TOFEDU: The Future of Education Journal
Volume 5 Number 2 (2026) Page: 2513-2522
E-ISSN 2961-7553 P-ISSN 2963-8135
https://journal.tofedu.or.id/index.php/journal/index

The Integration of Test and Non-Test Assessment in Islamic Education toward Holistic Evaluation for Sustainable Learning

[1]Diantoro, [2]Aji Joko Budi Pramono
Universitas Islam Negeri Raden Mas Said Surakarta
dianalbayati505@gmail.com, ajijoko@gmail.com

ABSTRACT
Islamic Religious Education (PAI) carries a sacred mission to develop students' cognitive, affective, and psychomotor aspects in a balanced manner. However, in practice, learning evaluation in PAI remains dominated by test-based assessments that primarily measure the cognitive domain. This study is grounded in the increasing urgency to understand how the imbalance of evaluation instruments hinders the internalization of comprehensive Islamic character values (*kaffah*). This article aims to analyze the theological-pedagogical urgency, implementation mechanisms, and sociological-technical challenges of integrating test and non-test assessments in PAI in order to achieve holistic evaluation and sustainable learning. The study employed a Systematic Literature Review (SLR) method by adapting the PRISMA protocol to literature published between 2023 and 2025. Data were reduced and synthesized from various reputable journals. Of the 58 initial articles identified through scientific databases such as Google Scholar, Garuda, and Scopus, 15 final articles met the strict inclusion criteria for analysis. The findings indicate that the integration of test instruments (written and oral examinations) and non-test instruments (observation, self-assessment, and worship portfolios) simultaneously is capable of capturing both students' theoretical and practical piety. This evaluation arrangement contributes directly to the long-term internalization of Islamic values. Although challenges remain, including teacher subjectivity, the halo effect, and administrative burdens, the digitalization of assessment instruments through application-based systems and the development of measurable and observable behavioral rubrics have emerged as strategic solutions to ensure evaluation objectivity in the modern era.

Keywords: holistic evaluation; assessment integration; Islamic Religious Education (PAI); sustainable learning; Systematic Literature Review (SLR)

PENDAHULUAN
Pendidikan Agama Islam (PAI) pada hakikatnya bukan sekadar proses transfer pengetahuan (knowledge transmission), melainkan sebuah upaya sistematis dan komprehensif dalam mentransformasi nilai, membentuk karakter (character building), serta membimbing perilaku peserta didik berlandaskan Al-Qur'an dan As-Sunnah (Hikmah & Maulida, 2025). Munir (2023) menegaskan bahwa tujuan akhir pendidikan Islam adalah mencetak manusia yang bertakwa secara utuh (kaffah), yang memadukan keunggulan intelektual (fikr) dan kematangan spiritual (zikr). Keberhasilan dalam pembelajaran PAI tidak boleh direduksi hanya menjadi deretan angka mati di atas kertas hasil ujian semesteran (Anwar & Siregar, 2024).

This work is licensed under a Creative Commons Attribution 4.0 International Licens
E-ISSN 2961-7553
P-ISSN 2963-8135

Kendati demikian, Fatmawati dan Rahmawati (2023) menyoroti fenomena ketimpangan yang cukup tajam di lembaga pendidikan saat ini, di mana praktik evaluasi PAI masih didominasi oleh penilaian berbasis tes konvensional. Dominasi ujian tulis ini cenderung hanya menguji hafalan materi, pemahaman tekstual, dan kemampuan konseptual tingkat rendah (Putri et al., 2023). Akibatnya, esensi dari capaian pembelajaran PAI yang sebenarnya menjadi kabur, sebab dimensi afektif seperti keikhlasan, kejujuran, dan kesalehan sosial serta dimensi psikomotorik seperti keterampilan beribadah luput dari pemantauan (Zainuddin, 2024). Evaluasi yang bersifat parsial dan reduksionis ini melahirkan bias dalam menilai kompetensi keagamaan seorang peserta didik (Sholeh, 2025). Kondisi ini dibuktikan oleh temuan Hakim dan Rohman (2025) yang menyatakan bahwa seorang siswa bisa saja mendapatkan nilai sempurna dalam ujian tertulis mengenai hukum shalat, namun di saat yang sama ia melalaikan kewajiban shalat tersebut di kehidupan nyata. Guna mewujudkan sistem evaluasi yang jujur (authentic assessment), diperlukan sebuah rekonstruksi metodologis berupa integrasi yang kokoh antara penilaian tes dengan penilaian non-tes secara seimbang (Anwar & Siregar, 2024). Integrasi ini sejalan dengan spirit Kurikulum Merdeka yang menuntut adanya pembelajaran berdiferensiasi dan evaluasi yang berpusat pada murid (Ramadhan, 2024). Melalui keterpaduan instrumen ini, konsep pembelajaran berkelanjutan (sustainable learning) dapat diwujudkan, di mana orientasi belajar agama tidak mandek ketika ujian sekolah usai, melainkan terus menjadi panduan moral sepanjang hayat (Hakim & Rohman, 2025). Meskipun kajian mengenai evaluasi PAI telah banyak dilakukan oleh para ahli, sebagian besar studi terdahulu cenderung memisahkan antara pengujian instrumen tes dengan instrumen non-tes (Munir, 2023). Kajian komprehensif yang membedah model integrasi sistemik-simultan kedua instrumen ini demi ketercapaian sustainable learning masih sangat terbatas (Hakim & Rohman, 2025). Oleh karena itu, penelitian ini hadir untuk mengisi kekosongan akademis tersebut. Kebaruan dari artikel ini terletak pada perumusan formula integrasi penilaian PAI yang praktis, aplikatif, sekaligus solutif dalam mengatasi bias penilaian perilaku keagamaan siswa melalui pemanfaatan teknologi dan standarisasi rubrik performa.

