TSAQOFAH
Jurnal Penelitian Guru Indonesia

p-ISSN: 2810-0395
e-ISSN: 2810-0042

Terindeks : Dimensions, Scilit, Lens, Semantic, Garuda, Google Scholar, Base, etc

https://doi.org/10.58578/tsaqofah.v6i1.8441

MENGUKUR RUH DAN RAGA: PERAN EVALUASI HOLISTIK PAI DALAM KURIKULUM MERDEKA

Measuring Soul and Body: The Role of Holistic Islamic Religious Education Evaluation in the *Kurikulum Merdeka*

Siti Sangadah, Rachma Rusdiana Dewi, Ahmad Rosyid Ridho
Institut Islam Mamba'ul 'Ulum Surakarta
sangadahsunardi@gmail.com; rachmarusdianadewi@gmail.com

Article Info:

Submitted:	Revised:	Accepted:	Published:
Nov 11, 2025	Dec 3, 2025	Dec 15, 2025	Dec 20, 2025

Abstract

Evaluation in Islamic Religious Education (PAI) often remains focused solely on measuring the cognitive domain, even though the primary goal of PAI is the formation of character and the holistic internalization of religious values. This study aimed to analyze the strategic role of holistic evaluation and the comprehensive evaluation models relevant to improving the quality of PAI, particularly in the context of the *Kurikulum Merdeka*. The method employed was library research with content analysis of various recent literature on PAI evaluation and curriculum development. The findings show that PAI evaluation should ideally be multidimensional, integrating the cognitive, affective, and psychomotor domains in a unified manner. Synergy among the CIPP Model (for systemic evaluation), the Kirkpatrick Model (for measuring behavioral change), and Tyler's Goal Attainment Model (for assessing the achievement of objectives) is required to ensure the effectiveness and relevance of PAI programs. This study concludes that evaluation functions not only as a tool for measuring learning outcomes but also as a strategic instrument for curriculum development, the

Tsaqofah journal is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License

strengthening of instructional practice, and the assurance of sustainable quality in Islamic education within the framework of *Kurikulum Merdeka*.

Keywords: Evaluation of Islamic Religious Education (PAI); CIPP Model; Kirkpatrick Model; *Kurikulum Merdeka*; Holistic Education

Abstrak: Evaluasi dalam Pendidikan Agama Islam (PAI) kerap terfokus pada pengukuran ranah kognitif semata, padahal tujuan utama PAI adalah pembentukan karakter dan internalisasi nilai keagamaan secara utuh. Penelitian ini bertujuan menganalisis peran strategis evaluasi holistik serta model-model evaluasi komprehensif yang relevan dalam peningkatan mutu PAI, khususnya dalam konteks *Kurikulum Merdeka*. Metode yang digunakan adalah studi pustaka (*library research*) dengan analisis isi terhadap berbagai literatur mutakhir terkait evaluasi PAI dan pengembangan kurikulum. Hasil penelitian menunjukkan bahwa evaluasi PAI idealnya bersifat multidimensi, mencakup ranah kognitif, afektif, dan psikomotorik secara terpadu. Sinergi antara Model CIPP (untuk evaluasi sistemik), Model Kirkpatrick (untuk pengukuran perubahan perilaku), dan Model *Goal Attainment* Tyler (untuk penilaian ketercapaian tujuan) diperlukan guna menjamin efektivitas dan relevansi program PAI. Penelitian ini menyimpulkan bahwa evaluasi tidak hanya berfungsi sebagai alat ukur hasil belajar, melainkan sebagai instrumen strategis untuk pengembangan kurikulum, penguatan praktik pembelajaran, dan penjaminan mutu pendidikan Islam yang berkelanjutan dalam kerangka *Kurikulum Merdeka*.

Kata Kunci: Evaluasi PAI; Model CIPP; Model Kirkpatrick; Kurikulum Merdeka; Pendidikan Holistik

PENDAHULUAN

Pendidikan Agama Islam (PAI) merupakan bagian inti dari kurikulum pendidikan nasional yang memiliki peranan fundamental dalam membentuk karakter, akhlak, dan kepribadian peserta didik. Sebagai mata pelajaran yang bertujuan utama menumbuhkan keimanan, ketakwaan, dan akhlak mulia, PAI tidak hanya berfungsi mentransfer pengetahuan keagamaan, tetapi juga membina peserta didik agar mampu menginternalisasikan nilai-nilai Islam dalam perilaku sehari-hari mereka. Oleh sebab itu, keberhasilan pembelajaran PAI tidak diukur semata dari kemampuan menghafal ayat atau konsep keagamaan, melainkan tercermin pada kemampuan peserta didik untuk memahami, menghayati, dan mengamalkan ajaran Islam dalam kehidupan nyata. Dalam kerangka inilah evaluasi mendapatkan posisi yang sangat strategis sebagai instrumen untuk memastikan bahwa seluruh proses pembelajaran berjalan sesuai arah dan tujuan yang telah ditetapkan. Evaluasi dalam pembelajaran PAI melampaui fungsi sebagai alat ukur hasil belajar, melainkan sebagai proses menyeluruh yang memantau perkembangan belajar peserta didik sejak awal, selama, hingga akhir pembelajaran (Ikbal dkk., 2025).

Permasalahan utama yang mendasari kajian ini adalah kesenjangan antara tuntutan tujuan PAI yang holistik dengan kebutuhan metode dan model evaluasi yang tepat dan komprehensif. Tujuan PAI menuntut keberhasilan diukur pada tiga ranah, yaitu ranah kognitif, ranah afektif (sikap dan akhlak), dan ranah psikomotorik (keterampilan praktik ibadah). Tuntutan pengukuran pada ranah afektif dan psikomotorik ini lebih kompleks dan sulit dibandingkan kognitif. Kesenjangan ini semakin diperkuat dalam konteks Kurikulum Merdeka, di mana evaluasi dituntut menggunakan asesmen autentik dan fokus pada perkembangan karakter serta proses berpikir (Majid, 2014). Tuntutan ini menimbulkan tantangan serius dalam memastikan evaluasi mampu memberikan gambaran komprehensif.

