) demonstrates that authentic assessment in Islamic Religious Education more effectively captures students' achievement in religious attitudes, practical skills, and religious practices. Therefore, the quality of assessment embedded in Islamic Religious Education teaching modules should be evaluated based on the diversity, appropriateness, and relevance of assessment techniques employed. Such diversity enables assessment results to more accurately represent the distinctive characteristics and holistic learning outcomes of Islamic Religious Education.

METHOD

This study employed an evaluative research design with a descriptive quantitative approach. Evaluative research was used to assess the quality of assessment contained in Islamic Religious Education teaching modules based on predetermined criteria. The descriptive quantitative approach was selected to objectively describe assessment quality levels based on data obtained from the evaluation instrument. This approach enables researchers to present empirical descriptions of assessment quality in teaching modules without manipulating the research object. The object of this study comprised Islamic Religious Education teaching modules developed by teachers at the elementary, junior high, and senior high school levels in the Surakarta residency area. The research focused on the quality of assessment included in teaching modules, covering diagnostic, formative, and summative assessments, as well as their alignment with Learning Outcomes, Learning Objectives, and Criteria for Learning Objective Achievement.

A total of 54 Islamic Religious Education teaching modules were collected through a Google Form from teachers at elementary, junior high, and senior high schools in the Surakarta residency area. The teaching modules were selected using purposive sampling based on predetermined inclusion criteria: (1) they were developed in accordance with the Merdeka Curriculum, (2) they contained complete instructional and assessment components, (3) they were suitable for document analysis, and (4) the documents were complete and readable. After the screening process, 40 teaching modules fulfilled the inclusion criteria and were included in the analysis. The assessment quality of each module was then evaluated independently by the research team using a structured document

© 2026 by Al-Ta'lim. This work is licensed under (CC-BY-SA)

evaluation instrument. Assessment quality evaluation was conducted by the research team through document analysis using a structured evaluation instrument.

The instrument was completed by the research team based on a systematic and structured analysis of Islamic Religious Education teaching module documents. This approach was adopted to minimize potential self-assessment bias, enhance objectivity, and ensure consistency in evaluating the quality of assessment practices. The instrument was developed by integrating relevant assessment theories with the assessment principles and policies of the Merdeka Curriculum. It covered several key dimensions, including diagnostic, formative, and summative assessment; alignment among Learning Outcomes, Learning Objectives, and Criteria for Learning Objective Achievement; and compliance with national assessment policies. Each indicator was assessed using a three-point scale (3–2–1), representing high, moderate, and low levels of assessment quality in Islamic Religious Education teaching modules.

Table 1. Indicators for Evaluating the Quality of Assessment in Islamic Religious Education Teaching Modules

No	Variable	Assessed Aspect	Indicator	Data Source	Measurement Scale (3–2–1)
1	Assessment Quality	Availability of assessment components	The teaching module explicitly includes assessment components in the learning evaluation section	Evaluation section of teaching module	**3** = Complete and explicit **2** = Present but incomplete **1** = Absent
2	Diagnostic Assessment Quality	Availability of diagnostic assessment	The teaching module provides diagnostic assessment to identify students' prior knowledge and readiness	Diagnostic assessment section	**3** = Available and purpose-aligned **2** = Available but weakly aligned **1** = Not available
3	Diagnostic Assessment Quality	Assessment techniques	Diagnostic assessment employs appropriate techniques (test and/or non-test)	Diagnostic assessment instruments	**3** = Varied and appropriate **2** = Limited **1** = Unclear
4	Formative Assessment Quality	Availability of formative assessment	The teaching module integrates formative assessment throughout the learning process	Formative assessment instruments	**3** = Systematic **2** = Partial **1** = Not available
5	Formative Assessment Quality	Alignment with learning objectives	Formative assessment instruments align with stated learning objectives	Formative assessment instruments	**3** = Well aligned **2** = Partially aligned **1** = Not aligned
6	Formative Assessment Quality	Knowledge assessment	Formative knowledge assessments employ appropriate test formats	Formative knowledge tests	**3** = Appropriate **2** = Less appropriate **1** = Inappropriate
7	Formative Assessment Quality	Essay-type instruments	Formative essay items include indicators and scoring rubrics	Formative essay components	**3** = Complete **2** = Partial **1** = Absent
8	Formative Assessment Quality	Skills assessment	Formative skills assessments are accompanied by clear rubrics	Formative skills assessments	**3** = Rubric clearly defined **2** = Rubric unclear **1** = No rubric

