

EVALUATION OF THE PAI GRADUATION ASSESSMENT POLICY IN SECONDARY SCHOOLS: A PERSPECTIVE FROM THE PAKIS DIVISION YOGYAKARTA REGIONAL OFFICE

Ainun Wafiq¹, Karwadi², Fahrudin³

Universitas Islam Negeri Sunan Kalijaga Yogyakarta, Indonesia¹⁻²

Kantor Wilayah Kementerian Agama Yogyakarta, Indonesia³

Corresponding author: 24204092021@student.uin-suka.ac.id

ABSTRACT

This study analyzes the strategies employed by the Islamic Religious and Religious Education Division (PAKIS) of the Yogyakarta Regional Office of the Ministry of Religious Affairs to assure the quality of regional assessments for Islamic Religious Education (PAI) in secondary schools. It also identifies policy implementation constraints and maps post-assessment professional development solutions. The study employed a descriptive qualitative approach with a policy-study design. Primary data were obtained through written interviews with key informants from the PAKIS Division, while secondary data were drawn from regulatory documents and institutional evaluation reports. Data were analyzed using the interactive model developed by Miles, Huberman, and Saldaña. The findings indicate that macro level quality assurance has been implemented systematically through a three stage control mechanism: determining the proportional distribution of lower, middle, and higher-order thinking skills (LOTS, MOTS, and HOTS) in the test blueprint; reviewing test drafts before printing; and conducting post-examination item analysis to integrate Religious Moderation content. Nevertheless, the effectiveness of this top-down policy is constrained by structural discrepancies, particularly disparities in PAI teachers' pedagogical competence and the persistence of pragmatic teaching to the test practices in peripheral areas. In response, the PAKIS Division has initiated post-assessment interventions through a large scale standardization movement for senior secondary PAI instruction, training in HOTS-oriented item writing, and stronger coordination through district and municipal PAI Subject Teacher Associations (MGMP). These measures are intended to promote the sustainable and equitable quality of Islamic education. **Keywords:** Quality Assurance, Regional Assessment, Islamic Religious Education.

Received: 02-02-2026

Revised: 17-02-2026

Accepted: 25-03-2026

INTRODUCTION

Regional final assessments, such as the Regional Education Standardization Assessment (Asesmen Standardisasi Pendidikan Daerah, ASPD) and the Regional-Based School Examination (Ujian Sekolah Berbasis Daerah, USBD), have become important contemporary policy instruments for mapping educational achievement at the macro level (Sari & Hartono, 2022). Indonesia's national assessment standards similarly require assessment procedures to be objective, fair, and educationally meaningful rather than merely administrative (Ministry of Education, Culture, Research, and Technology, 2022). International frameworks emphasize the alignment of constructs, instruments, interpretation, and policy use when assessment results are employed to evaluate system performance (OECD, 2023). In the Special Region of Yogyakarta (DIY), widely regarded as a national benchmark for education, expectations concerning the accountability of regional evaluation outcomes are particularly high (Nuryanto, 2021). In Islamic Religious Education (PAI) at the secondary-school level, these assessments are no longer intended merely to test cognition or memorization. They are also designed to evaluate the internalization of religious values, noble character, and contemporary priorities such as Religious Moderation, which is actively promoted by the Indonesian Ministry of Religious Affairs (Ministry of Religious Affairs of the Republic of Indonesia, 2019, 2022; Republic of Indonesia, 2023). Evaluation systems in general schools must therefore use instruments capable of measuring substantive religious literacy rather than simply recalling information (Munir, 2023).