Mengingat kompleksitas ini, evaluasi tidak dapat dilakukan secara sembarangan, sehingga muncul pertanyaan mengenai model evaluasi apa yang paling efektif dan bagaimana model-model seperti Model CIPP (*Context, Input, Process, Product*), Model Kirkpatrick, dan Model Goal Attainment (Tyler) dapat diterapkan guna menjamin tercapainya tujuan PAI secara optimal. Oleh karena itu, kajian mendalam mengenai peranan dan model-model evaluasi ini sangat penting dilakukan untuk memberikan pemahaman mengenai bagaimana evaluasi dapat digunakan sebagai instrumen strategis dalam meningkatkan mutu pembelajaran, sehingga PAI dapat dijalankan secara lebih efektif dan mampu membentuk peserta didik yang beriman, bertakwa, berilmu, dan berakhlak mulia.

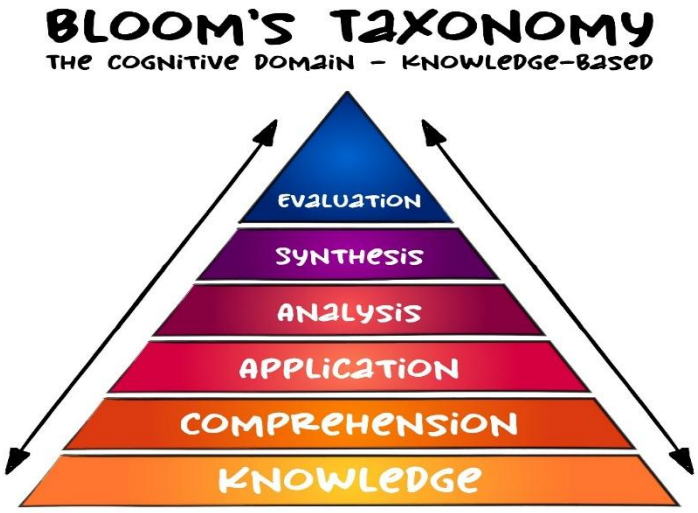
METODE

Penelitian ini menggunakan metode Studi Pustaka (*Library Research*) dengan pendekatan kualitatif deskriptif. Penelitian kepustakaan menjadi cara yang tepat untuk menghasilkan karya ilmiah (Sari, & Asmendri, 2020). Teknik dokumentasi adalah satu metode pengumpulan data dengan melihat atau menganalisis dokumen dokumen yang di buat oleh subjek sendiri atau oleh orang lain (Sugiyono 2020). Data dikumpulkan melalui teknik dokumentasi terhadap sumber sekunder berupa buku teks, jurnal ilmiah terindeks, dan dokumen kurikulum resmi. Kriteria sumber pustaka difokuskan pada literatur mutakhir (rentang 2015–2025) guna menjaga relevansi dengan Kurikulum Merdeka. Analisis data dilakukan secara sistematis melalui tiga tahap: reduksi data (memilah konsep evaluasi yang relevan), display data (membandingkan model CIPP, Kirkpatrick, dan Tyler), serta penarikan kesimpulan (sintesis model evaluasi holistik untuk PAI). Validitas data dijaga melalui teknik triangulasi sumber pustaka untuk memastikan konsistensi teori yang diangkat.

HASIL

Peran Evaluasi PAI Bersifat Holistik dan Multidimensi

Temuan ini merupakan kesimpulan krusial dari kajian literatur, menekankan bahwa peran evaluasi dalam Pendidikan Agama Islam (PAI) harus dilaksanakan secara menyeluruh dan terpadu, jauh melampaui sekadar mengukur kemampuan mengingat atau menghafal. Sifat holistik dan multidimensi ini mutlak diperlukan karena tujuan PAI adalah transformasi perilaku dan pembentukan karakter, yang mencakup seluruh aspek kepribadian siswa.



Evaluasi PAI wajib mengintegrasikan pengukuran pada tiga ranah pembelajaran utama yang saling terkait, sesuai dengan taksonomi tujuan pendidikan. Evaluasi merupakan komponen penting dalam proses pembelajaran, karena mengukur sejauh mana tujuan pembelajaran kognitif, afektif, dan psikomotor telah tercapai. Dalam praktiknya, evaluasi PAI masih cenderung menekankan aspek kognitif, sementara domain afektif dan psikomotor kurang mendapat perhatian (Faelasup & Astuti, 2025). Ranah tersebut meliputi Ranah Kognitif yang mengukur pemahaman peserta didik terhadap konsep-konsep dasar keagamaan dan pengetahuan Al-Qur'an serta Hadis. Namun, fokus utama terletak pada Ranah Afektif, yang merupakan jantung dari PAI, di mana evaluasi harus mampu mengukur penghayatan, internalisasi nilai, dan manifestasi sikap keagamaan siswa, seperti kejujuran, toleransi, kedisiplinan dalam beribadah, dan tanggung jawab sosial. Penting untuk menanamkan nilai-nilai agama pada siswa sekolah dasar. Anak-anak sekolah dasar mengembangkan sikap keagamaan mereka, hal itu akan sangat berpengaruh dalam

melanjutkan pertumbuhan dan perkembangan mereka di tingkat yang lebih tinggi (Nurdiyanto, et al., 2024). Terakhir adalah Ranah Psikomotorik, yang mengukur keterampilan praktik ibadah, seperti ketepatan gerakan salat, wudu, dan kemahiran membaca Al-Qur'an. Penilaian di ranah ini harus dilakukan melalui praktik langsung dan performance assessment.

Peran evaluasi dalam PAI bersifat ganda: ia tidak hanya menilai hasil akhir, tetapi juga berfungsi sebagai proses perbaikan yang berkelanjutan. Evaluasi bertindak sebagai alat diagnosis yang sangat penting. Evaluasi berperan sebagai instrumen yang tepat untuk menilai seberapa efektif proses pengajaran dan memberikan umpan balik formatif yang sangat penting bagi pendidik. (Utami et al., 2025). Peran dari evaluasi pembelajaran adalah untuk menilai sejauh mana tujuan pembelajaran tercapai, memberikan masukan, mengidentifikasi kelemahan dan kekuatan, serta memandu perbaikan dalam pembelajaran. (Ansya et al., 2024). Sebagai evaluator, guru harus mampu menilai kemampuan siswa selama proses pembelajaran (Ariefky & Inayati, 2023). Dengan menilai secara multidimensi, guru dapat mengidentifikasi secara spesifik akar permasalahan siswa, misalnya apakah kesulitan ada pada pemahaman konsep (*kognitif*) atau pada penerapannya (*afektif/psikomotorik*). Informasi diagnostik ini memungkinkan guru memberikan remedial atau pengayaan yang tepat sasaran, sejalan dengan prinsip *assessment for learning* dalam Kurikulum Merdeka.

