

DOI: <https://doi.org/10.37850/cendekia.v18i02.1410>
<https://journal.faibillfath.ac.id/index.php/cendekia>.

BEYOND COGNITIVE ACHIEVEMENT: ASSESSMENT INFRASTRUCTURE AND THE UNEQUAL VISIBILITY OF AFFECTIVE AND PSYCHOMOTOR LEARNING IN ISLAMIC RELIGIOUS EDUCATION

Nur Aini Farida^{1*}, Dea Tara Ningtyas², Nunung Hidayati³, Willa Putri⁴

¹ Universitas Singaperbangsa Karawang, Indonesia

² UIN Jurai Siwo Lampung, Indonesia

³ UIN K.H. Abdurrahman Wahid Pekalongan, Indonesia

⁴ Institut Ummul Quro Al-Islami Bogor, Indonesia

email: nfarida@fai.unsika.ac.id¹; deataraNINGTYAS@metrouniv.ac.id²; nunung.hidayati@uingusdur.ac.id³; willa.putri@iugibogor.ac.id⁴

* Corresponding Author

Received 28 August 2026; Received in revised form 13 September 2026; Accepted 30 October 2026

Abstract

This study examined documented assessment instruments and practices, teachers' reported constraints, and the unequal visibility of learning outcomes across junior-secondary educational settings in Karawang, Indonesia. Using a qualitative multiple-case study design, the study examined eight teacher-centered cases across public and private Islamic secondary schools (MTs) and junior secondary schools (SMP). Data were collected through semi-structured interviews with one Islamic education teacher per school, assessment-document review, and non-participant classroom observation, followed by within- and cross-case analysis. We identified written cognitive assessments in all eight cases but found no structured affective assessment instruments in the reviewed materials. Informal observation of students' attitudes was reported, while two private-school cases identified limited video-based practical assessment. Teachers reported constraints related to class size, assessment time, access to structured rubrics, professional training, and technology. Although cognitive assessment documentation predominated across cases, variations emerged in informal affective observations and isolated attempts at practical assessment. These findings suggest that unequal visibility of learning outcomes may reflect differences in access to usable instruments, documentation routines, and contextual assessment conditions. The study proposes assessment infrastructure as a provisional analytical lens for interpreting these patterns rather than as an empirically validated explanatory model. The findings highlight the need to further investigate sustainable, ethically appropriate approaches to multidimensional IRE assessment.

Keywords: Affective Assessment; Assessment Infrastructure; Assessment Literacy; Islamic Religious Education; Psychomotor Assessment.

INTRODUCTION

Islamic Religious Education (IRE; locally known as Pendidikan Agama Islam) seeks to develop religious knowledge together with values, attitudes, and practical religious competencies. Its assessment system should therefore generate evidence across multiple dimensions of learning rather than treat cognitive attainment as a sufficient proxy for overall achievement. Foundational taxonomies distinguish cognitive and affective outcomes (Bloom, 1956; Krathwohl et al., 1964) while while

subsequent revisions of the cognitive taxonomy strengthened the alignment among learning objectives, instruction, and assessment (Anderson & Krathwohl, 2001). Here, assessment infrastructure is proposed as a sensitizing analytical lens rather than a newly established or empirically validated theory. It denotes the interplay of usable instruments, teachers' assessment literacy, documentary routines, workload conditions, technological practices, and institutional expectations that shapes which student outcomes can become recorded assessment evidence. It is narrower than an assessment system as a whole: the emphasis is on the practical and organizational conditions of evidentiary visibility. This framing is related to assessment literacy and formative assessment but does not claim that either concept is replaced (Schildkamp et al., 2020; Yan & Pastore, 2022).

This multidimensional orientation is particularly important in IRE because knowledge of religious principles does not necessarily demonstrate their internalization or competent enactment. Exclusive reliance on written examinations may therefore provide an incomplete account of student learning. Recent studies similarly advocate combining test and non-test techniques to capture cognitive, affective, and psychomotor evidence through authentic assessment (Aisyah et al., 2024; Ramdhan et al., 2023). For example, previous research documented written and oral tests for cognitive outcomes, observation-based techniques for affective outcomes, and practical demonstrations supported by performance rubrics for psychomotor outcomes (Aisyah et al., 2024). This problem also calls for ethical restraint. Observable participation, respectful conduct, and demonstrated practice can supply limited evidence of learning, but do not establish the sincerity of belief or internal religious commitment. In IRE, assessment of attitudes should avoid intrusive judgments about faith, protect student privacy, make behavioral criteria transparent, and consider the risks of social desirability, stigma, and unequal interpretation. Accordingly, the argument for broader assessment is not an argument for scoring private religiosity. Research on holistic Islamic pedagogy likewise treats knowledge, attitudes, skills, and spiritual development as complementary dimensions, although evidence from higher education must be transferred cautiously to junior-secondary settings (Budyanti et al., 2024).