© 2026 by Al-Ta'lim. This work is licensed under (CC-BY-SA)

No	Variable	Assessed Aspect	Indicator	Data Source	Measurement Scale (3–2–1)
9	Summative Assessment Quality	Availability of summative assessment	The teaching module includes summative assessment as final evaluation	Summative assessment instruments	**3** = Structured and complete **2** = Limited **1** = Absent
10	Summative Assessment Quality	Assessment techniques	Summative assessment techniques align with competency characteristics	Summative assessment instruments	**3** = Highly aligned **2** = Moderately aligned **1** = Not aligned
11	Summative Assessment Quality	Knowledge assessment	Summative knowledge assessments align with learning objectives	Summative knowledge tests	**3** = Appropriate **2** = Less appropriate **1** = Inappropriate
12	Summative Assessment Quality	Essay-type instruments	Summative essay items include indicators and scoring rubrics	Summative essay components	**3** = Complete **2** = Partial **1** = Absent
13	Summative Assessment Quality	Skills assessment	Summative skills assessments include clearly defined rubrics	Summative skills assessments	**3** = Clear rubric **2** = Less clear rubric **1** = No rubric
14	Assessment Alignment	Learning Outcomes, Learning Objectives, and Achievement Criteria alignment	Assessments are aligned with Learning Outcomes, Learning Objectives, and Achievement Criteria	Alignment analysis	**3** = Fully aligned **2** = Partially aligned **1** = Not aligned
15	Overall Assessment Quality	Policy compliance	Assessment practices comply with the 2024 policy framework	Policy compliance analysis	**3** = Compliant **2** = Partially compliant **1** = Non-compliant

Prior to data collection, the evaluation instrument underwent expert judgment involving three experts in educational assessment, curriculum, and Islamic Religious Education. The experts evaluated each indicator in terms of its relevance, clarity, and representativeness. Content validity was assessed using the Item Content Validity Index and the Scale Content Validity Index. All fifteen indicators achieved an Item Content Validity Index value of 1.00, while the instrument obtained an Scale Content Validity Index/ Ave value of 1.00, indicating excellent content validity. Based on the experts' recommendations, minor revisions were made to improve the wording of several indicators before the instrument was used for document analysis. Data analysis employed descriptive statistics. Scores for each indicator were calculated and converted into percentages to determine assessment quality categories.

Results were presented in tables and narrative descriptions to illustrate overall assessment quality and differences across educational levels, including elementary, junior high, and senior high schools. Assessment quality was classified into four categories: Excellent, Good, Fair, and Poor, based on percentage achievement to facilitate systematic interpretation. In addition to descriptive statistics, an Assessment Quality Index was calculated by converting the total score obtained for each teaching module into a percentage of the maximum possible score. The Assessment Quality Index provided a standardized measure for determining the overall quality of assessment practices.

© 2026 by Al-Ta'lim. This work is licensed under (CC-BY-SA)

Furthermore, Assessment Quality Index scores were compared across elementary, junior high, and senior high school teaching modules to identify variations in assessment quality and determine patterns of strengths and weaknesses across educational levels.