Selain itu, hasil evaluasi berfungsi sebagai cermin dan umpan balik strategis bagi guru. Data yang diperoleh memberikan gambaran nyata mengenai efektivitas metode, media, dan strategi pengajaran yang digunakan. Metode alternatif seperti observasi, portofolio, jurnal reflektif, penilaian diri, dan evaluasi berbasis proyek keagamaan direkomendasikan untuk menangkap hasil belajar siswa secara lebih akurat (Faelasup & Astuti, 2025). Jika sebagian besar siswa menunjukkan pencapaian afektif yang rendah, guru memiliki dasar yang kuat untuk merefleksikan dan memperbaiki pendekatannya dalam perencanaan pembelajaran berikutnya. Evaluasi multidimensi juga memungkinkan guru mengukur konsistensi antara pengetahuan yang dimiliki siswa dengan perilaku yang ditampilkannya, memastikan PAI berhasil jika pengetahuan (kognitif) terinternalisasi menjadi sikap (afektif) dan diwujudkan dalam perbuatan (psikomotorik). Pengembangan pendekatan evaluasi yang komprehensif, inovatif, dan sesuai dengan karakteristik pendidikan agama Islam untuk memastikan evaluasi mampu memberikan gambaran utuh terhadap pemahaman, penghayatan, dan pengamalan peserta didik yang diarahkan terhadap nilai-nilai agama (Fadli, et al., 2024). Secara keseluruhan, temuan ini menegaskan bahwa sifat holistik dan multidimensi adalah kebutuhan

esensial untuk memastikan bahwa evaluasi PAI benar-benar menggambarkan keberhasilan pendidikan Islam secara utuh dan menjadi katalisator bagi perbaikan mutu pendidikan berkelanjutan.

Tabel 1. Peran Holistik dan Multidimensi Evaluasi PAI

Temuan Utama	Indikator Temuan	Keterangan Temuan
Peran Evaluasi PAI Bersifat Holistik dan Multidimensi	Penilaian mencakup tiga ranah: Kognitif, Afektif, dan Psikomotorik.	Evaluasi wajib mengukur tidak hanya pengetahuan (kognitif), tetapi juga sikap/akhlak (afektif) dan praktik ibadah (psikomotorik) secara terpadu.
Fungsi Diagnostik dan Umpan Balik	Pemberian remedial/pengayaan; refleksi perbaikan metode guru.	Evaluasi berfungsi sebagai alat diagnostik (assessment for learning) yang mengidentifikasi kesulitan siswa, dan sebagai cermin bagi guru untuk perbaikan strategi pengajaran berikutnya.
Konsistensi Karakter	Pengukuran internalisasi nilai dan perubahan perilaku.	PAI berhasil jika terdapat konsistensi antara pengetahuan teoretis dan manifestasi perilaku keagamaan sehari-hari, didukung oleh penilaian autentik dan observasi.

Kesimpulan utama dari kajian ini adalah bahwa evaluasi dalam Pendidikan Agama Islam (PAI) harus diakui sebagai proses esensial dan strategis yang berfungsi mengukur keberhasilan pembelajaran secara utuh, bukan sekadar menilai hasil akhir pengetahuan. Peran ini menuntut pelaksanaan evaluasi secara holistik dan multidimensi, yang berarti evaluasi wajib mengukur tiga ranah pembelajaran secara terpadu dan seimbang: ranah kognitif (pengetahuan), ranah afektif (sikap, akhlak, dan internalisasi nilai), serta ranah psikomotorik (keterampilan dan praktik ibadah). Ranah afektif dan psikomotorik, yang mencerminkan pembentukan karakter, menjadi penentu utama keberhasilan PAI dan menuntut penggunaan teknik penilaian autentik dan observasi. Evaluasi yang dilaksanakan secara holistik ini memberikan fungsi ganda yang krusial, yaitu sebagai alat diagnostik untuk mengidentifikasi kesulitan spesifik siswa guna perbaikan remedial, dan sebagai umpan balik strategis bagi guru untuk merefleksikan serta memperbaiki metode pengajaran mereka. Dengan demikian, peran holistik evaluasi memastikan adanya konsistensi antara pengetahuan teoretis dan manifestasi perilaku keagamaan, yang pada akhirnya menopang upaya peningkatan mutu pendidikan Islam secara berkelanjutan.

Model Evaluasi Komprehensif Diperlukan untuk PAI

Temuan ini menegaskan bahwa untuk menjamin akurasi, validitas, dan relevansi hasil evaluasi Pendidikan Agama Islam (PAI), penggunaan model evaluasi yang sifatnya

komprehensif, terstruktur, dan sistematis adalah suatu keharusan. Mengingat tujuan PAI yang bersifat holistik (mencakup kognitif, afektif, dan psikomotorik) serta tuntutan Kurikulum Merdeka, model evaluasi konvensional tidak lagi memadai. Implementasi paradigma pendidikan holistik dalam pembelajaran Pendidikan Agama Islam berperan dalam mengembangkan: kecerdasan intelektual, kecerdasan spiritual, dan kecerdasan emosional serta kekuatan fisik siswa melalui interaksi pembelajaran (Malili, et. al., 2023). Diperlukan adopsi model yang mampu menilai program secara keseluruhan dan mengukur dampak hingga ke tingkat perubahan perilaku.

Model yang paling komprehensif untuk menilai program PAI secara menyeluruh adalah Model CIPP (*Context, Input, Process, Product*). CIPP merupakan akronim yang terdiri dari : contexs evaluation, input evaluation, process evaluation, 15 product evaluatioan. Setiap tipe evaluasi terikat pada perangkat pegambilan keputusan yang menyangkut perencanaan dan operasi sebuah program (Rusman, 2018). Model yang dikembangkan oleh Stufflebeam ini mengukur program pada empat dimensi krusial. Evaluasi dimulai dengan menganalisis Context (Konteks), yaitu menilai kebutuhan dan kesesuaian tujuan PAI yang dirumuskan dengan kondisi lingkungan sekolah dan tuntutan masyarakat. Selanjutnya, evaluasi Input (Masukan) menilai kesiapan sumber daya, termasuk kualitas kurikulum, kompetensi guru PAI, dan ketersediaan media pembelajaran. Tahap Process (Proses) berfokus pada kualitas pelaksanaan pembelajaran PAI di kelas, menilai implementasi metode pengajaran dan interaksi guru-siswa. Terakhir, evaluasi Product (Produk) menilai hasil akhir program, yang mencakup pencapaian kognitif, afektif, dan psikomotorik siswa. Konteks (perencanaan pembelajaran) guru PAI dipandu oleh surat edaran dinas pendidikan. Masukan (penyusunan dan prosedur) dilakukan secara periodik dan berkelanjutan. Proses dilakukan dengan memfasilitasi pengembangan minat dan potensi siswa. Sedangkan produk (evaluasi hasil belajar siswa) dilakukan secara lisan dan praktik langsung. Sedangkan evaluasi (produk) hasil belajar siswa dilakukan secara lisan dan praktik langsung (Jamali, 2018). Model CIPP memungkinkan identifikasi kelemahan pada tahap mana pun, sehingga memberikan informasi yang akurat untuk perbaikan program, bukan hanya perbaikan hasil belajar.