Within the bureaucratic structure, the Yogyakarta Regional Office of the Ministry of Religious Affairs, particularly its Islamic Religious and Religious Education Division (PAKIS/PAIS), functions as regulator, facilitator, and supervisor of instructional quality (Yogyakarta Regional Office of the Ministry of Religious Affairs, 2023). The PAKIS Division holds administrative and academic responsibility for ensuring that PAI graduation competency standards in general schools remain aligned with prevailing national curriculum and governance policies (Anwar & Hasan, 2023). Its quality-assurance responsibilities include standardizing test blueprints, aligning examination indicators, and supervising assessment implementation to ensure consistency with the national vision of Islamic education (Asiyah et al., 2024). This responsibility also requires attention to PAI teacher competence because valid assessment practices depend on teachers' professional and personal capacity to translate curricular values into observable evidence (Ayu & Marzuki, 2017).

In practice, however, regional standardized assessments for PAI in secondary schools often encounter instructional dilemmas and marked

discrepancies at the school level. Many secondary schools remain trapped in pragmatic, text-based memorization and teaching-to-the-test practices intended primarily to achieve high regional graduation scores. This tendency obscures the fundamental purpose of Islamic education, which emphasizes character and moral development (Supriadi, 2022). The discrepancy between provincial curriculum policy and pedagogical practice is commonly driven by variations in PAI teachers' ability to develop higher-order thinking skills (HOTS) items and by unequal academic facilities across schools (Sholeh & Rohmah, 2025). Indonesian evidence shows that teachers may recognize the importance of HOTS while still experiencing difficulty translating the construct into instructional strategies and valid assessment tasks (Retnawati et al., 2018; Widana, 2017). PAI teachers frequently work without consistent professional guidance, causing the quality of regional examination items to fluctuate (Zulkifli et al., 2023).

This discrepancy confirms that quality assurance in religious education cannot be separated from macro-level policy intervention by regional authorities (Pratama & Setiawan, 2023). The quality-control strategy of the PAKIS Division is therefore crucial. If supervision and post-examination evaluation are not implemented effectively, regional assessment results risk becoming a collection of scores with no meaningful impact on instructional improvement (Yusuf & Kurniawan, 2024). Assessment becomes educationally productive only when evidence is interpreted and used to adjust instruction, provide feedback, and support subsequent learning (Black & Wiliam, 2009; Panadero, Jönsson, & Botella, 2017). A critical analysis of how the religious-education bureaucracy manages the quality of regional examinations is essential because the success of religious-education policy in general schools depends substantially on effective oversight by its supervising institution (Wibowo, 2022).

To date, the literature on regional-based school examinations has been dominated by quantitative studies that separately examine student achievement or teacher readiness in the classroom (Fauzi, 2022; Ramadhani & Fitri, 2025). Qualitative evaluative research from the perspective of policymakers remains extremely limited. In particular, few studies have investigated how the PAKIS Division of the Yogyakarta Regional Office designs macro-level quality control for regional PAI assessments during the current transition in the national curriculum (Hidayat et al., 2022). This study addresses that gap by focusing directly on the principal actors responsible for assuring the quality of religious-education policy at the provincial level.

Against this background, the study analyzes the strategies employed by the PAKIS Division to assure the quality of regional standardized PAI assessments in secondary schools. It also identifies bureaucratic and human-resource constraints

encountered during policy monitoring and maps appropriate follow-up measures. The study is expected to contribute theoretically to Islamic educational management, particularly the evaluation of religious instructional policy (Mulyadi, 2021). In practical terms, it offers strategic recommendations for regional offices of the Ministry of Religious Affairs throughout Indonesia seeking to improve the accountability, transparency, relevance, and quality of religious-education assessments.

METHOD

This study employed a descriptive qualitative approach with a policy-study design. It focused on evaluating macro-level strategies for assuring the quality of regional PAI examinations. The primary informants were selected purposively and comprised officials and functional staff members of the PAKIS Division responsible for curriculum and assessment. Data were collected through structured in-depth interviews and analysis of relevant regulatory documents. The data were analyzed using the interactive model of Miles, Huberman, and Saldaña (2014), which consists of data condensation, data display, and conclusion drawing and verification.