METODE PENELITIAN

Kajian ini didesain menggunakan metode Systematic Literature Review (SLR) yang bertujuan untuk mengidentifikasi, mengevaluasi, memetakan, dan menyintesis seluruh hasil penelitian primer secara objektif (Sari & Handayani, 2022). SLR dipilih karena kemampuannya menghasilkan sintesis yang komprehensif, berbasis bukti, serta dapat direplikasi. Proses seleksi artikel dalam SLR ini sepenuhnya mengadaptasi protokol PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) untuk menjamin transparansi dan meminimalkan bias seleksi. Langkah-langkahnya dioperasikan melalui empat tahapan sistematis, meliputi identifikasi (identification), penyaringan (screening), kelayakan (eligibility), dan inklusi (included).

2.1 Strategi Pencarian Literatur

Pencarian literatur dilakukan secara elektronik pada pangkalan data ilmiah (database) Google Scholar, Garuda (Garba Rujukan Digital), dan Scopus. Formula pencarian dikembangkan menggunakan operator Boolean dengan kata kunci: ("integrasi penilaian" OR "evaluasi holistik") AND "PAI" AND ("tes" AND "non-tes"). Rentang waktu publikasi artikel dibatasi secara ketat dari tahun 2023 hingga 2025 demi menjaga aktualitas, relevansi, dan kebaruan temuan (Sari & Handayani, 2022). Hasil pencarian awal menghasilkan 58 artikel yang secara umum memiliki kemiripan topik dengan fokus kajian ini. Selanjutnya, proses seleksi dilakukan secara bertahap mengikuti alur PRISMA sebagaimana disajikan pada tabel sebagai berikut:

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

Tabel 1. Alur Seleksi Artikel Menggunakan Pendekatan PRISMA.

Tahap PRISMA	Keterangan	Jumlah Artikel
Identifikasi (*Identification*)	Pencarian awal pada database ilmiah (Google Scholar, Garuda, Scopus) menggunakan kata kunci dan operator Boolean.	58 artikel
Penyaringan (*Screening*)	Eliminasi artikel berdasarkan duplikasi data antar-database serta ketidaksesuaian judul dan isi abstrak dengan fokus penelitian.	32 artikel *(26 artikel dieliminasi)*
Kelayakan (*Eligibility*)	Analisis teks lengkap (*full-text*) dengan mengeksklusi 17 artikel (9 topik umum tanpa kekhasan PAI, 5 kebijakan kurikulum makro, 3 kajian teoretis pra-2021).	15 artikel *(17 artikel dieksklusi)*
Inklusi (*Included*)	Artikel final yang memenuhi seluruh kriteria inklusi ketat dan siap untuk dianalisis lebih lanjut.	**15 artikel**

2.2 Kriteria Inklusi dan Eksklusi

Kriteria inklusi dalam penelitian ini mewajibkan artikel: (1) berasal dari jurnal ilmiah orisinal berstatus peer-reviewed baik nasional terakreditasi maupun internasional, (2) fokus pada mata pelajaran PAI di tingkat sekolah (dasar dan menengah), (3) menyajikan model operasional penilaian tes dan non-tes, dan (4) tersedia dalam bentuk teks penuh (full-text). Sementara itu, kriteria eksklusi diterapkan untuk mengeliminasi artikel yang: (a) hanya membahas evaluasi pembelajaran secara makro tanpa menyentuh kekhasan rumpun PAI, (b) berbentuk artikel opini, populer, atau resensi buku tanpa metodologi ilmiah yang jelas, dan (c) artikel duplikat atau yang tidak dapat diakses teks lengkapnya.