Selain CIPP, Model Kirkpatrick terbukti sangat relevan bagi PAI karena secara spesifik berfokus pada pengukuran dampak program, terutama pada dimensi perubahan perilaku, yang merupakan tujuan utama pendidikan agama. Dalam pembelajaran PAI, model ini dipandang lebih relevan untuk menilai ranah afektif, seperti sikap religius, kedisiplinan dalam beribadah, serta penerapan akhlak dalam kehidupan sehari-hari (Majid, 2014).Model

yang paling relevan dan efektif adalah integrasi dari Model CIPP dan Model Kirkpatrick, didukung oleh Evaluasi Autentik. Model Evaluasi Kirkpatrick dapat dijadikan fondasi dalam dunia evaluasi program, terus beradaptasi hingga memberikan kerangka kerja yang pada akhirnya tepat sasaran untuk memastikan bahwa setiap program Bimbingan dan Konseling, benar-benar memberikan nilai dan mencapai tujuannya yang telah ditetapkan (Khalisah, 2025) Model evaluasi seperti CIPP dan Kirkpatrick memberikan kerangka kerja, namun implementasinya memerlukan adaptasi terhadap aspek kontekstual yang bervariasi. Solusi untuk tantangan ini melibatkan pengembangan kapasitas evaluasi di tingkat sekolah, partisipasi aktif pemangku kepentingan, dan pembentukan kebijakan evaluasi yang komprehensif (Munandar, et al., 2023). Model Kirkpatrick, efektif dalam meningkatkan kompetensi profesional guru secara keseluruhan sekaligus memberikan dampak positif pada praktik pembelajaran dan menumbuhkan iklim kelas yang lebih toleran dan inklusif (Hidayati & Adien, 2025). Model Evaluasi Kirkpatrick yang mencakup tingkat empat: reaksi, pembelajaran, perilaku, dan hasil (Purwati, et al., 2025). Model ini membantu mengukur Reaction (respons awal siswa), Learning (peningkatan pengetahuan), dan yang paling penting untuk PAI adalah Behavior (Perilaku), yang mengukur apakah ada perubahan nyata pada sikap, akhlak, dan praktik keagamaan siswa dalam kehidupan sehari-hari. Model ini juga mengukur Result (Hasil), yakni dampak jangka panjang program terhadap lingkungan sekolah. Dengan fokus pada pengukuran perilaku dan hasil, Model Kirkpatrick memberikan data empiris yang kuat mengenai efektivitas PAI dalam membentuk karakter siswa. Model evaluasi program yaitu model Kirkpatrick yang akan mengevaluasi secara mendalam, dengan empat tingkatan yaitu tingkat reaksi, tingkat dampak program, tingkat perubahan pembelajaran, dan tingkat hasil (Ridho, et al., 2020).

Model Goal Attainment (Tyler) menekankan bahwa evaluasi merupakan proses sistematis untuk mengetahui sejauh mana tujuan pembelajaran yang telah dirumuskan dapat dicapai oleh peserta didik. Dalam pandangan Tyler, tujuan pembelajaran adalah standar utama yang menjadi acuan bagi seluruh kegiatan pendidikan, mulai dari perencanaan, pelaksanaan, hingga evaluasi. Evaluasi program berorientasi tujuan Tyler dirancang untuk menggambarkan tujuan program yang sukses. Evaluasi juga bermakna untuk menarik kesimpulan yang mencakup tiga hal utama dalam prosesnya, yaitu merumuskan pertanyaan, mengumpulkan data, dan menyajikan atau menampilkan informasi (Bellia, 2022) Oleh karena itu, evaluasi dilakukan dengan cara *membandingkan antara tujuan Pendidikan Agama Islam (PAI) yang telah ditetapkan dengan hasil belajar nyata yang ditunjukkan siswa*. Pendekatan ini tidak hanya

menilai aspek kognitif, tetapi juga afektif dan psikomotor, sehingga hasil evaluasi benar-benar mencerminkan pemahaman keagamaan, sikap, serta keterampilan ibadah yang berkembang pada diri peserta didik. M*enurut Tyler dalam Aguokogbuo (2000), rasionalitas pendidikan, keterampilan belajar untuk mencapai tujuan, keterampilan pengorganisasian pembelajaran, dan penilaian/evaluasi adalah empat hal mendasar yang berkaitan dengan pengembangan kurikulum (Ibeh2022).*

Selain itu, Tyler menegaskan bahwa evaluasi harus bersifat kontinu, berkesinambungan, dan terintegrasi dengan proses pembelajaran. Dengan memanfaatkan evaluasi Tyler, para pendidik memiliki alternatif dalam mendesain evaluasi dan asesmen pembelajaran. (Betu, 2023). Guru PAI perlu memantau perkembangan siswa secara berkelanjutan melalui evaluasi formatif agar dapat mengidentifikasi kesulitan belajar, kelemahan metode mengajar, serta kebutuhan perbaikan strategi pembelajaran. Dalam konteks PAI, pendekatan ini membantu memastikan bahwa kegiatan belajar mengarah pada tercapainya kompetensi yang diharapkan, seperti kemampuan membaca Al-Qur'an dengan baik, memahami hadis, memperkuat akidah, dan mengamalkan akhlak mulia. Jika hasil belajar belum sesuai dengan tujuan, guru dapat melakukan remedial, pengayaan, atau penyesuaian pendekatan pengajaran untuk meningkatkan efektivitas pembelajaran.