The data sources were classified into primary and secondary categories. Primary data were obtained through in-depth interviews with key informants selected through purposive sampling. These informants included a functional official or head of the Curriculum and Evaluation Working Team and implementation staff in the PAKIS Division who were directly responsible for standardizing ASPD or USBD assessments for PAI. Secondary data were obtained through document analysis, including regulatory texts, decrees of the Director General of Islamic Education, the 2020-2024 Strategic Plan of the Yogyakarta Regional Office as revised in 2023, regional examination blueprints, and annual reports on the quality of secondary-school PAI examination outcomes in Yogyakarta.

RESULTS AND DISCUSSION

1. PAI Assessment Quality-Standardization Policy

The policy analysis indicates that the quality-assurance strategy for ASPD or USBD assessments in secondary-school PAI is centered on aligning examination instruments with national curriculum regulations. The PAKIS Division actively standardizes test blueprints to ensure that assessed content extends beyond low-level textual cognition and incorporates HOTS indicators and Religious Moderation content (Ministry of Religious Affairs of the Republic of Indonesia, 2022). Quality is controlled through three rigorous technical

stages. Mr. Fahrudin, Head of the PAI Team at the Yogyakarta Regional Office, confirmed this procedure in a written interview on 18 June 2026:

“From the outset, item quality is determined by specifying the standard proportions of LOTS, MOTS, and HOTS in the test blueprint. Once the items have been written, they are reviewed before printing. Finally, we conduct an item analysis after students have completed the examination.” (Fahrudin, personal communication, 18 June 2026)

This account demonstrates that the Regional Office applies a comprehensive quality-control mechanism to maintain the accountability of regional evaluation instruments. Establishing the proportions of lower-order thinking skills (LOTS), middle-order thinking skills (MOTS), and HOTS at the blueprint-development stage is intended to enable a fair and measurable assessment of the depth of students’ religious literacy (Munir, 2023). The distinction among levels of cognitive demand should be grounded in an explicit taxonomy and translated into observable item specifications rather than treated as a numerical quota alone (Anderson & Krathwohl, 2001; Brookhart, 2010). Collective review before printing and post-examination item analysis further demonstrate the regional authority’s dominant and systematic role in safeguarding item quality. Through this end-to-end strategy, the PAKIS Division treats examinations not merely as administrative graduation requirements but as evidence-based instruments for assessing the extent to which secondary-school students in Yogyakarta have internalized religious values, tolerance, and national commitment (Nuryanto, 2021).

2. Identifying Discrepancies and Implementation Constraints

Although the PAKIS Division’s quality-assurance policy is procedurally well designed, the initial implementation of the USBP revealed a substantial discrepancy between macro-level regulations and pedagogical realities. The primary constraint is the unequal pedagogical competence of PAI teachers in translating high-standard, HOTS-oriented blueprints into everyday instructional practices (Sholeh & Rohmah, 2025). This finding is consistent with an evaluation report from the Yogyakarta Regional Office, which emphasizes the need for standardization because the quality of PAI instruction continues to vary among senior secondary schools in the region (Langit Ilmu Edu, 2026). The discrepancy indicates that formal quality standards are frequently constrained by uneven infrastructure and human-resource capacity across districts and municipalities.

Unequal competence and inconsistent standardization have led many schools, particularly those in peripheral areas with limited access to direct professional support, to retain conventional and pragmatic teaching practices. Teachers tend to orient instruction primarily toward achieving regional graduation-score targets, thereby perpetuating teaching-to-the-test practices and obscuring the essential purpose of Islamic education in cultivating character, morality, and noble conduct (Supriadi, 2022). Fluctuating examination quality has therefore become a critical concern for the Yogyakarta Regional Office during the early implementation of the USBD. Institutional evaluation confirms that substantive quality assurance for regional examination outcomes cannot be achieved without first improving and standardizing classroom instruction (Anwar & Hasan, 2023).