Despite this conceptual consensus, translating holistic assessment into routine IRE practice remains difficult. Existing studies show that affective and psychomotor assessment is feasible when explicit procedures are available. Previous research has documented observation, self-assessment, journals, peer assessment, simulations, and checklists (Wijaya et al., 2023). Affective assessment becomes more systematic when teachers establish indicators, assessment scales, structured observations, and follow-up procedures (Hazyimara et al., 2024) while worship-practice assessment can employ performance tasks with explicit criteria (Syamsiah et al., 2022).

Evidence from other IRE settings reinforces this point. Previous research reported written tests for cognitive outcomes, observations and journals for



affective outcomes, and practical activities for psychomotor outcomes (Dona et al., 2024). Related research likewise identified combinations of written and oral tests, observation, self- and peer assessment, practical demonstrations, performance rubrics, and feedback (Aisyah et al., 2024). More recent research showed that assessment across the three domains can be adapted to learner characteristics when evidence collection, instructional follow-up, and contextual responsiveness are deliberately connected (Lismaya et al., 2026).

Teacher assessment literacy offers an additional explanatory perspective. Formative assessment literacy encompasses conceptual, practical, and socio-emotional dimensions (Yan et al., 2022), while teacher knowledge, psychological factors, and social conditions have been identified as important prerequisites for effective classroom assessment (Schildkamp et al., 2020). In Indonesia, a correlational study of 267 teachers found a strong positive relationship between formative assessment literacy and formative assessment practice (Hidayati et al., 2024). These findings shift attention beyond instructional time alone toward teachers' assessment competence, access to usable instruments, professional support, and the routines through which assessment evidence is generated, documented, and used. A systematic review of teacher assessment literacy further shows persistent difficulties translating assessment knowledge into classroom practice (Pastore, 2023). This literature suggests examining not only what teachers know but also which assessment procedures they can sustain under their working conditions. Affective assessment also raises ethical concerns because judgments about students' attitudes may be influenced by teacher bias and subjective interpretation. Fair assessment therefore requires transparent criteria, consistent procedures, and sensitivity to students' developmental contexts (Rasooli et al., 2018). In junior-secondary Islamic Religious Education, assessment should focus on observable behavior rather than presumed religious sincerity, while protecting students' dignity and confidentiality (Liu et al., 2022).

Under Indonesia's 2025 curriculum framework, Islamic Religious Education (IRE) and Character Education in general junior secondary schools (SMP) is allocated two 40-minute intracurricular lesson periods per week, equivalent to 72 periods annually, alongside a separate allocation of 36 co-curricular periods per year. In Islamic junior secondary schools (MTs), Islamic education is organized into four subjects, namely Qur'an and Hadith Studies, Islamic Creed and Ethics, Islamic Jurisprudence, and Islamic Cultural History, each generally allocated two weekly intracurricular periods, totaling approximately eight periods across the four subjects (Ministry of Primary and Secondary Education, 2025; Ministry of Religious Affairs, 2025). However, these aggregate allocations do not represent equivalent assessment time for individual teachers because MTs instruction is distributed across distinct subjects. Rather than serving as a direct measure of teacher workload, these curricular differences provide a comparative context for examining

whether different instructional arrangements correspond to distinct assessment practices. Similar patterns across these settings would suggest that instructional time alone may not adequately explain cognitive-centered assessment, highlighting the potential relevance of assessment instruments, teacher competence, workload, documentation routines, and institutional expectations.

Although prior studies offer valuable insights into holistic IRE assessment, an important comparative gap remains. Research has examined authentic assessment in vocational schools (Aisyah et al., 2024), affective assessment in elementary settings (Hazyimara et al., 2024), affective and psychomotor assessment in a junior-secondary setting (Wijaya et al., 2023), psychomotor assessment in elementary IRE (Syamsiah & Prastowo, 2022), and barriers to holistic assessment at the junior-secondary level (Muharrozi et al., 2025). However, limited evidence has examined whether the same imbalance persists across MTs and SMP and across public and private schools within a shared geographical context while explicitly accounting for differences in curriculum organization. Junior-secondary Islamic education research also underscores the importance of connecting school culture, religious habituation, and character formation (Wildani et al., 2026). The present study differs by examining whether such aims generate formal, traceable evidence within assessment routines.