Sejalan dengan tuntutan Kurikulum Merdeka dan karakter PAI, model evaluasi komprehensif harus diwujudkan melalui implementasi Evaluasi Autentik (*Authentic Assessment*). Metode ini adalah yang paling sesuai untuk mengukur ranah afektif dan psikomotorik, karena melibatkan penilaian kemampuan siswa dalam konteks nyata. Penerapannya mencakup penggunaan portofolio (mengumpulkan karya terbaik siswa), penilaian proyek (mengerjakan tugas yang membutuhkan aplikasi nilai-nilai agama), dan observasi praktik (penilaian langsung terhadap pelaksanaan ibadah). Penggunaan model-model komprehensif ini memastikan bahwa evaluasi PAI tidak sekadar mengukur apa yang diketahui siswa, tetapi lebih penting, mengukur bagaimana siswa bersikap dan berperilaku berdasarkan pengetahuan agama yang telah dipelajarinya, menjadikannya basis data kredibel untuk perbaikan program dan penjaminan mutu pendidikan Islam. Model pembelajaran Pendidikan Agama Islam (PAI) berbasis kontekstual merupakan pendekatan yang efektif dalam meningkatkan pemahaman dan internalisasi nilai-nilai Islami pada siswa Sekolah Dasar (Ningsih, 2025).

Tabel 2. Model Evaluasi Komprehensif Diperlukan untuk PAI

Temuan Utama	Indikator Temuan	Keterangan Temuan
1. Implementasi Model Evaluasi Sistemik (CIPP)	Context, Input, Process, dan Product.	Model ini mengevaluasi PAI secara menyeluruh, mulai dari kesesuaian kurikulum dengan lingkungan sekolah hingga kualitas interaksi guru-siswa dan hasil akhir belajar.
2. Pengukuran Dampak Perubahan Perilaku (Kirkpatrick)	Reaction, Learning, Behavior, dan Result.	Menekankan pada perubahan sikap dan akhlak siswa di kehidupan nyata (Behavior), yang merupakan inti dari keberhasilan pendidikan karakter dalam PAI.
3. Relevansi Evaluasi Autentik & Goal Attainment	Portofolio, Proyek, Observasi, dan keselarasan tujuan Tyler.	Menggunakan konteks nyata untuk menilai kemampuan praktik ibadah dan memastikan hasil belajar siswa sejalan dengan standar kompetensi yang ditetapkan.

Evaluasi sebagai Instrumen Pengembangan Kurikulum

Temuan ini menempatkan evaluasi pada posisi yang lebih tinggi dari sekadar alat pengukuran kelas. Evaluasi pendidikan Islam sangat erat kaitannya dengan apa yang telah difirmankan dalam Al-Qur'an, dimana evaluasi dalam pendidikan Islam berkaitan dengan keilmuan dan perilaku keagamaan (Yusuf & Nata,2023). Evaluasi merupakan bagian dari sistem manajemen, yaitu perencanaan, pengorganisasian, pelaksanaan, pemantauan, dan evaluasi (Rusmani, & Arifmiboy, 2023). Evaluasi dipahami sebagai instrumen strategis dan vital yang hasil-hasilnya digunakan sebagai dasar empiris untuk perbaikan, penyesuaian, dan pengembangan kurikulum PAI secara berkelanjutan. Proses evaluasi dilakukan untuk menilai kemajuan dari 3 ranah, kognitif, afektif, dan psikomotorik. (Ikbal et al., 2025). Dengan kata lain, evaluasi adalah jembatan yang menghubungkan hasil belajar di tingkat mikro (kelas) dengan keputusan kebijakan di tingkat makro (program sekolah atau sistem pendidikan). Melalui evaluasi formatif dan sumatif yang tepat, siswa dapat mengetahui pencapaian diri dan berusaha meningkatkan kemampuan. Fungsi dari penilaian formatif dan sumatif yaitu untuk memantau, mengidentifikasi, menganalisis, dan memperbaiki proses belajar, serta menilai keberhasilan siswa dalam mencapai tujuan pembelajaran yang diharapkan. (Putri et al., 2023).

Fungsi utama hasil evaluasi adalah menyediakan data dan informasi yang diperlukan untuk pengambilan keputusan yang rasional atau *data-driven decision making*. Ketika evaluasi dilakukan secara komprehensif, ia tidak hanya menghasilkan nilai akhir, tetapi juga data rinci mengenai efektivitas *Input* dan kualitas *Proses* pembelajaran. Misalnya, jika hasil evaluasi

menunjukkan pencapaian afektif yang rendah, data ini menjadi bukti empiris yang mendorong revisi Kurikulum dan materi ajar atau peningkatan Pelatihan Guru. Hasil evaluasi yang sistematis dan terpercaya menghilangkan dugaan-dugaan dalam pengambilan keputusan, menggantikannya dengan bukti nyata. Evaluasi yang tepat dapat membantu mengidentifikasi kekuatan dan kelemahan kurikulum serta memberikan rekomendasi untuk perbaikan kurikulum berikutnya. (Maemunah, 2025). Menurut Stufflebeam (2007), evaluasi adalah proses menyediakan informasi yang diperlukan untuk pengambilan keputusan.

Selain itu, evaluasi bertindak sebagai mekanisme Penjaminan Mutu Berkelanjutan bagi program PAI. Evaluasi membantu program PAI untuk selalu relevan dan efektif dengan mengidentifikasi kelemahan sistemik. Evaluasi pembelajaran berfungsi sebagai alat yang penting untuk refleksi bagi pendidik, siswa, dan pembuat kebijakan dalam meningkatkan praktik pendidikan yang lebih tepat, inklusif, dan bermanfaat. (Arbeni et al., 2024). Ketika data menunjukkan bahwa sebagian besar siswa kesulitan dalam ranah psikomotorik, ini mengindikasikan kelemahan pada cara penyampaian atau panduan praktik, bukan sekadar kesalahan individu siswa. Selain itu, evaluasi menilai sejauh mana kurikulum PAI saat ini relevan dengan tantangan kontemporer, misalnya isu moderasi beragama. Jika hasil evaluasi menunjukkan bahwa siswa kurang memiliki literasi digital yang Islami, maka kurikulum PAI perlu disesuaikan untuk memasukkan topik tersebut.

Dalam konteks akuntabilitas program, hasil evaluasi menunjukkan kepada pemangku kepentingan, seperti orang tua dan kementerian, bahwa program PAI telah dilaksanakan dan diukur dampaknya secara profesional. Hasil dari proses penilaian ini selanjutnya dievaluasi untuk menentukan tingkat pencapaian pribadi atau terprogram (Arfah, 2021). Laporan evaluasi yang transparan menunjukkan sejauh mana dana, waktu, dan sumber daya telah berhasil diubah menjadi capaian belajar yang sesuai dengan tujuan agama dan negara. Secara keseluruhan, temuan ini menegaskan bahwa evaluasi tidak boleh dilihat sebagai kegiatan akhir yang terpisah, melainkan sebagai bagian integral dari siklus pengembangan kurikulum. Data yang dihasilkannya adalah aset berharga yang memastikan PAI terus beradaptasi, berevolusi, dan mencapai tujuan utamanya secara optimal dan terukur.