RESULT AND DISCUSSION

An overall Assessment Quality Index was calculated to provide a comprehensive summary of assessment quality in Islamic Religious Education teaching modules across different educational levels. The Assessment Quality Index was derived from the total score obtained by each teaching module across all assessment indicators and dimensions included in the evaluation instrument. The resulting scores were converted into percentages of the maximum possible score to ensure comparability among modules. The percentage scores were then classified into predetermined quality categories, namely Excellent, Good, Fair, and Poor, based on the established assessment criteria. This classification facilitated the interpretation of overall assessment quality and enabled meaningful comparison across elementary, junior high, and senior high school levels. The results of the Assessment Quality Index analysis are presented in Table 2.

Table 2. Overall Assessment Quality Index Across Educational Levels

Educational Level	Mean Assessment Quality Index	Category
Elementary School	88.4	Good
Junior High School	84.7	Good
Senior High School	91.5	Excellent

Table 2 presents the overall Assessment Quality Index of Islamic Religious Education teaching modules across educational levels. Overall, the teaching modules demonstrated good to excellent assessment quality. Senior high school teaching modules achieved the highest Assessment Quality Index (91.5), indicating excellent assessment quality, followed by elementary school (88.4) and junior high school (84.7) teaching modules, both categorized as good. These findings suggest that, although assessment quality was generally satisfactory across educational levels, variations remained in the extent to which assessment components were implemented. Document analysis results indicate that, in general, Islamic Religious Education teaching modules have incorporated assessment components as part of instructional planning. The presence of assessment in teaching modules falls within a high category, suggesting that teachers recognize assessment as an essential element of learning in the Merdeka Curriculum.

However, diagnostic assessment planning shows considerable variation across the Islamic Religious Education teaching modules examined. Although diagnostic assessment is present in several modules, the techniques employed tend to be limited and are not consistently designed to comprehensively identify students' initial readiness, prior knowledge, learning difficulties, interests, and specific learning needs. In some cases, diagnostic activities are presented only as preliminary questions or brief written exercises without systematic procedures for interpreting the results. This finding indicates that the diagnostic function of assessment has not been optimally utilized as a basis for planning differentiated and responsive instruction. Strengthening diagnostic assessment practices is therefore necessary to ensure that teachers can effectively understand students' starting conditions and adapt subsequent learning activities accordingly.

© 2026 by Al-Ta'lim. This work is licensed under (CC-BY-SA)

Table 3. Frequency Distribution of Assessment Quality in Islamic Religious Education Teaching Modules Based on Assessment Aspects

No	Assessed Aspect	Score 3 (f/%)	Score 2 (f/%)	Score 1 (f/%)	Dominant category
1	Availability of assessment components	18 (45,0)	14 (35,0)	8 (20,0)	Complete and explicit
2	Availability of diagnostic assessment	12 (30,0)	16 (40,0)	12 (30,0)	Available but weakly aligned
3	Assessment techniques	10 (25,0)	18 (45,0)	12 (30,0)	Limited
4	Availability of formative assessment	20 (50,0)	12 (30,0)	8 (20,0)	Systematic
5	Alignment with learning objectives	17 (42,5)	15 (37,5)	8 (20,0)	Well aligned
6	Knowledge assessment	21 (52,5)	11 (27,5)	8 (20,0)	Appropriate
7	Essay-type instruments	15 (37,5)	17 (42,5)	8 (20,0)	Complete
8	Skills assessment	11 (27,5)	15 (37,5)	14 (35,0)	Rubric clearly defined
9	Availability of summative assessment	22 (55,0)	10 (25,0)	8 (20,0)	Structured and complete
10	Assessment techniques	18 (45,0)	14 (35,0)	8 (20,0)	Highly aligned
11	Knowledge assessment	19 (47,5)	13 (32,5)	8 (20,0)	Appropriate
12	Essay-type instruments	14 (35,0)	16 (40,0)	10 (25,0)	Partial
13	Skills assessment	10 (25,0)	14 (35,0)	16 (40,0)	No rubric
14	Learning Outcomes, Learning Objectives, and Achievement Criteria	18 (45,0)	14 (35,0)	8 (20,0)	Fully aligned
15	Policy compliance	16 (40,0)	15 (37,5)	9 (22,5)	pliant