Beyond describing differences among cognitive, affective, and psychomotor assessment, this study develops the concept of assessment infrastructure: the interconnected configuration of usable instruments, teacher assessment literacy, documentation routines, technological practices, workload conditions, and institutional expectations through which learning outcomes become observable, recordable, and reportable as formal assessment evidence.

Accordingly, this study aimed to (1) identify the assessment instruments and practices documented across the cognitive, affective, and psychomotor domains of Islamic Religious Education; (2) explore teachers' reported constraints in implementing and documenting affective and psychomotor assessment; (3) examine how assessment instruments, documentation routines, and contextual conditions may contribute to the unequal visibility of learning outcomes; and (4) describe similarities and differences in assessment practices across teacher-centered cases situated in MTs and SMP, as well as public and private schools in the study context. Drawing on these findings, the study proposes assessment infrastructure as a provisional analytical lens for understanding how classroom assessment conditions shape the documentation of multidimensional learning outcomes.

METHOD

This study employed a qualitative multiple-case study design to examine Islamic Religious Education (IRE) assessment practices across eight schools in



Karawang Regency, West Java, Indonesia (Stake, 2006; Yin, 2018). Each case focused on one IRE teacher and the associated assessment evidence, with the school serving as the institutional context rather than the unit of analysis. This design enabled exploratory comparisons across Islamic and general junior secondary schools and public and private institutions. Given the teacher-centered approach, findings support analytical interpretation rather than school-wide generalization or causal conclusions (Creswell et al., 2018).

Eight schools were purposively selected to represent institutional diversity within a shared geographical context, comprising two public and two private Islamic junior secondary schools (MTs) and two public and two private general junior secondary schools (SMP). One teacher responsible for teaching and assessing IRE or a relevant Islamic education subject participated from each school, for a total of eight participants. Where possible, teacher accounts were corroborated through documentary and observational evidence. Schools and participants were anonymized using coded identifiers such as MTsN-A, MTsS-A, SMPN-A, and SMPs-A.

Data were collected through semi-structured interviews, assessment-document analysis, and non-participant classroom observation to compare reported practices with accessible assessment evidence. Interviews explored cognitive, affective, and psychomotor assessment instruments and practices, implementation challenges, technology use, feedback strategies, and documentation procedures. Document analysis examined question banks, daily assessments, examinations, available attitude records, practical assessment instruments, rubrics, and grading documentation to identify formal, traceable evidence across the cognitive, affective, and psychomotor domains. Non-participant classroom observations focused on assessment activities, teacher-student interactions, and the documentation of learning outcomes during instruction.

Data analysis involved data condensation, data display, and conclusion drawing and verification, following an established qualitative analytical framework (Miles et al., 2020). Interview transcripts, observation notes, and documentary evidence were initially coded deductively according to the three learning domains, followed by inductive coding of emerging themes, including assessment constraints, teacher training, technology use, stakeholder expectations, and documentation routines. Within-case analysis compared teacher accounts with observational and documentary evidence, while cross-case analysis identified recurring patterns and differences across Islamic and general junior secondary schools and public and private institutions. Assessment was classified as systematically documented only when identifiable criteria were accompanied by recorded evidence of their application to student learning, distinguishing formal assessment from informal observations or unused instruments.

Research credibility was strengthened by triangulating interviews, observations, and documents, including comparisons between reported practices

and accessible evidence. Participants and schools were anonymized using coded identifiers. Participation was voluntary, and participants and relevant school authorities granted permission for interviews, document access, and classroom observations.

RESULTS AND DISCUSSION

RESULTS

Unequal Visibility of Assessment Domains

Across the eight teacher-centered cases, cognitive assessment generated more consistently identifiable records than affective and psychomotor assessment, a pattern supported by interview, observation, and document data. Each focal case contained identifiable cognitive instruments, particularly multiple-choice and essay questions for daily, midterm, and final assessments. Teachers described these formats as practical because they were relatively easy to administer, score, document, and report. No structured affective instruments were identified in the accessible case materials. Although teachers reported observing students' attitudes and behavior, these observations were generally informal and were not consistently documented using explicit criteria or structured records. The teachers' responses show a similar pattern in their assessment practices. One teacher explained,

"Nearly 90% of my assessment uses written questions. For attitudes, I briefly notice who is active or respectful, but I do not record it specifically." (Teacher, MTsN-A)

Similarly, another teacher stated,

"I rely mostly on written tests and questions to assess my students. For their attitudes, I simply pay attention to how they behave and participate in class, but I do not make any formal notes about it." (Teacher, SMPS-B)

Both teachers therefore rely mainly on written questions to assess students' knowledge, while they observe attitudes informally during classroom activities rather than documenting them in a structured way. Figure 1 presents examples of written assessment activities documented in the participating Islamic and general junior secondary schools, illustrating the implementation of cognitive assessment across the two institutional contexts.





