

# Test and Non-Test Evaluation Methods in Islamic Religious Education Learning in Senior High School: A Critical Analysis Based on Literature Review to Strengthen Student-Centered Learning

**Syarifah Musanna** ✉, Universitas Islam Negeri Ar-Raniry Banda Aceh, Indonesia

**Jamaluddin**, Universitas Islam Negeri Ar-Raniry Banda Aceh, Indonesia

**M. Duskri**, Universitas Islam Negeri Ar-Raniry Banda Aceh, Indonesia

✉ [syarifahmusanna@gmail.com](mailto:syarifahmusanna@gmail.com)

**Abstract:** This study critically examines test and non-test evaluation methods in learning Islamic Religious Education (PAI) at the high school level. This research uses the literature review method with primary data in the form of 30 national and international scientific articles published between 2013-2023 from Google Scholar, ScienceDirect, DOAJ, and Moraref databases, as well as policy documents such as Merdeka Curriculum and KMA No. 183/2019. Data were collected through systematic selection using inclusion-exclusion criteria based on relevance to the topic of PAI evaluation in Indonesia. Data analysis was conducted using thematic synthesis technique to identify patterns, challenges, and recommendations in the implementation of PAI evaluation. The results showed that PAI evaluation is still dominated by conventional test methods based on memorization, while non-test methods such as observation, portfolio, and self-assessment that are more effective in assessing noble character and psychomotor skills have not been optimally implemented due to limited technical guidance and teacher training. In fact, evaluation in line with the principle of learner-centered learning demands instruments that are able to measure deep understanding, critical thinking skills, and character strengthening. However, only 30% of PAI teachers understand the preparation of Higher Order Thinking Skills (HOTS) questions, and 65% have difficulty in designing valid instruments for affective aspects. Other challenges identified include unpreparedness of school infrastructure and teachers' administrative burden. Therefore, this study recommends a hybrid evaluation model that combines HOTS-based tests with non-test methods, teacher training in instrument design, and collaboration with religious institutions to strengthen contextual assessment and support more meaningful learning for students.

**Keywords:** PAI evaluation, test method, non-test method, HOTS, learner-centered learning.

**Received** February 27, 2025; **Accepted** May 11, 2025; **Published** June 30, 2025

Published by Mandailing Global Edukasia © 2025.



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

## INTRODUCTION

Islamic Religious Education (PAI) in high schools plays a crucial role in shaping students' noble character, personality, and competency. PAI learning evaluation aims not only to measure mastery of cognitive material but also to assess the development of attitudes, skills, and the internalization of Islamic values in daily life (Putri et al., 2024). However,

current evaluation practices are still dominated by conventional testing methods, such as written exams, that focus on memorization and theoretical understanding (Warisno, 2017). This situation has the potential to undermine the holistic goals of PAI, which emphasize character formation based on faith, piety, and noble character (Widiyanto & Inayati, 2023).

The development of the modern educational paradigm demands student-centered assessment, where evaluation methods not only measure learning outcomes but also encourage active participation, reflection, and the development of soft skills. This concept aligns with the Independent Curriculum, which emphasizes contextual and project-based learning. Unfortunately, the application of non-test methods such as observation, portfolios, self-assessment, and peer assessment in Islamic Religious Education (PAI) remains limited (Suryadi, 2021). Yet, these methods are capable of providing a comprehensive picture of student development, particularly in terms of morals and religious behavior.

Previous studies have shown that integrating test and non-test methods can improve the effectiveness of Islamic Religious Education (PAI) evaluation. For example, the use of Higher Order Thinking Skills (HOTS)-based instruments in written tests encourages students to think critically in understanding religious texts (Yuliani, 2024). Meanwhile, observation techniques and daily journals help teachers assess students' consistency in practicing Islamic values (Ilmi & Ashari, 2024). However, the main challenge is teachers' unpreparedness in designing valid and reliable evaluation instruments, as well as the lack of operational guidelines for combining the two methods.

On the other hand, a global literature review confirms that learner-centered evaluation can increase student motivation and engagement through continuous feedback (Pereira et al., 2016). This principle is relevant to the context of Islamic Religious Education (PAI) in high schools, where evaluation must be an integral part of the learning process to ensure the transfer of religious values into practical practice. Therefore, a critical analysis of test and non-test evaluation methods in PAI is necessary to identify best practices that support holistic and sustainable learning.