The tension between regulatory ideals and grassroots pragmatism, or street-level bureaucracy, further exacerbates this condition. The movement to standardize senior secondary PAI instruction initiated by the Yogyakarta Regional Office (Langit Ilmu Edu, 2026) indirectly acknowledges that classroom-level quality control has remained weak, fragmented, and unsupported by a consistent professional-development model. When regional authorities demand HOTS-based items to demonstrate regional quality while teachers remain bound to a teaching-to-the-test culture, the central purpose of Islamic education developing noble character and internalizing Religious Moderation is reduced to administrative scores (Supriadi, 2022).

The PAKIS Division should therefore avoid limiting its efforts to formal compliance with quality-audit procedures. Post-assessment intervention through the PAI Subject Teacher Association (MGMP) should move beyond technical-cognitive training in item writing toward sustained transformative pedagogical development that addresses the underlying problems of classroom teaching methodology (Pratama & Setiawan, 2023). Evidence on formative-assessment implementation indicates that teachers require not only technical knowledge but also data literacy, collaborative support, professional confidence, and favorable organizational conditions (Schildkamp et al., 2020). Professional development should therefore connect assessment design with deeper-learning pedagogy and opportunities for teachers to examine student work collectively (Darling-Hammond et al., 2019).

The PAKIS Division's strategy reflects a transition from conventional bureaucratic supervision toward continuous, responsive evaluation. Its multilayered quality-control process—from determining the proportions of LOTS, MOTS, and HOTS in the blueprint, to peer review before printing, and post-examination item analysis, as explained by Fahrudin (2026)—suggests an effort to

apply Total Quality Management principles. Nevertheless, this multilayered macro-level system continues to exhibit a critical structural gap. Standardization remains predominantly top-down: uniform evaluation instruments are imposed without sufficiently accounting for differences in human-resource readiness and infrastructure among schools and regions (Sari & Hartono, 2022). Consequently, an assessment standard that appears ideal at the regional level may produce inequitable evaluation outcomes for schools in peripheral areas. This concern is particularly important for Religious Moderation, whose educational success depends on coherent alignment among policy, curriculum, pedagogy, school culture, and evaluation rather than the inclusion of isolated test items (Ma'arif et al., 2025).

3. Post-Assessment Professional Development and Policy Recommendations

To address human-resource constraints and quality disparities, the PAKIS Division converts examination-evaluation data into a large-scale, structured quality-standardization initiative. Rather than treating answer sheets and student scores as administrative records, the Regional Office promotes comprehensive standardization of PAI instruction at the senior secondary level to improve the overall quality of Islamic education (Langit Ilmu Edu, 2026). Post-examination item-analysis data serve as a baseline for identifying school clusters that require urgent institutional intervention (Asiyah et al., 2024).

Based on this mapping, the PAKIS Division implements targeted interventions through intensive workshops, revitalization of the coordinating role of senior secondary PAI MGMPs, specialized training in HOTS-oriented item-writing techniques, and the use of innovative and inclusive teaching methods (Yusuf & Kurniawan, 2024). These measures respond directly to the results of item evaluation and reflect the Regional Office's position that evaluation must not end with test administration but should lead to continuous quality improvement. Through structured professional development, stronger teacher working groups, and standardized post-assessment processes, the PAKIS Division seeks to narrow the gap between government quality standards and actual Islamic Religious Education performance across secondary schools in Yogyakarta (Pratama & Setiawan, 2023).

CONCLUSION

The PAKIS Division's strategy for assuring the quality of regional PAI assessments in secondary schools has been implemented through a systematic three-stage process: determining the proportional distribution of LOTS, MOTS, and HOTS in the test blueprint; reviewing drafts before printing; and conducting post-examination item analysis. This macro-level mechanism aligns regional

examination instruments with national curriculum standards and integrates Religious Moderation content. Its effectiveness, however, remains constrained by structural discrepancies, including differences in PAI teachers' pedagogical competence across schools and the persistence of pragmatic teaching-to-the-test practices in peripheral areas. These practices prioritize administrative graduation scores and ultimately diminish the fundamental role of Islamic education in student character development.