Tabel 3. Evaluasi sebagai Instrumen Pengembangan Kurikulum

Temuan Utama	Indikator Temuan	Keterangan Temuan
Evaluasi sebagai Dasar Pengambilan Keputusan	Hasil evaluasi digunakan untuk merevisi kurikulum, materi ajar, dan metode pengajaran.	Evaluasi menyediakan data empiris yang rasional (data-driven) untuk memutuskan perbaikan pada input (kurikulum) dan proses (metode) PAI, menghilangkan keputusan yang didasarkan pada asumsi.
Mekanisme Penjaminan Mutu PAI	Identifikasi kelemahan sistemik dan penilaian relevansi kurikulum.	Evaluasi bertindak sebagai alat quality assurance yang memastikan program PAI tetap efektif, relevan dengan tantangan zaman, dan sesuai dengan tujuan pendidikan Islam.
Dukungan Akuntabilitas Program	Laporan evaluasi yang transparan kepada pemangku kepentingan.	Hasil evaluasi menunjukkan sejauh mana sumber daya telah diubah menjadi capaian belajar yang berhasil, mendukung akuntabilitas sekolah terhadap kualitas pendidikan agama.

Temuan ini menempatkan evaluasi pada peran strategis, yaitu sebagai instrumen vital dan berkelanjutan dalam pengembangan dan peningkatan mutu Kurikulum PAI. Evaluasi tidak hanya berhenti pada penilaian hasil siswa, melainkan berfungsi menyediakan data empiris yang kredibel untuk pengambilan keputusan yang rasional (*data-driven decision making*). Informasi yang dihasilkan dari evaluasi, terutama yang bersifat komprehensif, menjadi dasar utama untuk merevisi dan memperbaiki elemen *input* (seperti kurikulum dan materi ajar) serta *proses* (metode pengajaran guru). Selain itu, evaluasi bertindak sebagai mekanisme penjaminan mutu (*quality assurance*) yang memastikan program PAI selalu relevan dengan tantangan kontemporer dan efektif dalam mencapai tujuan pembentukan karakter. Dengan menyajikan laporan yang transparan, evaluasi juga mendukung akuntabilitas program kepada pemangku kepentingan. Secara ringkas, evaluasi adalah bagian integral dari siklus pengembangan kurikulum, menjamin PAI terus berevolusi dan optimal dalam mencapai tujuan pendidikan Islam.

PEMBAHASAN

Pembahasan ini berfungsi mengintegrasikan dan menganalisis secara mendalam ketiga temuan utama yang dihasilkan dari kajian mengenai peran evaluasi dalam Pendidikan Agama Islam (PAI): yaitu sifat holistik dan multidimensi, kebutuhan akan model komprehensif, serta fungsi evaluasi sebagai instrumen pengembangan kurikulum.

Temuan pertama menegaskan bahwa evaluasi PAI tidak dapat lagi dibatasi pada pengukuran kognitif semata. Implikasinya adalah PAI yang bertujuan pada transformasi karakter, harus memverifikasi keberhasilan program melalui pengukuran yang mencakup ranah afektif (sikap dan akhlak) dan psikomotorik (praktik ibadah) secara utuh dan berkelanjutan. Tuntutan holistik ini sejalan dengan filosofi Kurikulum Merdeka yang menekankan pada Projek Penguatan Profil Pelajar Pancasila (P5), yang secara inheren memerlukan penilaian perilaku dan etika. Tantangan utama dalam implementasi temuan ini adalah mengembangkan instrumen non-tes yang valid, seperti observasi terstruktur, jurnal refleksi, dan portofolio, yang mampu menangkap perubahan perilaku siswa dalam konteks nyata dan otentik. Evaluasi berguna untuk mediagnosis permasalahan atau tantangan belajar yang dihadapi oleh siswa, dan mencari cara untuk mengatasi permasalahan yang mereka hadapi. (Hasibuan, et al., 2024). Proses ini memungkinkan guru memberikan remedial atau pengayaan. Hal ini sejalan dengan prinsip assessment for learning dalam Kurikulum Merdeka (Kemendikbud, 2022).

Kebutuhan akan Model Evaluasi Komprehensif muncul sebagai respons langsung terhadap tuntutan holistik tersebut. Evaluasi PAI harus bersifat sistematis, dan Model CIPP menyediakan kerangka kerja yang ideal untuk menilai efektivitas program secara menyeluruh—mulai dari kesiapan sumber daya (*Input*), kualitas pelaksanaan (*Process*), hingga dampak akhir (*Product*). Penerapan CIPP memastikan bahwa perbaikan tidak hanya menyentuh hasil belajar siswa, tetapi juga akar penyebab masalah yang mungkin terletak pada kualitas kurikulum atau kompetensi guru. Lebih lanjut, sinergi dengan Model Kirkpatrick adalah kunci, khususnya pada level *Behavior*. Model Kirkpatrick" telah menjadi pendekatan yang paling banyak digunakan untuk evaluasi pelatihan di dunia korporasi, pemerintahan, dan akademis (Kirkpatrick, D., & Kirkpatrick, J. (2006). Dalam PAI, indikator keberhasilan tertinggi bukanlah nilai ujian, melainkan perubahan nyata dalam perilaku dan praktik keagamaan siswa. Dengan memfokuskan evaluasi hingga level perilaku, data yang dihasilkan menjadi sangat kredibel, membuktikan bahwa PAI benar-benar memberikan dampak nyata, yang diwujudkan melalui adopsi Evaluasi Autentik. Evaluasi autentik juga mendorong peserta didik untuk menunjukkan kemampuan mereka dalam konteks kehidupan nyata, misalnya melalui pembiasaan ibadah atau penerapan akhlak di sekolah. Karena mencerminkan kompetensi secara menyeluruh, evaluasi ini dianggap paling sesuai untuk menggambarkan perkembangan keagamaan peserta didik secara komprehensif (Kemendikbud, 2022). Jika hasil evaluasi menunjukkan adanya ketidaksesuaian antara tujuan pembelajaran dan

pencapaian siswa, maka guru memiliki dasar yang kuat untuk melakukan perbaikan terhadap rencana pembelajaran berikutnya, baik dari aspek pemilihan metode, penggunaan media, penyusunan materi, maupun strategi penilaian yang digunakan (Sudijono, 2011).