This study aims to analyze the effectiveness of test and non-test methods in evaluating Islamic Religious Education (PAI) learning in high schools, identify challenges and opportunities for integrating both methods to strengthen learner-centered learning, and formulate practical recommendations for PAI teachers in designing comprehensive and contextual evaluations. Therefore, the results of this study are expected to serve as a reference for developing PAI evaluations that not only measure competency but also shape individuals with noble character and are ready to face the challenges of the times.

## **METHODS**

This research uses a literature review approach to examine studies related to the evaluation of Islamic Religious Education (PAI) learning in high schools, using both test and non-test methods. The selection of this literature review was based on its ability to systematically and measurably collect, analyze, and synthesize literature, thus enabling the identification of trends, challenges, and innovations in PAI evaluation (Snyder, 2019). This approach is frequently used in educational research to explore various evaluation methods that are effective and relevant to learning needs (Boote & Beile, 2005). The primary focus of this research is the integration of test methods, such as Higher Order Thinking Skills (HOTS)-based written exams (Brookhart, 2010), and non-test methods, such as observation, portfolios, and self-assessment (Brown & Abeywickrama, 2019) within the context of learner-centered learning.

Data sources were obtained from national and international journal articles indexed in reputable databases such as Scopus, Sinta, and DOAJ, as well as textbooks and educational policies, including the Merdeka Curriculum and KMA No. 183/2019. A literature search was conducted through databases such as Google Scholar, ScienceDirect,

and Moraref using keywords such as "evaluation of Islamic Religious Education in Senior High Schools," "evaluation of Islamic Religious Education test methods," and "non-test assessment of Islamic Religious Education." Inclusion criteria for this study included publications between 2013 and 2023, a focus on the Indonesian context, and relevance to the integration of test and non-test methods in Islamic Religious Education evaluation (Xiao & Watson, 2019).

Data analysis was conducted using a narrative synthesis approach, which allows for the identification of patterns, themes, and gaps in the literature (Popay et al., 2006). The collected data were grouped into several main categories: (1) test methods (written exams, HOTS), (2) non-test methods (observation, portfolio), (3) implementation challenges, and (4) evaluation innovations. This approach assisted in designing evidence-based recommendations to improve the effectiveness of Islamic Religious Education evaluation in senior high schools.

## **RESULTS**

Based on a comprehensive literature review, this study presents key findings related to the evaluation of Islamic Religious Education (IRE) learning in senior high schools (SMA), using both test and non-test methods. The analysis focuses on four main themes: (1) the dominance of conventional test methods, (2) the effectiveness of non-test methods, (3) the integration of Higher Order Thinking Skills (HOTS) in written assessments, and (4) implementation challenges.

### **Dominance of Conventional Test Methods**

The dominance of conventional test methods in IRE assessment at the high school level remains prevalent in Indonesia. Evaluation practices tend to rely heavily on written examinations such as midterm and final exams, as well as multiple-choice or essay questions targeting lower-order cognitive skills, primarily memorization and basic understanding. A study by Widiyanto & Inayati (2023) found that IRE teachers in Central Java predominantly use written tests as the main assessment tool, with contextual or project-based assessments rarely implemented. Many teachers prioritize rote memorization of Qur'anic verses, religious arguments, or fiqh concepts without linking them to real-life contexts.

Several factors contribute to this trend. First, there is a lack of teacher training in designing more complex evaluation instruments. Rahma et al. (2024) reported that 60% of IRE teachers in Indonesia rarely attend workshops on HOTS-based assessment or affective domain evaluation. Second, high administrative burdens lead teachers to prefer written tests due to their ease of grading, while non-test methods such as observation and portfolios require more intensive documentation. Third, limited school resources, such as insufficient computer labs or internet access, hinder the implementation of technology-based or project-based assessments.

### **Effectiveness of Non-Test Methods**

Non-test methods in IRE evaluation include observation, portfolios, self-assessment, and peer assessment. Each method plays a specific role in assessing students' affective and psychomotor domains. Observations are conducted using monitoring sheets to assess prayer habits, honesty, and participation in social activities—for example, students' consistency in attending congregational prayers or involvement in social programs such as blood donation.

Portfolios include documentation of students' religious activities, such as Ramadan reports, reflective journals, or creative da'wah projects. Self and peer assessments involve

students reflecting on their behavior through questionnaires or group discussions. For instance, students assess their fairness or cooperation during group projects.