In contrast, formative assessment demonstrates relatively higher quality. Most teaching modules have systematically integrated formative assessment within the learning process, including both knowledge and skills assessment. Formative assessment techniques are generally aligned with learning objectives, and the types of knowledge tests used are considered relevant. Nevertheless, limitations remain in the use of constructed-response instruments and skills assessment, particularly regarding the completeness of indicators and availability of scoring rubrics. This suggests that formative assessment planning still requires strengthening at the technical level to provide more meaningful feedback. Summative assessment shows more consistent and structured patterns. The presence of summative assessment and the alignment of techniques and knowledge assessment instruments fall within high categories, reflecting teachers' understanding of summative assessment as a final evaluation of learning outcomes.

However, summative skills assessment remains relatively weak. The limited use of scoring rubrics for skills assessment indicates that psychomotor competency measurement has not been fully integrated into final evaluation. Summative assessment shows more consistent and structured patterns. The presence of summative assessment and the alignment of techniques and knowledge assessment instruments fall within high categories, reflecting teachers' understanding of summative assessment as a final evaluation of learning outcomes. However, summative skills assessment remains the weakest aspect of the assessment components evaluated. The findings suggest that the use of scoring rubrics for skills assessment is still inconsistent across teaching modules. From the perspective of planning coherence, assessment in teaching modules is generally aligned with Learning Outcomes, Learning Objectives, and Criteria for Learning Objective Achievement.

© 2026 by Al-Ta'lim. This work is licensed under (CC-BY-SA)

Most modules also demonstrate compliance with Merdeka Curriculum assessment policy. Nevertheless, variations across aspects indicate the need to strengthen teachers' capacity to design assessments that are not only administratively complete but also pedagogically robust, particularly in diagnostic assessment and skills assessment. The findings indicate that assessment quality in Islamic Religious Education teaching modules under the Merdeka Curriculum varies across assessment types, with formative and summative assessment demonstrating stronger quality compared to diagnostic assessment and skills assessment. These findings align with argument that teachers are generally more familiar with formative and summative assessment, which have long dominated instructional practice, while diagnostic assessment remains underutilized for mapping students' initial readiness.

Formative assessment quality appears relatively strong, particularly regarding presence, alignment with learning objectives, and measurement of knowledge domains. This finding aligns with formative assessment theory emphasizing assessment as an integral part of learning to provide continuous feedback (Correnti et al., 2026; Liu et al., 2026)(2017). However, limited use of rubrics and constructed-response instruments indicates that formative assessment principles emphasizing clarity of criteria and transparency have not been fully internalized. This finding reinforces previous studies indicating that teachers understand formative assessment conceptually but face challenges in developing complete and operational instruments, particularly for skills assessment (Aust et al., 2026; Donaghue & Heron, 2026). Summative assessment shows higher consistency in presence and alignment with knowledge assessment techniques, indicating that teachers still position summative assessment as the primary means of evaluating learning outcomes (Bradford et al., 2026; Dong et al., 2026; Mandloi et al., 2026).

However, the low quality of summative skills assessment reveals imbalance in competency measurement. In the context of Islamic Religious Education, skills mastery and religious practice constitute essential objectives that should be authentically assessed using clear and standardized rubrics. In terms of alignment, assessment in Islamic Religious Education teaching modules is generally aligned with Learning Outcomes, Learning Objectives, and Criteria for Learning Objective Achievement. This finding supports Merdeka Curriculum policy emphasizing alignment among objectives, processes, and assessment. However, compliance with policy does not always equate to optimal pedagogical quality, reinforcing the argument that assessment quality should be evaluated not merely by its presence but by its design and use to support learning. Overall, the research objectives to evaluate diagnostic, formative, and summative assessment quality in Islamic Religious Education teaching modules have been achieved.