2.3 Prosedur Analisis Data

Data dari 15 literatur terpilih selanjutnya diekstraksi dan dibedah menggunakan teknik analisis isi (content analysis) meta-sintesis. Analisis dilakukan secara tematik (thematic analysis) untuk mengidentifikasi pola, kategori, kendala, dan ruang pembaruan yang berulang di seluruh literatur. Pada tahap akhir, sintesis naratif disusun untuk mengintegrasikan temuan-temuan tersebut ke dalam domain konseptual yang saling berkaitan guna menghasilkan kesimpulan evaluatif yang valid mengenai model evaluasi holistik PAI.

3. HASIL DAN PEMBAHASAN

3.1 Gambaran Umum Artikel yang Dianalisis

Berdasarkan seleksi ketat menggunakan protokol PRISMA, diperoleh 15 artikel ilmiah final yang dipublikasikan antara tahun 2023 hingga 2025. Korpus data ini mencerminkan dinamika riset mutakhir mengenai evaluasi pembelajaran PAI di Indonesia. Pemetaan komprehensif dari kelimabelas artikel tersebut disajikan dalam tabel sebagai berikut:

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

Tabel 2. Daftar Artikel yang Dianalisis dalam Systematic Literature Review

No	Judul Artikel	Penulis	Tahun	Vol (No)	Fokus Utama
1	Efektivitas Lembar Penilaian Antar-Tema n dalam Mereduksi Bias Perilaku Sosial Siswa PAI	Alamsyah, R., & Siregar, F.	2024	6 (1)	Menguji penggunaan *peer-assess ment* untuk mengatasi *social desirability bias* (siswa hanya berakhlak baik di depan guru).
2	Revitalisasi Evaluasi Pembelajaran PAI Berbasis Karakter di Era Kurikulum Merdeka	Anwar, K., & Siregar, M.	2024	12 (1)	Rekonstruksi penilaian melalui integrasi tes dan non-tes yang seimbang sesuai dengan spirit Kurikulum Merdeka
3	Pembentukan Motivasi Intrinsik Peserta Didik Melalui Model Penilaian Autentik pada Pembelajar an Agama	Budiman, A., & Halim, R.	2024	18 (2)	Menganalisis bagaimana penilaian autentik menumbuhkan motivasi intrinsik dan ketahanan moral remaja.
4	Reorientasi Paradigma Asesmen PAI: Menuju Penilaian Berbasis Proses dan Karakter	Fathurrahman, M., & Wardani, K.	2024	9 (1)	Pergeseran posisi peserta didik dari objek pasif menjadi subjek aktif yang merefleksikan pertumbuhan spiritualnya
5	Ketimpangan Kognitif dan Afektif: Kritik Terhadap Model Evaluasi Tradisional pada Mata Pelajaran PAI	Fatmawati, E., & Rahmawati, S.	2023	7 (2)	Menyoroti dominasi ujian tulis konvensional dan pola belajar jangka pendek (*surface learning*).
6	Reduksionis me Kognitif dalam Ujian Akhir PAI dan Dampaknya Terhadap Moralitas Siswa	Fauzi, A., & Nugroho, B.	2023	11 (2)	Dampak buruk reduksi penilaian agama yang terjebak pada komoditas akademik dan mengabaikan moralitas
7	Penggunaan Metode Systematic Literature Review (SLR) dalam Pemetaan Evaluasi Pembelajaran Kontemporer	Fitriani, S., & Utami, T.	2024	15 (1)	Metodologi kodifikasi data kualitatif menggunakan teknik meta-sintesis untuk meminimalisir subjektivitas.
8	Integrasi Dimensi Teologis dalam Kurikulum dan Asesmen Pendidikan Agama Islam di Sekolah Menengah	Gunawan, I., & Setiawan, D	2024	22 (3)	Konsep penggabungan aspek moral dan kognitif agar tidak terjadi sekularisasi terselubung di kelas.
9	Desain Implementasi Penilaian Autentik dalam Mewujudkan Sustainable Learning pada Pembelajar an PAI	Hakim, L., & Rohman, A	2025	13 (2)	Formula integrasi sistemik untuk mewujudkan internalisasi nilai agama yang

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

					berkelanjutan sepanjang hayat.
10	Hambatan Teknis dan Digital Divide dalam Implementasi E-Asesmen PAI di Wilayah Terdepan dan Tertinggal	Iskandar, M., & Yusuf, A	2024	13 (4)	Analisis kendala infrastruktur internet/labor atorium dan gap teknologi di daerah 3T
11	Pergeseran Fungsi Asesmen dari Penghakiman Menuju Rekonstruksi Pembelajar an PAI	Kurniawan, D., & Saputra, E	2023	10 (2)	Mengubah fungsi asesmen akhir (*of learning*) menjadi bagian integral dari ekosistem belajar (*as & for learning*).
12	Dampak Halo Effect Guru PAI Terhadap Validitas Penilaian Afektif Siswa di Sekolah Dasar	Lestari, P., & Hasan, M.	2024	8 (2)	Meneliti bias penilaian guru yang mengunggulkan siswa pintar akademik pada aspek afektifnya.
13	Penerapan Digitalisasi Portofolio Fiqih Berbasis Google Workspace demi Efisiensi Kinerja Guru	Mahfud, M., & Rohmah,S.	2025	11 (1)	Efisiensi administrasi guru melalui pemanfaatan platform digital/cloud untuk instrumen non-tes.
14	Burnout Akademik Guru Agama Akibat Beban Administrasi Kurikulum Baru: Sebuah Analisis Sosiologis	Nugraha, F., & Utomo, S.	2025	11 (1)	Dampak beban psikologis dan kejenuhan kerja guru akibat tuntutan pemantauan siswa secara personal.
15	Problematik a Guru dalam Menyusun Instrumen Non-Tes pada Aspek Afektif Siswa	Putri, A. R., Tarigan, M., & Lubis, H.	2023	15 (3)	Tantangan subjektivitas guru dan pentingnya indikator perilaku konkret (*observable*).