Figure 1. Written assessment activities in Islamic and general junior secondary schools. (A) Students completing a written Islamic Religious Education assessment in an Islamic junior secondary school (MTs). (B) Students completing a written Islamic education assessment in a general junior secondary school (SMP). Source: Field documentation, 2026.

Formalization Gaps in Affective and Psychomotor Assessment

The reviewed documents revealed different levels of traceable evidence across the three assessment domains.

Table 1. Presence of Assessment Instruments by Domain

School	Cognitive	Affective	Psychomotor
MTsN-A	Multiple-choice & essay	No formal instrument identified in reviewed materials	Not identified in reviewed materials
MTsN-B	Multiple-choice & essay	No formal instrument identified in reviewed materials	Not identified in reviewed materials
MTsS-A	Multiple-choice & essay	Informal observation reported; no structured record identified	One-time video attempt
MTsS-B	Multiple-choice & essay	Informal observation reported; no structured record identified	Not identified in reviewed materials
SMPN-A	Multiple-choice & essay	No formal instrument identified in reviewed materials	Not identified in reviewed materials
SMPN-B	Multiple-choice & essay	No formal instrument identified in reviewed materials	Not identified in reviewed materials
SMPS-A	Multiple-choice & essay	No formal instrument identified in reviewed materials	One-time video attempt
SMPS-B	Multiple-choice & essay	No formal instrument identified in reviewed materials	Not identified in reviewed materials

Source: Interviews, document analysis, and observation, 2026.

Note: Affective categories describe reported or accessible evidence, not verified school-wide absence. The two practical attempts were one-time videos without identified structured rubrics. Counts: written multiple-choice and essay instruments 8/8; structured affective records identified 0/8; video-based practical attempts 2/8.

As Table 1 shows, each sampled school had identifiable cognitive instruments, whereas the reviewed affective evidence was predominantly informal and psychomotor assessment was rare. Only MTsS-A and SMPS-A had attempted video-based practical assessment; in both cases, the practice was used once and was not supported by a structured performance rubric.

Constraints within the Assessment Infrastructure

Interviews identified a recurring set of constraints that they associated with difficulties implementing and documenting affective and psychomotor assessment.

Table 2. Factors Constraining Holistic Assessment Reported by Teachers

No	Barrier	Teachers Reporting the Issue
1	Large class size	8/8
2	Limited time for assessment activities	8/8
3	Absence of standardized rubrics	8/8
4	Limited assessment training	8/8
5	Parental emphasis on academic scores	6/8
6	Technology mainly used for cognitive quizzes	6/8
7	Unequal student access to smartphones	4/8

Source: Teacher interviews, 2026.

All eight teachers reported four major constraints to implementing affective and psychomotor assessment: large class sizes, limited assessment time, a lack of structured rubrics, and insufficient professional training. Classes of approximately 35–40 students complicated systematic observation and performance assessment. As one teacher explained,

"With 38 students in one class, I cannot observe every student in every meeting" (Teacher, MTsS-A; translated from Indonesian).

Teachers also reported difficulty integrating assessment and documentation into routine instruction, and they did not identify standardized affective and psychomotor rubrics in the accessible records. Furthermore, all participants reported limited access to specialized training in designing and implementing assessment across these domains.

Curricular Organization and Instructional Time as Contextual Factors

The institutional contexts examined in this study differed in the curricular organization of Islamic education. Under the national curriculum framework, Islamic Religious Education (IRE) and Character Education in general junior secondary schools (SMP) are allocated approximately two 40-minute intracurricular lesson periods per week, with additional co-curricular time allocated



separately. In Islamic junior secondary schools (MTs), four distinct Islamic education subjects collectively account for approximately eight weekly intracurricular lesson periods. These allocations describe institutional curriculum structures rather than the assessment time available to individual teachers.

Across the MTs and SMP cases, written cognitive assessments were consistently documented, whereas accessible records showed limited evidence of structured affective and psychomotor assessment. Two private-school cases attempted video-based practical assessment, but neither established a sustained, rubric-supported procedure. Thus, the cases shared a predominantly cognitive-centered documentation pattern despite differences in curricular organization, without establishing that instructional time caused the observed practices.