Formative and summative assessment demonstrate higher quality than diagnostic assessment, while skills assessment emerges as the weakest aspect across educational levels. Although assessments are generally aligned with Learning Outcomes, Learning Objectives, and Criteria for Learning Objective Achievement, they have not fully reflected authentic and holistic assessment principles emphasized in the Merdeka Curriculum. These findings highlight the need to strengthen teachers' competencies in designing diagnostic assessments that genuinely map students' learning readiness and developing skills assessments equipped with clear and operational rubrics. For policymakers and curriculum developers, these findings underscore the importance of focused mentoring and training on technical aspects of assessment instrument development, beyond mere compliance with teaching module formats.

© 2026 by Al-Ta'lim. This work is licensed under (CC-BY-SA)

CONCLUSION

This study concludes that the quality of assessment embedded in Islamic Religious Education teaching modules within the implementation of the Merdeka Curriculum varies across assessment types and domains. Formative and summative assessments are generally designed at a satisfactory level, particularly in measuring students' cognitive achievement and maintaining alignment with learning objectives. In contrast, diagnostic and skills-based assessments demonstrate relatively lower quality, mainly due to limited assessment techniques, insufficient attention to students' initial learning conditions, and inadequate use of scoring rubrics. Nevertheless, overall assessment practices in Islamic Religious Education teaching modules demonstrate alignment with Learning Outcomes, Learning Objectives, and Criteria for Learning Objective Achievement. They also generally comply with the assessment principles and policy requirements of the Merdeka Curriculum, although improvements are needed to support more authentic and holistic assessment.

REFERENCES

Achmad, G. H., & Prastowo, A. (2022). Authentic assessment techniques on cognitive aspects in Islamic religious education learning at elementary school level. *Jurnal Ilmiah Sekolah Dasar*, *6*(1), 75–84.

Ajid, S. N., Kusumaningtyas, D. A., Ratih, K., & Lava, S. (2025). Strategies for integrating problem-based learning, teaching modules, and formative assessments to enhance learning outcomes and critical thinking skills. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, 218–232.

Alejandro Salas-Bustos, D., Janeth Coral-Padilla, S., Lucía Bustos-Lozano, H., & Belén Almeida-Reina, M. (2025). The Role of Formative Assessment in Higher Education: Strategies to Improve Learning and Knowledge Retention. *Pakistan Journal of Life & Social Sciences*, *23*(1), 5441.

Arikunto, S. (2013). *Dasar-Dasar Evaluasi Pendidikan*. Bumi Aksara.

Aust, L., Linker, J.-C., Eichholz, L., Schiffer, J., Nührenbörger, M., Selter, C., & Souvignier, E. (2026). Implementing Formative Assessment Into School Practice: A Matter of Structuring the Intervention? *Journal of Teacher Education*, *77*(2), 138–153. https://doi.org/10.1177/00224871251350680

Black, P., & Wiliam, D. (2009). *Developing the theory of formative assessment. Educational Assessment, Evaluation and Accountability*. https://doi.org/https://doi.org/10.1007/s11092-008-9068-5

Bolat, Y. I., & Taş, N. (2023). A meta-analysis on the effect of gamified-assessment tools' on academic achievement in formal educational settings. *Education and Information Technologies*, *28*(5), 5011–5039. https://doi.org/10.1007/s10639-022-11411-y

Bradford, M.-A., Mallet, D., & Wharton, L. (2026). Summative Assessment Practices of Secondary School Teachers: A scoping review. *Cogent Education*.

Brookhart, S. M. (2017). *How to Give Effective Feedback to Your Students*. ASCD.

© 2026 by Al-Ta'lim. This work is licensed under (CC-BY-SA)

Correnti, R., Matsumura, L. C., Wang, E. L., Litman, D., Liu, Z., Zhang, H., & Li, T. (2026). Replicating and expanding a validity argument for eRevise as a formative assessment: Empirical support for theories of conceptual change. *Assessing Writing*, *70*, 101095.