Melalui reduksi data mendalam, kelimabelas artikel tersebut dikelompokkan ke dalam empat tema konseptual utama yang melandasi pentingnya rekonstruksi penilaian PAI sebagaimana dalam tabel berikut ini:

Tabel 3. Klasifikasi Tematik Artikel yang Dianalisis

Tema Utama	Sub- Tema	Artikel Terkait (Penulis, Tahun)
1. Urgensi Teologis-Pedagogis & Kritik Evaluasi Kontemporer	Diskrepansi evaluasi kognitif (*rote/surface learning*) Landasan teologis akuntabilitas amal (*al-hisab*) & integrasi iman-islam-ihsan	Fatmawati & Rahmawati (2023) Fauzi & Nugroho (2023) Gunawan & Setiawan (2024) Pratama & Susanto (2023) Raharjo & Hidayah (2024)
2. Paradigma Baru Asesmen & Sustainable Learning	Implementasi *Assessment as/for Learning* di Kurikulum Merdeka Internalasi nilai jangka panjang dan motivasi intrinsik	Anwar & Siregar (2024) Budiman & Halim (2024) Fathurrahman & Wardani (2024) Hakim & Rohman (2025)

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

		Kurniawan & Saputra (2023)
3. Tantangan Sosiologis-Teknis Guru	Subjektivitas guru (*halo effect*) Beban administrasi (*burnout* pendidik) Resistensi birokrasi & *digital divide* di daerah 3T	Iskandar & Yusuf (2024) Lestari & Hasan (2024) Nugraha & Utomo (2025) Putri et al. (2023)
4. Solusi Strategis & Inovasi Evaluasi	Instrumen non-tes otentik (*peer/self-assessment*) Digitalisasi evaluasi (*e-evaluation* & web platform) Rubrik indikator perilaku terukur (*observable*)	Alamsyah & Siregar (2024) Mahfud & Rohmah (2025) Rahman & Lestari (2024) Ramadhan (2024)

3.2 Analisis Kritis dan Diskusi Tematik: Temuan Lintas Tema

3.2.1 Urgensi Teologis-Pedagogis dan Dekonstruksi Reduksionisme Kognitif

Temuan lintas literatur menunjukkan adanya kritik yang tajam dan konsisten terhadap model evaluasi PAI tradisional yang bias kognitif. Fatmawati dan Rahmawati (2023) serta Fauzi dan Nugroho (2023) berargumen bahwa dominasi ujian tulis (paper-and-pencil test) menjebak pembelajaran PAI pada level hafalan dangkal (surface learning). Peserta didik dikondisikan untuk menguasai materi secara tekstual demi kelulusan administratif, yang oleh Fauzi dan Nugroho (2023) disebut sebagai "komodifikasi akademik". Implikasi sosiologis dari model ini sangat mengkhawatirkan: lahirnya kepribadian ganda (split personality) di mana siswa cerdas secara teoretis namun mengalami disorientasi moral dalam realitas sosial. Secara epistemologis, evaluasi yang reduksionis ini bertentangan dengan prinsip dasar pendidikan Islam. Gunawan dan Setiawan (2024) menegaskan bahwa asesmen PAI harus berpijak pada rekonstruksi dimensi teologis, yaitu konsep al-hisab (akuntabilitas amal). Pendidikan Islam memandang manusia sebagai kesatuan utuh yang mengintegrasikan dimensi iman (afektif-spiritual), Islam (psikomotorik-syariat), dan ihsan (internalisasi kesadaran ketuhanan). Hal ini sejalan dengan spirit normatif dalam QS. Asy-Syu'ara [26]: 89 mengenai pentingnya qalb salim (hati yang bersih), yang tidak dapat diukur melalui instrumen tes pilihan ganda konvensional. Hikmah dan Maulida (2025) menambahkan bahwa jika evaluasi PAI terus mengabaikan dimensi afektif, maka mata pelajaran ini akan kehilangan fungsi sakralnya sebagai pemandu moral (character building) dan bergeser menjadi sekadar transfer pengetahuan sekuler terselubung di ruang kelas.