Cross-Context Consistency in Assessment Patterns

Cross-case analysis identified written multiple-choice and essay assessment in all eight teacher-centered cases across public and private Islamic and general junior secondary settings. Some cases reported informal affective observation, while MTsS-A and SMPS-A included one-time video-based practical-assessment attempts. No structured affective instrument was identified in the reviewed materials. Structured affective observation rubrics and standardized psychomotor-performance rubrics were not identified in the accessible records. Teachers nevertheless described observing students' behavior informally during instruction. Teachers reportedly used digital platforms mainly for cognitive quizzes and online forms. The two video-based practical-assessment attempts were not sustained; participating teachers cited unequal student access, workload, and the time needed to review submissions.

DISCUSSION

Theoretical Implications: Assessment Visibility and Infrastructure

The findings reveal a discrepancy between the multidimensional objectives of Islamic Religious Education (IRE) and the assessment evidence documented across the eight cases. While written instruments consistently supported cognitive outcomes, affective assessment relied largely on informal observation, and psychomotor assessment remained limited. This imbalance reflects a misalignment between intended learning outcomes and documented assessment practices, despite national expectations for comprehensive assessment. Foundational learning taxonomies emphasize the distinct nature of cognitive and affective outcomes (Anderson & Krathwohl, 2001). Consequently, written tests alone cannot adequately represent observable attitudes or practical religious competencies, regardless of their cognitive complexity. From an ethical perspective, assessment must distinguish observable conduct from personal religious conviction, ensuring transparent criteria, fairness, and respect for students' dignity (Rasooli et al., 2025).

The unequal visibility of learning outcomes may be understood through the interaction of assessment instruments, documentation routines, and contextual

conditions. Established written-test formats generated traceable cognitive evidence, whereas informal affective observations lacked consistent recording procedures, and practical assessments were infrequent. Teachers also reported constraints involving class size, assessment time, rubric availability, professional training, and technology. These findings align with research identifying multidimensional barriers to holistic IRE assessment (Muharrozi et al., 2025) and associations between assessment literacy and formative practices (Hidayati et al., 2024). Broader scholarship likewise emphasizes that effective assessment depends not only on conceptual knowledge but also on practical competence, usable instruments, and supportive conditions (Schildkamp et al., 2020; Yan & Pastore, 2022).

International evidence further strengthens this interpretation. A systematic review identifies persistent challenges in translating assessment literacy into classroom practice (Pastore, 2023), while research from Tanzania reports constraints involving large classes, examination-oriented routines, and limited assessment competence (Shafii & Berger, 2025). Although these settings differ from the present cases, their findings suggest that assessment limitations cannot be explained by instructional time alone. Instead, the interaction between instruments, documentation practices, and implementation conditions offers a more comprehensive, albeit provisional, interpretation of unequal assessment visibility. This perspective positions *assessment infrastructure* as an analytical lens rather than a causally validated explanatory model.

The present study extends previous research by demonstrating that the imbalance among assessment domains persists across different institutional settings. Although established techniques, including observation, attitude journals, self- and peer assessment, practical demonstrations, and performance rubrics, are available (Aisyah et al., 2024) their availability does not guarantee routine implementation. Previous studies emphasize the importance of explicit indicators, systematic observation, and performance criteria in assessing affective and psychomotor outcomes (Hazyimara et al., 2024; Syamsiah & Prastowo, 2022; Wijaya et al., 2023). In contrast, teachers in the present study frequently observed students' attitudes without producing systematic records. This suggests that the central challenge lies in translating existing assessment approaches into manageable procedures that generate traceable evidence.

This gap highlights the problem of *assessment visibility*, whereby established written tests make cognitive achievement more readily documentable, while less structured affective and psychomotor practices remain underrepresented in accessible assessment records. Such imbalance raises concerns about potential construct underrepresentation because written knowledge tests provide incomplete evidence of multidimensional IRE outcomes. Nevertheless, this interpretation must acknowledge both the ethical limits of assessing religious



dispositions and the possibility that relevant records were unavailable for review. Moreover, an umbrella review of K–12 formative assessment research reports heterogeneous effects and limited certainty across the synthesized evidence, suggesting that introducing additional assessment tasks does not necessarily improve learning (Sortwell et al., 2024). The priority is therefore to establish defensible criteria and meaningful documentation rather than simply increase assessment activities.

Technology likewise did not substantially diversify assessment evidence in the participating cases. Digital tools primarily supported cognitive quizzes, while video-based practical assessment was attempted in only two cases without sustained implementation. This distinction between digitization and assessment innovation is consistent with research emphasizing that effective technology-supported assessment depends on purposeful task design, teacher competence, and contextual support (Vlachopoulos & Makri, 2024). Similarly, digital assessment literacy has been conceptualized through the interaction of professional competencies and implementation conditions (Estaji et al., 2024). Accordingly, technological adoption should be evaluated by the quality and diversity of learning evidence generated rather than the digital format alone.