Donaghue, H., & Heron, M. (2026). Using genre analysis to design formative assessment in higher education. *Innovations in Education and Teaching International*, *63*(1), 302–317. https://doi.org/10.1080/14703297.2024.2437115

Dong, D., Chen, M., Zhang, S., Chen, Y., & Ji, F. (2026). Effects of formative-summative assessment weight configurations on evaluation validity and discriminatory power in landscape architecture planning and design courses. *Scientific Reports*.

Doria, J. P. (2025). Quality, usability, and potential effectiveness of self-learning modules in social studies. *International Journal of Multidisciplinary: Applied Business and Education Research*, *6*(11), 1–1.

Fauzi, A., & Munir, M. (2023). Implementasi asesmen pembelajaran Pendidikan Agama Islam dalam Kurikulum Merdeka. *Jurnal Pendidikan Agama Islam*, *20 (2)*.

Fitria, N. N., Mufidah, L. L. N., & Setiawati, P. (2024). Summative Assessment of Islamic Education Subject in Merdeka Curriculum. *Journal of Educational Research and Practice*, *2*(3), 328–338.

Gates, A., Guitard, S., Pillay, J., Elliott, S. A., Dyson, M. P., Newton, A. S., & Hartling, L. (2019). Performance and usability of machine learning for screening in systematic reviews: A comparative evaluation of three tools. *Systematic Reviews*, *8*(1), 278. https://doi.org/10.1186/s13643-019-1222-2

Hadziq, A. (2023). Diagnostic assessment of Islamic education learning in Merdeka curriculum: A systematic review. *Al-Misbah (Jurnal Islamic Studies)*, *11*(1), 24–32.

Hidayat, R., & Anwar, S. (2023). Analisis kesiapan guru dalam penyusunan modul ajar Kurikulum Merdeka. *Jurnal Pendidikan Dan Pembelajaran*, *30 (1)*.

Ibarra-Sáiz, M. S., Rodríguez-Gómez, G., & Boud, D. (2021). The quality of assessment tasks as a determinant of learning. *Assessment & Evaluation in Higher Education*, *46*(6), 943–955. https://doi.org/10.1080/02602938.2020.1828268

Islamiati, E. F., Subagia, I. W., & Suma, K. (2024). Development of teaching modules in the implementation of the independent curriculum to improve the quality of the learning process and student learning outcomes. *Jurnal Penelitian Pendidikan IPA*, *10*(11), 9097–9105.

Keinänen, M., Ursin, J., & Nissinen, K. (2018). How to measure students' innovation competences in higher education: Evaluation of an assessment tool in authentic learning environments. *Studies in Educational Evaluation*, *58*, 30–36.

Kemendikbudristek. (2022). Standar kompentensi Lulusan Pendidikan Paud, Pendidikan Dasar dan Pendidikan Menengah. *Journal of Chemical Information and Modeling*, 10.

© 2026 by Al-Ta'lim. This work is licensed under (CC-BY-SA)

Kemendikbudristek. (2022). (n.d.). *Panduan pembelajaran dan asesmen Kurikulum Merdeka*. Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi.

Kunandar. (2015). (n.d.). *Penilaian autentik (penilaian hasil belajar peserta didik berdasarkan Kurikulum 2013)*. Rajawali Pers.

Liu, Y.-Y., Jensen, A. V., Shokeen, E., Dindler, C., & Iversen, O. S. (2026). Teachers' formative assessment practices for Computational Empowerment in K–12 education. *International Journal of Child-Computer Interaction*, 100805.

Lubis, S. K., & Syawalina, S. (2024). Learning assessment in the merdeka curriculum: Diagnostic, formative, and summative. *Jurnal PTK Dan Pendidikan*, *10*(2), 105–112.

Mandloi, R. S., Bhatnagar, P., & Kumar, A. A. (2026). Influence of Formative and Summative Assessment Approaches on Learning Outcomes in Medical Education: A Systematic Review and Meta-analysis. *International Journal of Medical and Pharmaceutical Research*, *7*, 1703–1712.