Temuan terakhir, yang menempatkan evaluasi sebagai Instrumen Pengembangan Kurikulum, mengintegrasikan hasil dari kedua temuan sebelumnya ke tingkat kebijakan. Pembahasan ini menekankan bahwa hasil evaluasi holistik yang diperoleh melalui model komprehensif harus diartikan sebagai basis data empiris untuk penentuan arah kebijakan. Model komprehensif Pendidikan Agama Islam (PAI) adalah pendekatan pembelajaran yang menyeluruh, tidak hanya fokus pada aspek normatif (ibadah) tetapi juga mencakup pemahaman Islam secara luas. Model penanaman nilai-nilai karakter kebangsaan di madrasah membutuhkan pendekatan yang komprehensif integrative supaya dapat berjalan efektif dan berhasil baik (Hasanah, 2022). Evaluasi bertindak sebagai mekanisme *quality assurance* bagi program PAI. Jika evaluasi menunjukkan kelemahan sistemik, misalnya, rendahnya praktik moderasi beragama, data ini secara otomatis memicu revisi kurikulum untuk memperkuat topik tersebut. Hal ini menciptakan siklus perbaikan berkelanjutan, di mana evaluasi mengidentifikasi kelemahan program, kurikulum direvisi, dan program dilaksanakan kembali untuk dievaluasi. Dengan demikian, evaluasi memastikan program PAI tidak stagnan, tetapi terus beradaptasi dengan tantangan kontemporer dan mencapai akuntabilitas publik yang diperlukan. Secara keseluruhan, pembahasan menyimpulkan bahwa evaluasi PAI harus bertransformasi menjadi fungsi manajemen strategis, memastikan bahwa PAI efektif dalam membentuk karakter siswa yang berakhlak mulia.

KESIMPULAN

Penelitian ini menjawab tujuan untuk mengkaji urgensi transformasi evaluasi dalam Pendidikan Agama Islam (PAI) dengan menunjukkan bahwa evaluasi tidak lagi dapat diposisikan semata sebagai alat penilaian hasil akhir (*assessment of learning*), melainkan harus berfungsi sebagai instrumen strategis yang multidimensi. Temuan kajian menegaskan bahwa keberhasilan PAI tidak cukup diukur melalui penguasaan kognitif, tetapi harus memverifikasi internalisasi nilai pada ranah afektif (akhlak) dan kemahiran pada ranah psikomotorik (ibadah) melalui asesmen autentik. Implementasi model evaluasi yang komprehensif menjadi kunci, di mana sinergi Model CIPP yang memantau aspek sistemik (konteks hingga produk), Model Kirkpatrick yang menekankan perubahan perilaku nyata (*behavior*), serta Model *Goal*

Attainment Tyler yang menjaga konsistensi tujuan, menyediakan kerangka kerja yang utuh bagi pendidik dalam merancang dan melaksanakan evaluasi PAI. Secara strategis, hasil evaluasi berfungsi sebagai basis *data-driven* bagi pengembangan kurikulum dan penjaminan mutu berkelanjutan, sehingga memastikan PAI tetap relevan dengan tantangan zaman dan efektif dalam membentuk peserta didik yang beriman, bertakwa, dan berakhlak mulia sesuai amanat Kurikulum Merdeka.

Studi ini berkontribusi secara teoretis dengan memperkaya pemahaman tentang evaluasi PAI sebagai proses holistik yang mengintegrasikan dimensi kognitif, afektif, dan psikomotorik dalam satu kerangka evaluasi yang menyeluruh, melalui pemanfaatan terpadu Model CIPP, Model Kirkpatrick, dan Model *Goal Attainment* Tyler. Secara metodologis, penelitian ini menegaskan pentingnya pemanfaatan hasil evaluasi sebagai dasar empiris bagi perancangan dan penyempurnaan kurikulum, sekaligus sebagai instrumen penjaminan mutu yang berkelanjutan. Secara praktis, temuan ini memberikan rujukan konseptual bagi pendidik, pengelola satuan pendidikan, dan perancang kebijakan untuk mengembangkan sistem evaluasi PAI yang lebih akurat, autentik, dan selaras dengan profil kompetensi lulusan yang diamanatkan *Kurikulum Merdeka*.

Berdasarkan cakupan analisis yang disajikan dalam penelitian ini, kajian selanjutnya disarankan untuk mengembangkan dan menguji secara empiris model evaluasi PAI yang holistik pada berbagai jenjang dan konteks satuan pendidikan, guna menilai efektivitas penerapan kombinasi Model CIPP, Model Kirkpatrick, dan Model *Goal Attainment* Tyler dalam praktik. Penelitian lanjutan juga perlu diarahkan pada pengembangan instrumen asesmen autentik yang lebih terukur untuk memotret ranah afektif dan psikomotorik, serta pada eksplorasi lebih jauh pemanfaatan hasil evaluasi sebagai dasar pengambilan keputusan dalam pengembangan kurikulum dan sistem penjaminan mutu PAI yang berkelanjutan.

DAFTAR PUSTAKA

Ansya, Y. A. U., Alfianita, A., Syahkira, H. P., & Syahrial, S. (2024). Peran Evaluasi Pembelajaran pada Mata Pelajaran Matematika Kelas V Sekolah Dasar. *Indiktika: Jurnal Inovasi Pendidikan Matematika, 6*(2), 173–184.

Arbeni, W., Indrianti, N., Fahlevi, F., Septiawan, D., Wahyu, M., Sari, R. P., & Nasution, M. R. (2024). Analisis Hasil Evaluasi dan Perkembangan. *Mudabbir Journal Research and Education Studies, 4*(2), 376–387.

Arfah, M. A. (2021). Evaluasi Hasil Belajar Pendidikan Agama Islam (PAI). *Jurnal Literasiologi, 7*(2), 211–236.

Ariefky, M. M., & Inayati, N. L. (2023). Peran Guru PAI dalam Pelaksanaan Evaluasi Pembelajaran Sumatif Siswa di SMK Negeri 6 Sukoharjo. *Edukasia Jurnal Pendidikan dan Pembelajaran, 4*(2), 2343–2350.