As a follow-up measure, the PAKIS Division has transformed evaluation data into a broad initiative to standardize senior secondary PAI instruction throughout Yogyakarta. Post-assessment interventions include mapping teacher competence, providing intensive HOTS item-writing training, and strengthening coordination through district and municipal senior secondary PAI MGMPs. Policymakers should redirect teacher development from formal cognitive-bureaucratic audit compliance toward transformative pedagogical mentoring that addresses the underlying problems of classroom methodology. Equitable access to assessment and instructional-development workshops is particularly necessary for teachers in peripheral areas if substantive and evenly distributed quality assurance in Islamic education is to be achieved.

REFERENCES

- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Ayu, S. M., & Marzuki, M. (2017). An assessment model of Islamic religion education teacher personality competence. *Research and Evaluation in Education*, 3(1), 77-91. <https://doi.org/10.21831/reid.v3i1.14029>
- Anwar, K., & Hasan, M. (2023). Evaluasi Kebijakan Standardisasi Kompetensi Guru Ngaji oleh Kantor Wilayah Kementerian Agama. *Jurnal Kebijakan dan Pengembangan Pendidikan Islam*, 4(2), 89-104.
- Asiyah, S., Bashori, B., & Mu'ammarr, M. A. (2024). Monitoring dan Evaluasi Program Keagamaan Kementerian Agama dalam Peningkatan Kualitas SDM Keagamaan. *Jurnal Administrasi Pendidikan Islam*, 6(1), 55-70.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31. <https://doi.org/10.1007/s11092-008-9068-5>
- Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. ASCD.
- Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2019). Implications for educational practice of the science of learning and development. *Applied Developmental Science*, 24(2), 97-140. <https://doi.org/10.1080/10888691.2018.1537791>

- Fahrudin. (2026, June 18). Strategi Pengendalian Mutu Soal USBK PAI pada Bidang PAKIS Kanwil Kemenag DIY [Personal communication/written WhatsApp interview].
- Fauzi, A. (2022). Manajemen Pemberantasan Buta Aksara Al-Qur'an Berbasis Komunitas: Studi Peran Kelompok Tadarus Masyarakat. *Jurnal Kemasyarakatan dan Keagamaan*, 14(2), 110-123.
- Hidayat, R., Syafaruddin, S., & Nurmawati, N. (2022). Implementasi Model Evaluasi Kesenjangan (Discrepancy Model) pada Program Pendidikan Keagamaan. *Jurnal Akuntabilitas Manajemen Pendidikan*, 10(1), 45-56.
- Kantor Wilayah Kementerian Agama Daerah Istimewa Yogyakarta. (2023). *Rencana Strategis (Renstra) Peningkatan Mutu Pendidikan Keagamaan dan Pondok Pesantren Provinsi DIY 2020-2024 (Revisi 2023)*. Kanwil Kemenag DIY.
- Kementerian Agama RI. (2022). *Keputusan Direktur Jenderal Pendidikan Islam Nomor 91 Tahun 2022 tentang Petunjuk Teknis Penyelenggaraan Lembaga Pendidikan Al-Qur'an*. Direktorat Jenderal Pendidikan Islam.
- Kementerian Agama RI. (2019). *Moderasi beragama*. Badan Litbang dan Diklat Kementerian Agama RI.
- Langit Ilmu Edu. (2026). Kanwil Kemenag DIY Dorong Standardisasi Pembelajaran PAI SMA untuk Tingkatkan Mutu Pendidikan PAI. <https://www.langitilmuedu.id/cakrawala-utama/3562803555/kanwil-kemenag-diy-dorong-standardisasi-pembelajaran-pai-sma-untuk-tingkatkan-mutu-pendidikan-pai>
- Mulyadi, M. (2021). *Evaluasi Program Pembelajaran dan Pemberdayaan Masyarakat Berbasis Model Prokus*. Rekam Media.
- Ma'arif, M. A., Rokhman, M., Fatikh, M. A., Kartiko, A., & Hasan, M. S. (2025). Kiai's leadership strategies in strengthening religious moderation in Islamic boarding schools. *Jurnal Ilmiah Peuradeun*, 13(1), 23-48. <https://doi.org/10.26811/peuradeun.v13i1.1168>
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE.
- Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia. (2022). *Regulation Number 21 of 2022 concerning educational assessment standards for early childhood, primary, and secondary education*.
- Munir, A. (2023). Eksistensi Metode Iqra' Manual di Era Digital: Studi Kasus Efektivitas Pembelajaran Tajwid dan Makhraj Santri. *Jurnal Studi Al-Qur'an*, 19(1), 75-92.
- Nuryanto, H. (2021). Dinamika Pengajaran Al-Qur'an Tradisional di Yogyakarta: Tantangan Standar Mutu Kemenag. *Jurnal Tamaddun Islam*, 9(1), 12-27.
- OECD. (2023). *PISA 2022 assessment and analytical framework*. OECD Publishing. <https://doi.org/10.1787/dfeobf9c-en>
- Panadero, E., Jönsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, 22, 74-98. <https://doi.org/10.1016/j.edurev.2017.08.004>

