



MODEL-MODEL EVALUASI PROGRAM DALAM PENDIDIKAN AGAMA ISLAM: KAJIAN KOMPARATIF MODEL CIPP DAN MODEL STAKE

Muhammad Hafi¹ & Dina Hermina²

Program Studi Pendidikan Agama Islam, Pascasarjana Universitas Islam Negeri Antasari
Banjarماسin^{1,2}

elfatih0777@gmail.com¹

ARTICLE INFO

History

Published: 30 Sep 2026

Keywords:

program evaluation, CIPP
Model, Stake Model, Islamic
Religious Education



© Author(s) 2026

This work is licensed under
a [Creative Commons Attribution
4.0 International License](https://creativecommons.org/licenses/by/4.0/).



ABSTRACT

Indonesia:

Penelitian ini mengkaji model-model evaluasi program dalam konteks Pendidikan Agama Islam (PAI), dengan fokus pada perbandingan Model CIPP (Context, Input, Process, Product) yang dikembangkan oleh Daniel L. Stufflebeam dan Model Stake (Countenance Model) yang dikembangkan oleh Robert E. Stake. Evaluasi dalam PAI menuntut pendekatan yang holistik karena mencakup aspek kognitif, afektif, dan psikomotorik secara terpadu. Metode yang digunakan dalam penelitian ini adalah library research (kajian kepustakaan) dengan pendekatan deskriptif-analitis dan komparatif. Hasil kajian menunjukkan bahwa Model CIPP unggul dalam aspek manajerial dan sistemik, sehingga efektif untuk mengevaluasi kurikulum dan sistem pembelajaran PAI secara menyeluruh, sedangkan Model Stake lebih efektif dalam memahami pengalaman religius dan dinamika pembentukan karakter peserta didik melalui pendekatan responsif dan naturalistik. Kajian ini menyimpulkan bahwa integrasi kedua model evaluasi tersebut dapat menghasilkan sistem evaluasi PAI yang komprehensif, objektif, dan sekaligus berlandaskan nilai-nilai Islam, guna mendukung terwujudnya tujuan pendidikan Islam: pembentukan insan yang beriman, bertakwa, dan berakhlak mulia.

English:

This study examines program evaluation models in the context of Islamic Religious Education (PAI), focusing on a comparative analysis of the CIPP Model (Context, Input, Process, Product) developed by Daniel L. Stufflebeam and the Stake Model (Countenance Model) developed by Robert E. Stake. Evaluation in PAI requires a holistic approach encompassing cognitive, affective, and psychomotor aspects in an integrated manner. This research employs a library research method with a descriptive-analytical and comparative approach. The findings indicate that the CIPP Model excels in managerial and systemic aspects, making it effective for comprehensively evaluating PAI curricula and learning systems, while the Stake Model is more effective in understanding students' religious experiences and the dynamics of character formation through a responsive and naturalistic approach. This study concludes that integrating both evaluation models can produce a comprehensive, objective, and values-based PAI evaluation system, supporting the realization of Islamic education's ultimate goal: forming individuals who are faithful, pious, and of noble character.

PENDAHULUAN

Pendidikan Agama Islam (PAI) memiliki peran strategis dalam membentuk karakter, moral, dan spiritual peserta didik. Secara normatif, tujuan pendidikan nasional sebagaimana tertuang dalam Undang-Undang Nomor 20 Tahun 2003 adalah mengembangkan potensi peserta didik agar menjadi manusia yang beriman, bertakwa kepada Tuhan Yang Maha Esa, dan berakhlak mulia. Dalam konteks ini, PAI tidak hanya berfungsi sebagai proses *transfer of knowledge*, tetapi juga sebagai proses internalisasi nilai (*value internalization*) yang membentuk kepribadian peserta didik secara utuh.