Bellia, F. (2022). Evaluasi Studi Mata Kuliah Metodologi Penelitian (Implementasi Evaluasi Program Model Berorientasi Tujuan Ralph W. Tyler). *Jurnal Ilmiah Mandala Pendidikan*.

Betu, F. S. (2023). Objectives-Oriented Evaluation: The Tylerian Tradition sebagai Tawaran Evaluasi terhadap Satuan Pendidikan Sekolah Dasar. *Sapa: Jurnal Kateketik dan Pastoral, 8*(2), 147–156.

Fadli, M. R., Hizbulloh, P., & Loji, N. S. (2024). Tantangan dalam Evaluasi Pendidikan Agama Islam. *Jurnal Intelek Insan Cendikia, 1*(10), 7461–7468.

Faelasup, F., & Astuti, A. (2025). Evaluasi Hasil Belajar Pendidikan Agama Islam (PAI) melalui *Library Research*. *Jurnal Riset Rumpun Agama dan Filsafat, 4*(1), 621–635.

Hasanah, A. (2022). Penguatan Karakter Kebangsaan melalui Pendekatan Integratif pada Mapel Rumpun PAI di Madrasah. *Edukasi Islami: Jurnal Pendidikan Islam, 11*(1).

Hasibuan, S. M. (2024). Evaluasi Diagnostik dan Remedial dalam Pembelajaran. *TADRIBUNA: Journal of Islamic Management Education, 4*(2), 62–71.

Hidayati, R. E., & Adien, M. N. (2025). Evaluasi Workshop Peningkatan Kemampuan Penyusunan Perencanaan Pembelajaran Bermuatan Literasi Keagamaan Lintas Budaya (LKLB) Menggunakan Model Kirkpatrick pada Guru-Guru di Kota Malang. *Wawasan: Jurnal Kediklatan Balai Diklat Keagamaan Jakarta, 6*(1), 39–54.

Ibeh, A. I. (2022). Teori Kurikulum oleh Ralph Tyler dan Implikasinya terhadap Pembelajaran Abad Ke-21. *Jurnal Penelitian dan Studi Kebijakan Pendidikan Unizik*, 52–61.

Ikbal, A., Andrianto, A., Suryaningrum, E., Muniroh, B., Gusnita, F., & Julhadi, J. (2025). Pengukuran, Evaluasi dan Asesmen serta Tujuan dan Fungsi Evaluasi dalam Pembelajaran PAI. *Dharmas Education Journal (DE_Journal), 6*(1), 12–17.

Jamali, Y. (2018). Evaluasi Kompetensi Pedagogik Guru PAI pada SMA Negeri se-Kota Pangkalpinang (Penerapan Model Evaluasi CIPP). *Jurnal Ilmiah Islam Futura, 17*(2), 340–358.

Kemendikbud. (2022). *Panduan Pembelajaran dan Asesmen Kurikulum Merdeka*. Kemendikbud Ristek.

Khalisah, L. (2025). *Makalah Evaluasi Model Kirkpatrick*.

Kirkpatrick, D., & Kirkpatrick, J. (2006). *Mengevaluasi Program Pelatihan: Empat Tingkatan*. Berrett-Koehler.

Maimunah, M. (2024). Inovasi Evaluasi Pembelajaran dalam Kurikulum PAI. *Al-Mikraj Jurnal Studi Islam dan Humaniora, 4*(2), 1374–1387.

Majid, A. (2014). *Penilaian Autentik dalam Pembelajaran*. PT Remaja Rosdakarya.

Malili, A., Setiawati, Y. H., & Primarnie, A. (2023). Implementasi Pendidikan Holistik Islami pada Pembelajaran Pendidikan Agama Islam di Sekolah Dasar Muhammadiyah Bojong Gede Bogor. *Jurnal Dirosah Islamiyah, 5*(1), 95–121.

Munandar, A., Nurholizah, S., Artika, D. T., Mahroja, S., Nurholizah, R., Anggraini, M., … Adetya, A. F. (2023). Evaluasi Program Pendidikan: Tinjauan terhadap Efektivitas dan Tantangan. *Jurnal Manajemen Pendidikan Islam, 9*(2), 128–136.

Ningsih, W. (2025). Model Pembelajaran PAI yang Relevan dengan Kehidupan Sehari-Hari Siswa. *Komprehensif, 3*(1), 66–73.

Nurdiyanto, N., Basri, H. B. H., & Suhartini, A. S. A. (2024). Internalisasi Nilai Religius pada Mata Pelajaran PAI Jenjang SD untuk Mengembangkan Sikap Keberagamaan Siswa. *Ta Dib Jurnal Pendidikan Islam, 13*(1), 97–112.

Purwati, H., Aini, A. N., Wulandari, D., Handayanto, A., & Nugrahani, D. (2025). Evaluasi Penyelenggaraan Pelatihan Pemanfaatan AI dalam Perencanaan Pembelajaran di Era Revolusi Industri 4.0 bagi Guru SMA. *Jurnal Pelatihan Pendidikan, 4*(1), 15–26.

Putri, F., & Zakir, S. (2023). Mengukur Keberhasilan Evaluasi Pembelajaran: Telaah Evaluasi Formatif dan Sumatif dalam Kurikulum Merdeka. *Dewantara: Jurnal Pendidikan Sosial Humaniora, 2*(4), 172–180.

Ridho, A., Kusaeri, K., Nasaruddin, N., & Rohman, F. (2020). Evaluasi Program Gerakan Furudhul Ainiyah (GEFA) dengan Menggunakan Model Kirkpatrick. *Fikrotuna: Jurnal Pendidikan dan Manajemen Islam, 11*(1), 1480–1495.

Rusman, M. P. I. (2018). Efektifitas Evaluasi Model CIPP (Context, Input, Process, Product) terhadap Hasil Belajar Siswa pada Bidang Studi Pendidikan Agama Islam di SMP Jati Agung Wage Sidoarjo.

Rusmani, M. A., & Arifmiboy, A. (2023). Evaluasi Kurikulum. *Anthor: Education and Learning Journal, 2*(3), 410–415.

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

Sudijono, A. (2011). *Pengantar Evaluasi Pendidikan*. RajaGrafindo.

Sugiyono. (2020). *Metode Penelitian Pendidikan*. Alfabeta.

Utami, A. P., Balqis, Z., Apriani, L., Rahmadani, I., & Amanda, R. D. (2025).

Yusuf, E., & Nata, A. (2023). Evaluasi dan Pengembangan Kurikulum Pendidikan Islam. *Edukasi Islami: Jurnal Pendidikan Islam, 12*(1).