- Pratama, A. Y., & Setiawan, B. (2023). Analisis Evaluasi Kesenjangan (Discrepancy Evaluation Model) dalam Penjaminan Mutu Pendidikan Islam. *Jurnal Manajemen Pendidikan Islam (JMPI)*, 8(2), 112-125.
- Ramadhani, S., & Fitri, A. (2025). Kriteria Keberhasilan Program Pendidikan Non-Formal Menggunakan Pendekatan Malcolm Provus. *Jurnal Evaluasi Pendidikan*, 13(1), 18-29.
- Republic of Indonesia. (2023). *Presidential Regulation Number 58 of 2023 concerning the strengthening of religious moderation*.
- Retnawati, H., Djidu, H., Kartianom, Apino, E., & Anazifa, R. D. (2018). Teachers' knowledge about higher-order thinking skills and its learning strategy. *Problems of Education in the 21st Century*, 76(2), 215-230. <https://doi.org/10.33225/pec/18.76.215>
- Sari, D. P., & Hartono, R. (2022). Aplikasi Discrepancy Evaluation Model (DEM) dalam Mengevaluasi Standar Proses Pembelajaran. *Kependidikan: Jurnal Hasil Penelitian dan Kajian Kepustakaan*, 8(3), 510-522.
- 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>
- Sholeh, M., & Rohmah, N. (2025). Implementasi Kebijakan Kurikulum Pendidikan Al-Qur'an di Tingkat Wilayah Provinsi: Antara Idealita dan Realita Lapangan. *Jurnal Pendidikan Islam Al-Ilmi*, 12(1), 60-74.
- Supriadi, E. (2022). *Metodologi Pengajaran Al-Qur'an Kontemporer: Dari Iqra' hingga Tahsin*. Pustaka Edukasi.
- Wibowo, T. (2022). *Sistem Evaluasi Program Keagamaan Non-Formal di Indonesia*. Laksana.
- Widana, I. W. (2017). Higher order thinking skills assessment (HOTS). *Journal of Indonesian Student Assessment and Evaluation*, 3(1), 32-44. <https://doi.org/10.21009/JISAE.031.04>
- Yusuf, M., & Kurniawan, D. (2024). Evaluasi Kinerja Organisasi Sukarelawan Keagamaan dalam Pembelajaran Instrawilayah Kemenag. *Jurnal Tata Kelola Pendidikan Keagamaan*, 5(2), 132-147.
- Zulkifli, Z., Rahman, A., & Fatimah, S. (2023). Manajemen Mutu Instruksional Pengajar Al-Qur'an Sukarelawan pada Lembaga Non-Formal. *Jurnal Manajemen Pendidikan Islam*, 7(2), 143-156.