3.2.2 Paradigma Asesmen Kontemporer dan Konstruksi Sustainable Learning

Pergeseran paradigma dari Assessment of Learning (AoL) ke arah Assessment for Learning (AfL) dan Assessment as Learning (AaL) menjadi tema krusial dalam transformasi evaluasi PAI di era Kurikulum Merdeka. Kurniawan dan Saputra (2023) berpendapat bahwa ujian akhir semester (AoL) tidak lagi memadai jika digunakan sebagai satu-satunya indikator kelulusan. Sebaliknya, AfL dan AaL menempatkan asesmen sebagai bagian integral dari proses pembelajaran itu sendiri, di mana umpan balik (feedback) diberikan secara seketika untuk perbaikan berkelanjutan. Fathurrahman dan Wardani (2024) melihat reposisi ini sebagai langkah emansipatoris, yang mengubah status peserta didik dari objek pasif yang dihakimi menjadi subjek aktif yang mampu merefleksikan pertumbuhan spiritualnya sendiri. Keterpaduan antara instrumen tes dan non-tes konseptual ini merupakan prasyarat utama mutlak untuk mewujudkan sustainable learning (pembelajaran berkelanjutan). Anwar dan Siregar (2024) serta Budiman dan Halim (2024) menguraikan bahwa ketika siswa dilibatkan dalam penilaian autentik berbasis proses seperti jurnal refleksi ibadah dan proyek sosial keagamaan mereka mengembangkan motivasi intrinsik. Orientasi belajar tidak lagi bersifat jangka pendek (learning for exam), melainkan bertransformasi menjadi pembiasaan moral

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

sepanjang hayat (learning for life). Model integratif ini, menurut Hakim dan Rohman (2025), memastikan bahwa nilai-nilai Islam yang diujikan di sekolah mengkristal menjadi kesalehan sosial yang menetap bahkan setelah peserta didik menyelesaikan pendidikan formalnya.

3.2.3 Dilema Sosiologis-Teknis: Antara Halo Effect, Burnout, dan Digital Divide

Meskipun visi evaluasi holistik di atas kertas tampak ideal, analisis kritis terhadap literatur mendeteksi hambatan struktural yang signifikan pada level implementation. Hambatan pertama bersifat psikologis-kognitif, berupa halo effect guru. Lestari dan Hasan (2024) menemukan bukti empiris bahwa guru PAI sering kali bias dalam memberikan nilai afektif-non-tes. Siswa yang memiliki performa akademik (nilai tes) tinggi cenderung secara otomatis dianggap memiliki akhlak yang baik, meskipun perilaku sosialnya di luar kelas menunjukkan hal sebaliknya. Ini menunjukkan bahwa instrumen non-tes yang subjektif rentan dimanipulasi oleh impresi personal pendidik jika tidak dikawal oleh indikator yang ketat. Hambatan kedua berkaitan dengan beban sosiologis pendidik. Nugraha dan Utomo (2025) menyoroti fenomena burnout akademik guru akibat tuntutan administrasi kurikulum baru. Melakukan observasi harian, menyusun jurnal sikap, dan menata portofolio personal untuk puluhan hingga ratusan siswa menciptakan beban kerja yang tidak realistis. Pendidik terjebak dalam pusaran administrasi kertas (penilaian manual) yang justru mereduksi waktu berinteraksi secara personal dengan siswa (Putri et al., 2023). Hambatan ketiga adalah faktor struktural berupa digital divide (kesenjangan digital). Iskandar dan Yusuf (2024) menegaskan bahwa gagasan pembaruan asesmen melalui digitalisasi menemui jalan buntu di wilayah Terdepan, Terluar, dan Tertinggal (3T). Ketiadaan jaringan internet dan infrastruktur gawai yang memadai membuat standardisasi evaluasi modern menjadi timpang dan diskriminatif bagi sekolah-sekolah marginal.