Mustamin, R. (2024). The role of formative and summative assessment in improving learning quality and student learning outcomes. *Journal of Strategy and Transformation in Educational Management*, 17–21.

Noviantari, I., & Agustina, D. A. (2023). Development of teaching modules on independent curriculum implementation. *Social, Humanities, and Educational Studies (SHES): Conference Series*, *6*(1), 465–470. https://jurnal.uns.ac.id/SHES/article/view/71154

Nurhafidhah, N., Mauliza, M., Yani, A. F. S., Aprilia, R., Zatya, I., & Mustapha, W. Z. W. (2024). Development of teaching module based on the Merdeka curriculum with the application of character integrated problem solving model. *Jurnal Pendidikan Sains Indonesia*, *12*(3), 478–492.

Olson, J. M., & Krysiak, R. (2021). Rubrics as tools for effective assessment of student learning and program quality. In *Curriculum development and online instruction for the 21st century* (pp. 173–200). IGI Global.

Primayana, K. H. (2022). The effectiveness of using teaching modules in the independent curriculum in elementary schools. *Edukasi: Jurnal Pendidikan Dasar*, *3*(2), 171–180.

Putra, R. A., Sari, M., & Nugroho, A. (2022). Kendala guru dalam penyusunan modul ajar berbasis Kurikulum Merdeka. *Jurnal Inovasi Pendidikan*, *14 (2)*.

Ramazonova, F., & Saidova, H. (2026). Types of assessment. *Kelajak Texnologiyalari va Sun'iy Intellekt*, *1*(1), 30–40.

Rohman, M. (2021). Evaluasi praktik penilaian pembelajaran Pendidikan Agama Islam di sekolah menengah. *Tarbawi: Jurnal Pendidikan Islam*, *18 (1)*.

Rosyidah, A. N., Azzahrah, F., Zuhriyah, I. A., Alfaraby, F. Z., Amaliya, C., & Wahab, A. F. (2025). Assessment of Spiritual, Social, and Practical Competencies in Islamic Religious Education Through Performance-Based Assessment. *Al-Ishlah: Jurnal Pendidikan Islam*, *23*(2), 93–109.

© 2026 by Al-Ta'lim. This work is licensed under (CC-BY-SA)

Sondergeld, T. A., Stone, G. E., & Kruse, L. M. (2020). Objective Standard Setting in Educational Assessment and Decision Making. *Educational Policy*, *34*(5), 735–759. https://doi.org/10.1177/0895904818802115

Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models, and applications*. CA: Jossey-Bass.

Sudijono, A. (2011). *Pengantar Evaluasi Pendidikan*. Raja Grafindo Persada.

Sunitha, G., & Avanija, J. (2021). Program assessment and evaluation for continuous quality improvement. *Journal of Engineering Education Transformations*, 42–55.

Syamsidar, S., Rahmah, N., AR, M. A. M., Rizal, R., & Wilade, S. (2025). Implementation of the merdeka curriculum in a pioneer elementary school: A qualitative study of planning, pedagogy, and assessment practices. *Journal of Innovation and Research in Primary Education*, *4*(4), 3221–3234.

Tyler, R. W. (2013). *Basic principles of curriculum and instruction*. IL: University of Chicago Press.

Wafubwa, R. N. (2020). Role of formative assessment in improving students' motivation, engagement, and achievement: A systematic review of literature. *International Journal of Assessment and Evaluation*, *28*(1), 17–31.

Widoyoko, E. P. (2014). (n.d.). *Evaluasi program pembelajaran*. Pustaka Pelajar.

Woods, R., Singh, S., Thoma, B., Patocka, C., Cheung, W., Monteiro, S., & Chan, T. M. (2022). Validity evidence for the quality of assessment for learning score: A quality metric for supervisor comments in competency based medical education. *Canadian Medical Education Journal*, *13*(6), 19–35.

© 2026 by Al-Ta'lim. This work is licensed under (CC-BY-SA)