Collectively, these findings suggest that holistic IRE assessment requires better alignment among learning objectives, classroom activities, assessment criteria, documentation, and feedback rather than the accumulation of additional tasks. The study therefore proposes *assessment infrastructure* as a provisional analytical lens for understanding how instruments, assessment literacy, recording routines, workload, technology, and institutional expectations may shape the visibility of learning outcomes. However, the evidence does not establish causal relationships among these components or validate assessment infrastructure as a general explanatory model. Future research should incorporate additional institutional informants, broader assessment records, and direct examination of assessment decision-making processes to evaluate this interpretation.

Practical Implications for Strengthening Holistic IRE Assessment

Strengthening holistic Islamic Religious Education (IRE) assessment requires practical and sustainable procedures that integrate cognitive, affective, and psychomotor evidence into routine instruction. Affective outcomes can be documented through structured observations of responsibility, discipline, and cooperation, while psychomotor competencies can be assessed through religious-practice demonstrations, simulations, and performance tasks (Aisyah et al., 2024; Syamsiah & Prastowo, 2022; Wijaya et al., 2023). Concise rubrics with observable indicators can improve assessment clarity, consistency, and feedback (English et al., 2022; Hazyimara et al., 2024). However, assessment must prioritize observable learning-related behavior rather than presumed religious sincerity, with transparent criteria, confidentiality, and equitable participation. Recent research on

assessment fairness and inclusion further supports accessible procedures that minimize subjective judgments and accommodate diverse student circumstances (Rasooli et al., 2025; Townend et al., 2025).

Teacher professional development should emphasize practical competencies in designing rubrics, collecting evidence, interpreting results, and providing feedback. Research highlights the importance of assessment literacy, professional collaboration, and contextual support in translating assessment knowledge into classroom practice (Hidayati et al., 2024; Schildkamp et al., 2020; Yan & Pastore, 2022). These findings support targeted professional development focused on assessment criteria, rubric design, and documentation procedures. In large classes, rotational observation and manageable recording procedures may reduce workload, while student reflection and peer assessment can provide complementary evidence without replacing teacher judgment. These approaches require clear purposes and safeguards, especially when evaluating affective learning (Nieminen & Boud, 2025).

Technology may support portfolios, video-based demonstrations, and assessment documentation when aligned with learning objectives and students' access to resources. Studies emphasize the importance of contextual adaptation, effective assessment design, teacher competence, and institutional support in technology-assisted assessment (Børte et al., 2023; Lismaya et al., 2026; Vlachopoulos & Makri, 2024). Non-digital alternatives should remain available to ensure equitable participation.

At the institutional level, quality assurance should recognize documented affective and psychomotor achievement alongside cognitive outcomes. The shared assessment patterns across Islamic and general junior secondary schools suggest that increasing instructional time alone may be insufficient without appropriate instruments, professional development, and manageable documentation routines. Accordingly, schools could pilot low-burden assessment strategies combining concise rubrics, rotational observation, and selective technology use. However, the present study has not tested their feasibility and effectiveness, and further empirical investigation is needed.

CONCLUSION AND SUGGESTION

This study identified unequal documentation of cognitive, affective, and psychomotor assessment across eight teacher-centered Islamic Religious Education (IRE) cases in public and private junior secondary schools in Karawang, Indonesia. Written cognitive assessments were consistently documented, whereas structured affective instruments were not identified in accessible records, and psychomotor assessment was limited to isolated video-based attempts in two private-school cases. Teachers reported constraints involving class size, assessment time, rubric availability, professional training, and technology. Despite variations in informal



and practical assessment, cognitive-centered documentation predominated across institutional settings. These findings suggest that the unequal visibility of learning outcomes may reflect interactions among assessment instruments, documentation routines, and contextual conditions. Accordingly, the study proposes assessment infrastructure as a provisional analytical lens rather than a validated causal model. Given its teacher-centered design and limited access to documentation, the findings cannot be generalized to entire institutions or used to establish causal relationships.

To strengthen holistic IRE assessment, schools should consider practical observation instruments, concise performance rubrics, targeted teacher training, and accessible technology-supported documentation. These strategies should emphasize observable competencies, assessment fairness, and manageable classroom implementation. Future research should involve additional institutional informants, broader assessment records, and empirical evaluation of proposed assessment strategies to examine their feasibility, effectiveness, and implications for multidimensional learning assessment.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the participating schools, IRE teachers, and school authorities in Karawang Regency for their cooperation during interviews, document access, and classroom observations. The authors also thank all parties who supported the completion of this study.