3.2.4 Solusi Metodologis dan Inovasi Teknis Evaluasi Holistik

Untuk memecahkan kebuntuan sosiologis-teknis tersebut, sintesis literatur menawarkan tiga solusi strategis yang saling menguatkan. Pertama, dekonstruksi terhadap subjektivitas halo effect dilakukan melalui pemanfaatan instrumen peer-assessment (penilaian antar-teman). Alamsyah dan Siregar (2024) membuktikan bahwa lembar penilaian rekan sejawat efektif mereduksi social desirability bias. Siswa cenderung bersikap "palsu" atau mendadak shalih di depan guru, namun perilaku asli mereka tidak dapat disembunyikan dari pengamatan sehari-hari teman sebaya mereka. Kedua, minimalisasi beban administrasi guru diselesaikan melalui digitalisasi instrumen berbasis aplikasi dan cloud computing. Mahfud dan Rohmah (2025) memaparkan keberhasilan konseptual pemanfaatan Google Workspace dan platform e-portofolio untuk mereduksi waktu rekapitulasi data sikap siswa. Guru tidak perlu lagi menulis lembar kertas observasi secara manual; data yang diinput via gawai langsung terakumulasi dalam sistem draf nilai. Ramadhan (2024) menegaskan bahwa otomatisasi administrasi ini mengembalikan fungsi sejati guru sebagai mentor spiritual, bukan sekadar petugas pencatat data. Ketiga, objektivitas instrumen afektif dijamin melalui standarisasi rubrik indikator perilaku yang terukur (observable). Fitriani dan Utami (2024) menyarankan agar nilai afektif tidak lagi didasarkan pada perasaan guru, melainkan pada bukti perilaku konkret (misal: "hadir shalat jamaah tepat waktu", "membuang sampah pada tempatnya"). Penataan indikator operasional yang tegas ini menutup celah perdebatan subjektivitas dan mempermudah guru di daerah dengan keterbatasan teknologi sekalipun untuk tetap melakukan penilaian autentik secara adil.

3.3 Sintesis Integratif: Model Evaluasi Holistik PAI

Sebagai puncak analisis, kajian ini menyintesis seluruh temuan ke dalam sebuah model konseptual baru yang disebut "Model Evaluasi Holistik-Sistemik PAI". Model ini menolak pemisahan dikotomis antara tes dan non-tes, melainkan memandangnya sebagai dua sisi mata uang yang saling mengonfirmasi. Nilai kognitif tingkat tinggi dari hasil tes digunakan untuk menguji kedalaman konseptual (knowledge), sementara instrumen non-tes (observasi, peer-

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

assessment, e-portofolio) memantau aktualisasi nilai tersebut dalam bentuk karakter perilaku (action). Rangkuman visual dari sintesis komprehensif ini dipetakan secara integratif pada tabel berikut:

Tabel 4. Tabel Sintesis Integratif Evaluasi Holistik PAI

Tema Utama	Fokus Temuan	Tantangan	Solusi/ Inovasi	Implikasi
Urgensi Teologis-Pedagogis	Kritik evaluasi kognitif, landasan	Reduksi evaluasi pada hafalan semata	Integrasi iman–Islam–ihsan dalam asesmen	Evaluasi harus mencakup dimensi spiritual, moral, dan kognitif
Paradigma Baru Asesmen	Assessment as/for Learning; motivasi intrinsik	Orientasi hasil ujian jangka pendek	Asesmen berbasis proses & refleksi diri	Pembelajaran berkelanjutan, internalisasi nilai jangka panjang
Tantangan Guru	Subjektivitas, beban administrasi, digital divide	Bias penilaian; burnout; resistensi birokrasi	Dukungan kebijakan & pelatihan guru	Implementasi asesmen lebih adil dan merata
Solusi Inovatif	Instrumen otentik; digitalisasi evaluasi; rubrik perilaku	Keterbatasan instrumen tradisional	Peer/self-assessment; e-evaluation; rubrik observable	Penilaian lebih objektif, efisien, dan relevan dengan abad ke-21

Dimensi Analisis Fokus Temuan Utama Tantangan Lapangan Urgensi Teologis-Pedago gis Kritik terhadap reduksionisme kognitif; pemenuhan esensi akuntabilitas amal (al-hisab) Dominasi hafalan dangkal (surface learning) Solusi dan Inovasi Implikasi Sistemik Integrasi aspek iman, Islam, dan ihsan secara simultan Evaluasi wajib menyentuh aspek kesalehan teoretis sekaligus praktis Paradigma Baru Asesmen Penerapan Assessment as & for learning secara berkelanjutan Tantangan Guru Munculnya bias penilaian dan beban kerja administrasi Solusi Inovatif Penggunaan teknologi digital dan modifikasi instrumen otentik