### **Integration of HOTS in Written Tests**

Higher Order Thinking Skills (HOTS) refer to students' critical, analytical, and creative thinking abilities involving synthesis, evaluation, and knowledge application to solve complex problems. In the context of IRE, HOTS-based assessments go beyond memorization, encouraging students to analyze contemporary Islamic legal cases or evaluate social issues such as intolerance and corruption from Qur'anic and Hadith perspectives.

Some teachers have begun incorporating HOTS-oriented questions, such as analyzing zakat obligations for digital workers based on *maqashid syariah* or evaluating the impact of social media on youth morality. However, the implementation of such assessments remains inconsistent due to limited teacher training, lack of relevant question banks, and inadequate technical guidance.

### **Challenges in Implementing IRE Evaluation in Senior High Schools**

IRE assessments in high schools face several challenges, particularly in balancing cognitive, affective, and psychomotor domains. Assessments are still dominated by cognitive aspects like memorization and theoretical understanding, while affective and psychomotor aspects are only sporadically assessed. Teachers often struggle to design valid affective instruments due to the subjective nature of traits such as respect or honesty, which are not supported by clearly defined indicators. Limited infrastructure in remote areas, high administrative workloads, and lack of technical guidance further complicate the implementation of comprehensive assessment systems.

## **DISCUSSION**

The dominance of conventional test methods contradicts the expectations of the Merdeka Curriculum, which emphasizes contextual learning and competency-based assessment. This curriculum requires evaluation instruments that measure students' analytical, creative, and applicative abilities. However, current IRE assessment practices still focus on lower-order thinking skills (Bloom's taxonomy levels C1–C2), such as recalling "Mention the pillars of prayer!" without exploring deeper analytical questions like "How can prayer help reduce stress among teenagers?"

IRE evaluations also lack integration of Islamic values with contemporary social issues, such as corruption or pluralism, which should be central to moral and character formation. This aligns with Pereira's (2016) findings that the dominance of traditional testing in various countries, including Indonesia, hinders the development of learner-centred assessment approaches that emphasize active student participation.

The consequences of conventional test dominance are significant. Students tend to focus only on test scores (surface learning) without deeply understanding Islamic values. Hayati (2019) found that high school students perceive IRE as a memorization subject. Additionally, affective aspects such as honesty and tolerance, and psychomotor skills such as worship practices, are difficult to assess objectively—despite the fact that the core objective of IRE is to foster students of strong faith and noble character. This assessment bias also limits the development of 21st-century skills such as critical thinking, collaboration, and problem-solving, which are essential in modern life.

To address the dominance of conventional tests, several strategic measures are required. First, teachers need training in HOTS-based item construction, such as analyzing Qur'anic verses related to environmental issues or contemporary Islamic law cases. Second, schools should adopt clear rubrics for non-test methods like observation and



portfolio assessment and integrate them into lesson plans. Third, collaboration with Islamic institutions or pesantren can support the evaluation of students' worship practices outside of school.

Non-test methods are effective in assessing students' character development in a contextual and holistic manner. These assessments capture students' social interactions and ethical responses—for example, in managing interpersonal conflicts. Such methods also allow teachers to assess practical skills, such as prayer performance and Qur'anic recitation, which written tests cannot measure. Moreover, peer assessment and direct observation provide valuable insights into students' social and collaborative abilities, while self-assessment promotes reflective learning.

Despite their potential, non-test methods still face several implementation challenges. One major barrier is the lack of technical guidelines for rubric development. Teachers often struggle to assess criteria such as "consistency in practicing Islamic values" due to their subjective nature. Moreover, many teachers are not trained in interpreting qualitative data. High administrative workloads also hinder thorough documentation of portfolios and observations.

Although the integration of HOTS into IRE assessments is promising, its implementation remains limited. Difficulties in constructing contextually relevant questions and insufficient teacher training are significant obstacles. To enhance HOTS-based assessments, recommended strategies include organizing teacher training, developing contextual item banks, utilizing platforms like Google Classroom and SIAP-PAI for data management, and collaborating with religious institutions to evaluate worship practices beyond the classroom.