AUTHOR CONTRIBUTIONS

All authors contributed to the conception and design of the study, data interpretation, and manuscript preparation. NAF coordinated the research and prepared the initial draft. DTN, NH, and WP contributed to data analysis, critical revision, and refinement of the manuscript. All authors read and approved the final manuscript.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest related to this study.

DECLARATION OF AI

During the revision of this manuscript, the authors used ChatGPT (OpenAI) to assist with language editing, structural organization, and clarity. The tool was not used to generate or analyze research data, determine research findings, or make scholarly judgments. All AI-assisted text was critically reviewed, verified, and revised by the authors, who take full responsibility for the final content.

REFERENCES

- Aisyah, A., Saona, S., & Sadaat, S. H. (2024). Implementation of authentic assessment of Islamic Religious Education in the Merdeka Curriculum: Qualitative studies in the cognitive, affective, and psychomotor domains in vocational high schools. *Jurnal Pendidikan Indonesia*, 5(4). <https://doi.org/10.59141/japendi.v5i4.9663>
- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. David McKay.
- Børte, K., Lillejord, S., Chan, J., Wasson, B., & Greiff, S. (2023). Prerequisites for teachers' technology use in formative assessment practices: A systematic review. *Educational Research Review*, 41, 100568. <https://doi.org/10.1016/j.edurev.2023.100568>
- Budiyanti, N., & Parhan, M. (2024). Evaluating the Ulul Ilmi model: Enhancing spiritual and moral character through holistic Islamic pedagogy in higher education. *Jurnal Pendidikan Islam*, 13(1), 55–68. <https://doi.org/10.14421/jpi.2024.131.55-68>
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.
- Dona, R., Puspa, P., Rahmayanti, R., & Arifmiboy, A. (2024). Evaluasi pembelajaran PAI: Tinjauan ranah kognitif, afektif dan psikomotorik di SDIT Cahaya Hati Bukittinggi. *Jurnal Pendidikan Tambusai*, 8(3), 43041–43052.
- English, N., Robertson, P., Gillis, S., & Graham, L. (2022). Rubrics and formative assessment in K–12 education: A scoping review of literature. *International Journal of Educational Research*, 113, 101964. <https://doi.org/10.1016/j.ijer.2022.101964>
- Estaji, M., Banitalebi, Z., & Brown, G. T. L. (2024). The key competencies and components of teacher assessment literacy in digital environments: A scoping review. *Teaching and Teacher Education*, 141, 104497. <https://doi.org/10.1016/j.tate.2024.104497>
- Hazymara, K., Nurasia, N., & Itsnainy, S. K. (2024). Implementation of affective domain assessment in Islamic Religious Education learning at SDN Rampal Celaket 2 Malang. *MANAZHIM*, 6(1), 64–77. <https://doi.org/10.36088/manazhim.v6i1.4191>
- Hidayati, Z., Malyuna, S. I., Tamami, B., & Aisyah, K. (2024). Literasi penilaian guru itu penting: Studi korelasional tentang implementasi penilaian formatif dalam Pendidikan Agama Islam. *Jurnal Cendekia Media Komunikasi Penelitian Dan Pengembangan Pendidikan Islam*, 16(02), 521–536. <https://doi.org/10.37850/cendekia.v16i02.744>



- Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). *Taxonomy of educational objectives: The classification of educational goals. Handbook II: Affective domain*. David McKay.
- Lismaya, S. S., Sumarna, E., & Anwar, S. (2026). Adaptive assessment in Islamic Religious Education learning for students with special needs high school. *Jurnal Cendekia Media Komunikasi Penelitian Dan Pengembangan Pendidikan Islam*, 18(01), 73–91. <https://doi.org/10.37850/cendekia.v18i01.1279>
- Liu, H., Zhu, J., Duan, Y., Nie, Y., Deng, Z., Hong, X., Haugen, M., Baker, J. S., & Liang, W. (2022). Development and students' evaluation of a blended online and offline pedagogy for physical education theory curriculum in China during the COVID-19 pandemic. *Educational Technology Research and Development*, 70(6), 2235–2254. <https://doi.org/10.1007/s11423-022-10131-x>
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). *Qualitative data analysis: A methods sourcebook* (4th ed.). SAGE Publications.
- Muharrozi, M., Anwar, S., & Romli, U. (2025). Strengthening holistic assessment in Islamic religious education: Cognitive, psychomotor, and affective domains. *At-Tarbawi: Jurnal Kajian Kependidikan Islam*, 10(2), 174–190. <https://doi.org/10.22515/attarbawi.v10i2.13364>
- Nieminen, J. H., & Boud, D. (2025). Student self-assessment: A meta-review of five decades of research. *Assessment in Education: Principles, Policy & Practice*, 32, 127–151. <https://doi.org/10.1080/0969594X.2025.2510211>
- Pastore, S. (2023). Teacher assessment literacy: A systematic review. *Frontiers in Education*, 8, 1217167. <https://doi.org/10.3389/educ.2023.1217167>
- Ramdhan, T. W., Mufaizin, M., & Putra, M. K. B. (2023). Kurikulum Merdeka's authentic assessment for multicultural Islamic Religious Education. *Jurnal Al-Murabbi*, 8(2), 60–76. <https://doi.org/10.35891/amb.v8i2.4196>
- Rasooli, A., Holden, M., & Sinval, J. (2025). Preservice teachers' assessment decisions: Exploring the role of fairness conceptions. *British Educational Research Journal*, 51, 2450–2473. <https://doi.org/10.1002/berj.4181>
- Rasooli, A., Zandi, H., & DeLuca, C. (2018). Re-conceptualizing classroom assessment fairness: A systematic meta-ethnography of assessment literature and beyond. *Studies in Educational Evaluation*, 56, 164–181. <https://doi.org/10.1016/j.stueduc.2017.12.008>
- Schildkamp, K., van der Kleij, F. M., Heitink, M. C., Kippers, W. B., & Veldkamp, B. P. (2020). Formative assessment: A systematic review of critical teacher prerequisites for classroom practice. *International Journal of Educational Research*, 103, 101602. <https://doi.org/10.1016/j.ijer.2020.101602>
- Shafii, R. A., & Berger, J.-L. (2025). Teacher assessment literacy, formative assessment practices, and their perceived efficacy in Tanzania: A scoping review. *Studies in Educational Evaluation*, 86, 101496. <https://doi.org/10.1016/j.stueduc.2025.101496>

- Sortwell, A., Trimble, K., Ferraz, R., Geelan, D. R., Hine, G., Ramirez-Campillo, R., Carter-Thuiller, B., Gkintoni, E., & Xuan, Q. (2024). A systematic review of meta-analyses on the impact of formative assessment on K-12 students' learning: Toward sustainable quality education. *Sustainability*, 16, 7826. <https://doi.org/10.3390/su16177826>
- Stake, R. E. (2006). *Multiple Case Study Analysis*. Guilford Press.
- Syamsiah, P. N., & Prastowo, A. (2022). Implementasi penilaian autentik pada ranah psikomotorik dalam pembelajaran PAI di sekolah dasar. *Journal of Research and Thought on Islamic Education (JRTIE)*, 5(2), 130–143. <https://doi.org/10.24260/jrtie.v5i2.2204>
- Townend, G., Alonzo, D., Knipe, S., & Baker, S. (2025). What does the international literature say about assessment practice for equitable learning outcomes for educationally disadvantaged high school students? *Frontiers in Education*, 10, 1536485. <https://doi.org/10.3389/educ.2025.1536485>
- Vlachopoulos, D., & Makri, A. (2024). A systematic literature review on authentic assessment in higher education: Best practices for the development of 21st century skills, and policy considerations. *Studies in Educational Evaluation*, 83, 101425. <https://doi.org/10.1016/j.stueduc.2024.101425>
- Wijaya, A., Andriani, R. T., Ismady, M. W., Suwarni, W. S. D., & Ahzim, R. (2023). The implementation of affective and psychomotor assessment in Islamic religious education learning. *TARBAWY: Indonesian Journal of Islamic Education*, 10(2), 195–206. <https://doi.org/10.17509/t.v10i2.65467>
- Wildani, A. S., Parhan, M., Rahmat, M., & Nugraha, R. H. (2026). Integrating Islamic Religious Education to foster students' noble character: A holistic school-based approach in junior high school. *Jurnal Cendekia: Media Komunikasi Penelitian Dan Pengembangan Pendidikan Islam*, 18(1), 152–164. <https://doi.org/10.37850/cendekia.v18i01.1245>
- Yan, Z., & Pastore, S. (2022). Are teachers literate in formative assessment? The development and validation of the Teacher Formative Assessment Literacy Scale. *Studies in Educational Evaluation*, 74, 101183. <https://doi.org/10.1016/j.stueduc.2022.101183>
- Yin, R. K. (2018). *Case Study Research and Applications: Design and Methods* (6th ed.). SAGE Publications.