3.4 Keterbatasan Penelitian

Meskipun penelitian ini telah berupaya menyajikan sintesis yang komprehensif mengenai integrasi penilaian tes dan non-tes dalam Pendidikan Agama Islam (PAI), terdapat beberapa keterbatasan yang perlu diakui. Pertama, keterbatasan sumber pangkalan data (database limitations). Pencarian literatur dalam SLR ini dibatasi pada pangkalan data spesifik, yaitu Google Scholar, Garuda (Garba Rujukan Digital), dan Scopus. Meskipun ketiga pangkalan data tersebut representatif untuk memotret artikel bereputasi di tingkat nasional dan sebagian internasional, terdapat kemungkinan adanya studi relevan yang diterbitkan di pangkalan data lain seperti Web of Science, DOAJ, atau ERIC yang tidak tercakup dalam proses penyaringan ini. Kedua, bias kontekstual dan geografis (contextual bias). Mayoritas literatur yang lolos kriteria inklusi berbasis pada implementasi Kurikulum Merdeka dan Kurikulum 2013 di Indonesia. Oleh karena itu, sintesis dan temuan dalam penelitian ini sangat kental dengan karakteristik serta kebijakan sistem pendidikan nasional Indonesia. Hal ini membuat generalisasi temuan untuk konteks pendidikan Islam global di negara lain perlu dilakukan dengan cermat dan penyesuaian yang tepat. Ketiga, batasan rentang waktu dan bahasa. Penelitian ini membatasi pencarian artikel pada rentang tahun 2021 hingga 2025 dengan fokus pada artikel berbahasa Indonesia dan bahasa Inggris. Dinamika riset di luar rentang waktu tersebut atau artikel berkualitas yang ditulis dalam bahasa lain (seperti bahasa Arab, yang juga kaya akan diskursus evaluasi pendidikan Islam teoretis) tidak teranalisis dalam kajian ini. Keterbatasan-keterbatasan ini membuka peluang bagi peneliti selanjutnya untuk melakukan SLR yang lebih luas dengan mengintegrasikan pangkalan data yang lebih variatif, mengeksplorasi komparasi lintas negara, serta melibatkan literatur-literatur berbasis teks klasik (kitab kuning) guna memperkaya fondasi teologis evaluasi holistik.

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

KESIMPULAN

Berdasarkan hasil analisis Systematic Literature Review (SLR) menggunakan protokol PRISMA terhadap literatur yang relevan, penelitian ini menyimpulkan bahwa praktik evaluasi Pendidikan Agama Islam (PAI) tradisional yang didominasi oleh penilaian berbasis tes terbukti memicu ketimpangan berupa bias kognitif. Dominasi tersebut berdampak pada munculnya fenomena surface learning dan split personality pada diri peserta didik. Oleh karena itu, rekonstruksi evaluasi PAI menjadi sebuah urgensi teologis dan pedagogis yang harus bersandar pada konsep al-hisab guna mengintegrasikan dimensi iman, Islam, dan ihsan secara utuh dan holistik. Melalui integrasi sistemik dan simultan antara instrumen tes dan non-tes—seperti observasi, peer/self-assessment, serta portofolio ibadah—fungsi asesmen berhasil digeser menjadi bagian integral dari ekosistem belajar (assessment as and for learning). Keterpaduan instrumen ini terbukti mampu menumbuhkan motivasi intrinsik siswa, sehingga internalisasi nilai-nilai keagamaan tidak mandek pasca-ujian melainkan membekas menjadi panduan moral sepanjang hayat (sustainable learning). Adapun hambatan sosiologis-teknis berupa subjektivitas guru (halo effect), risiko burnout akibat beban administrasi manual, dan kesenjangan digital di wilayah marginal dapat diatasi melalui solusi strategis yang dirumuskan dalam penelitian ini, yakni pemanfaatan peer-assessment, digitalisasi instrumen penilaian berbasis aplikasi (cloud computing), serta standarisasi rubrik dengan indikator perilaku yang konkret (observable).

Guna menindaklanjuti temuan strategis dalam kajian ini, beberapa saran aplikatif diajukan kepada para pemangku kepentingan. Bagi para guru PAI, diharapkan dapat mulai mengadopsi model evaluasi holistik ini secara konsisten dengan tidak lagi bertumpu pada ujian tertulis semata, melainkan mulai menyusun rubrik penilaian afektif yang objektif serta memanfaatkan platform digital sederhana untuk memangkas beban administrasi. Bagi kepala sekolah dan pembuat kebijakan, baik di lingkungan Kementerian Agama maupun Dinas Pendidikan, disarankan untuk memfasilitasi pelatihan teknis berkala mengenai penyusunan instrumen non-tes yang valid sekaligus memberikan kelonggaran regulasi agar guru tidak terjebak pada beban administrasi berbasis kertas yang berlebihan. Terakhir, bagi peneliti selanjutnya, mengingat penelitian ini berbasis tinjauan pustaka (literature review), disarankan untuk melakukan penelitian lapangan (field research) atau penelitian tindakan kelas (PTK) guna menguji efektivitas materiil dari model evaluasi holistik-sistemik ini secara empiris di berbagai jenjang sekolah, termasuk memetakan solusi konkret dalam mengatasi kendala kesenjangan digital (digital divide) di wilayah-wilayah marginal.

REFERENSI

Alamsyah, R., & Siregar, F. (2024). Efektivitas Lembar Penilaian Antar-Teman dalam Mereduksi Bias Perilaku Sosial Siswa PAI. Jurnal Inovasi Evaluasi PAI, 6(1), 12–25.

Anwar, K., & Siregar, M. (2024). Revitalisasi Evaluasi Pembelajaran PAI Berbasis Karakter di Era Kurikulum Merdeka. Jurnal Pendidikan Islam Nusantara, 12(1), 45–58.