Challenges in affective and psychomotor assessment, infrastructure limitations, and the dominance of academic achievement culture hinder the development of comprehensive evaluation. The lack of awareness regarding the Merdeka Curriculum and learner-centred assessment further exacerbates these issues. Without strategic reform, the gap between Indonesia's assessment practices and global trends such as holistic, competency-based evaluation—as implemented in Europe—will continue to widen.

## CONCLUSION

This study found that Islamic Religious Education (PAI) evaluation in high schools is still dominated by conventional memorization-based testing methods, while non-test methods such as observation and portfolios have not been optimized. Key obstacles include teacher unpreparedness, a lack of technical guidance, and administrative burdens. To improve evaluation quality, integration of HOTS-based tests with non-test methods, intensive teacher training, and a reduction in administrative burdens are needed. The transformation to learner-centered assessment that balances cognitive, affective, and psychomotor aspects is crucial for maintaining PAI's relevance in shaping students' character.

## REFERENCES

- Arfah, M. A. (2021). Evaluasi Hasil Belajar Pendidikan Agama Islam (Pai). *Jurnal Literasiologi*, 7(2), 211–236.
- Boote, D. N., & Beile, P. (2005). Scholars before researchers: On the centrality of the dissertation literature review in research preparation. *Educational Researcher*, 34(6), 3–15.
- Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. Ascd.
- Brown, H. D., & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices*. Pearson.
- Hayati, S. (2019). *Persepsi Siswa Terhadap Mata Pelajaran Pendidikan Agama Islam Di SMP*

- M. Darul Ilmi, & M. Yahya Ashari. (2024). Metode, Model dan Pendekatan Evaluasi dalam Pendidikan Islam. *Jurnal Budi Pekerti Agama Islam*, 2(1), 27–40. <https://doi.org/10.61132/jbpai.v2i1.52>
- Magdalena, I., Fatmawati, F., Luthfiyah, J., & Khoiriyah, S. (2021). Penggunaan Instrumen Evaluasi Non Tes Dalam Mata Pelajaran Aqidah Akhlak Kelas Iv Di Sdi Nur Insan. *EDISI*, 3(2), 197–209.
- Muamanah, H. (2017). *Pelaksanaan Penilaian Portofolio pada Mata Pelajaran Pendidikan Agama Islam Kelas x di Sekolah Menengah Atas Negeri 5 Palembang*. Doctoral Dissertation, Uin Raden Fatah Palembang.
- P. (2016). *Pereira, D., Flores, A., & Niklasson, L. (2016). Assessment revisited: a review of research in Assessment and Evaluation in Higher Education. 41.*
- Peu'uma, N. Y. (2025). Inovasi Metode Pengajaran Pai Di Era Digital. *Ash-Shahabah: Jurnal Pendidikan Dan Studi Islam*, 11(1), 34–40.
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006). Guidance on the conduct of narrative synthesis in systematic reviews. *A Product from the ESRC Methods Programme Version*, 1(1), b92.
- Putri, T. A., Wati, L. R., & Gusmaneli, G. (2024). Memahami Konsep Desain Pembelajaran Serta Evaluasi Penilaian dalam Pembelajaran Pendidikan Agama Islam. *IHSANIKA: Jurnal Pendidikan Agama Islam*, 2(4), 203–218.
- Rahma, S., Erniati, E., & Hamlan, H. (2024). Problematika Guru Pendidikan Agam Islam (Pai) Dalam Menyusun Soal Berbasis Higher Order Thinking Skill (HOTS) Di SMP Negeri 2 Palu. *Journal on Education*, 6(4), 21019–21027. <https://doi.org/10.31004/joe.v6i4.6251>
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339.
- Warisno, A. (2017). *Pengembangan sistem evaluasi pendidikan agama islam. 21200072*, 1–186.
- Widiyarto, A., & Inayati, N. L. (2023). Penerapan Evaluasi Pembelajaran Tes Dan Non-Tes Pendidikan Agama Islam di Sekolah Menengah Kejuruan. *Munaddhomah: Jurnal Manajemen Pendidikan Islam*, 4(2), 307–316. <https://doi.org/10.31538/munaddhomah.v4i2.439>
- Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. *Journal of Planning Education and Research*, 39(1), 93–112.
- Yuliani, Y. (2024). *Instrumen Evaluasi Bentuk Tes Berbasis HOTS Pada Mata Pelajaran Pendidikan Agama Islam. 02(01)*, 47–56.