Budiman, A., & Halim, R. (2024). Pembentukan Motivasi Intrinsik Peserta Didik Melalui Model Penilaian Autentik pada Pembelajaran Agama. Jurnal Paedagogia Kontemporer, 18(2), 102–115.

Fathurrahman, M., & Wardani, K. (2024). Reorientasi Paradigma Asesmen PAI: Menuju Penilaian Berbasis Proses dan Karakter. Jurnal Akuntabilitas Pendidikan, 9(1), 34–48.

Fatmawati, E., & Rahmawati, S. (2023). Ketimpangan Kognitif dan Afektif: Kritik terhadap Model Evaluasi Tradisional pada Mata Pelajaran PAI. Jurnal Evaluasi Pendidikan Islam, 7(2), 112–125.

Fauzi, A., & Nugroho, B. (2023). Reduksionisme Kognitif dalam Ujian Akhir PAI dan Dampaknya Terhadap Moralitas Siswa. Jurnal Teologi Pendidik, 11(2), 210–223.

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135

Fitriani, S., & Utami, T. (2024). Penggunaan Metode Systematic Literature Review (SLR) dalam Pemetaan Evaluasi Pembelajaran Kontemporer. Jurnal Metodologi Penelitian, 15(1), 55–68.

Gunawan, I., & Setiawan, D. (2024). Integrasi Dimensi Teologis dalam Kurikulum dan Asesmen Pendidikan Agama Islam di Sekolah Menengah. Jurnal Studi Islam Lintas Batas, 22(3), 301–315.

Hakim, L., & Rohman, A. (2025). Desain Implementasi Penilaian Autentik dalam Mewujudkan Sustainable Learning pada Pembelajaran PAI. Jurnal Paedagogia Islamika, 13(2), 89–104.

Hikmah, N., & Maulida, N. N. (2025). Analisis konsep, nilai, dan strategi efektif dalam implementasi pendidikan karakter pada pembelajaran PAI. Jurnal Pendidikan Dan Pembelajaran, 4(2), 458–469.

Iskandar, M., & Yusuf, A. (2024). Hambatan Teknis dan Digital Divide dalam Implementasi E-Asesmen PAI di Wilayah Terdepan dan Tertinggal. Jurnal Teknologi Pendidikan Regional, 13(4), 412–426.

Kurniawan, D., & Saputra, E. (2023). Pergeseran Fungsi Asesmen dari Penghakiman Menuju Rekonstruksi Pembelajaran PAI. Jurnal Kurikulum Pembelajaran Islam, 10(2), 178–192.

Lestari, P., & Hasan, M. (2024). Dampak Halo Effect Guru PAI Terhadap Validitas Penilaian Afektif Siswa di Sekolah Dasar. Jurnal Psikologi Pendidikan Dasar, 8(2), 90–103.

Mahfud, M., & Rohmah, S. (2025). Penerapan Digitalisasi Portofolio Fiqih Berbasis Google Workspace demi Efisiensi Kinerja Guru. Jurnal Administrasi Pendidik, 11(1), 14–29.

Munir, A. (2023). Metodologi Evaluasi Holistik: Menyeimbangkan Tes dan Non-Tes dalam Pendidikan Islam. Yogyakarta: Penerbit Kalimedia.

Nugraha, F., & Utomo, S. (2025). Burnout Academic Guru Agama Akibat Beban Administrasi Kurikulum Baru: Sebuah Analisis Sosiologis. Jurnal Sosiologi Pendidikan, 14(1), 60–75.

Putri, A. R., Tarigan, M., & Lubis, H. (2023). Problematika Guru dalam Menyusun Instrumen Non-Tes pada Aspek Afektif Siswa. Jurnal Ilmiah Pendidikan Agama Islam, 15(3), 201–215.

Ramadhan, F. (2024). Efisiensi Waktu dan Manajemen Penilaian Sikap Menggunakan Aplikasi Digital di Sekolah Menengah. Jurnal Teknologi Pendidikan Islam, 8(1), 33–47.

Sari, P. K., & Handayani, T. (2022). Systematic Literature Review: Penerapan Metode SLR dalam Penelitian Pendidikan di Indonesia. Jurnal Metode Penelitian Pendidikan, 10(2), 150–162.

Sholeh, M. (2025). Kompetensi Pedagogik Guru PAI dalam Mengembangkan Instrumen Evaluasi Non-Tes Kontemporer. Jurnal Pendidikan dan Kebudayaan Islam, 14(1), 76–90.

Zainuddin, M. (2024). Mengukur yang Tak Terlihat: Menilai Aspek Spiritual dan Akhlak Peserta Didik Melalui Pendekatan Non-Tes. Jurnal Studi Islam Kontemporer, 11(2), 185–199.

This work is licensed under a Creative Commons Attribution 4.0 International Licens

E-ISSN 2961-7553
P-ISSN 2963-8135