Keberhasilan pembelajaran PAI tidak cukup diukur dari aspek kognitif saja, melainkan juga harus mencakup aspek afektif dan psikomotorik sebagaimana klasifikasi tujuan pendidikan yang dikemukakan oleh Benjamin S. Bloom. Namun dalam praktiknya, evaluasi pembelajaran PAI di sekolah sering kali masih berorientasi pada hasil belajar (output) dalam bentuk tes tertulis, sehingga lebih menekankan pada ranah kognitif dan kurang menyentuh dimensi sikap serta praktik keagamaan peserta didik.

Dalam perkembangan teori evaluasi program pendidikan, terdapat berbagai model evaluasi yang dikembangkan oleh para ahli. Dua model yang paling berpengaruh adalah Model CIPP (Context, Input, Process, Product) yang dikembangkan oleh Daniel L. Stufflebeam dan Model Stake (Countenance Model) yang dikembangkan oleh Robert E. Stake. Model CIPP dirancang sebagai pendekatan evaluasi yang berorientasi pada pengambilan keputusan (*decision-oriented evaluation*), sedangkan Model Stake menekankan pendekatan responsif terhadap kebutuhan pemangku kepentingan. Kajian komparatif terhadap kedua model ini menjadi penting dalam rangka mengembangkan sistem evaluasi pembelajaran PAI yang lebih komprehensif dan kontekstual.

Berdasarkan latar belakang tersebut, artikel ini merumuskan beberapa permasalahan: (1) bagaimana konsep evaluasi program dalam PAI ditinjau dari teori evaluasi pendidikan?; (2) bagaimana karakteristik Model CIPP dalam evaluasi pembelajaran PAI?; (3) bagaimana karakteristik Model Stake dalam konteks pembelajaran PAI?; dan (4) bagaimana perbandingan antara Model CIPP dan Model Stake serta implikasinya terhadap peningkatan mutu pembelajaran PAI?

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kualitatif dengan metode kajian kepustakaan (*library research*). Data dikumpulkan melalui telaah terhadap sumber-sumber primer dan sekunder yang relevan, meliputi buku teks, artikel jurnal ilmiah, dan dokumen akademik yang berhubungan dengan teori evaluasi pendidikan, Model CIPP, Model Stake, dan evaluasi dalam Pendidikan Agama Islam. Analisis dilakukan secara deskriptif-analitis dan komparatif untuk memperoleh pemahaman yang mendalam mengenai karakteristik, kelebihan, dan kelemahan masing-masing model evaluasi, serta relevansinya dalam konteks pembelajaran PAI.

HASIL DAN PEMBAHASAN

1. Konsep Evaluasi dalam Pendidikan Agama Islam

Evaluasi dalam pendidikan pada hakikatnya merupakan proses sistematis untuk mengumpulkan, menganalisis, dan menafsirkan informasi guna menentukan tingkat ketercapaian tujuan pembelajaran. Secara harfiah, evaluasi berasal dari bahasa Inggris *evaluation* yang berarti penilaian dan penaksiran. Dalam bahasa Arab dijumpai istilah *imtiḥān* yang berarti ujian, dan *ḵbataman* yang berarti cara menilai hasil akhir dari proses kegiatan.¹

Dalam perspektif pendidikan modern, Stufflebeam memandang evaluasi sebagai suatu proses menyediakan informasi yang berguna bagi pengambilan keputusan pendidikan, baik terkait peserta didik, program, maupun kebijakan pendidikan secara umum.² Pandangan ini menegaskan bahwa evaluasi memiliki fungsi strategis dalam menjamin mutu pembelajaran. Sejalan dengan itu, Black dan Wiliam menegaskan bahwa evaluasi memiliki peran penting dalam meningkatkan kualitas pembelajaran melalui pemberian umpan balik (*feedback*) yang berkesinambungan kepada peserta didik dan pendidik.³

Dalam konteks PAI, evaluasi memiliki karakteristik yang khas karena tidak hanya menilai aspek kognitif berupa penguasaan materi keislaman, tetapi juga mencakup aspek afektif yang berhubungan dengan sikap dan nilai religius, serta aspek psikomotorik yang tampak dalam praktik ibadah dan keterampilan keagamaan peserta didik. Evaluasi PAI menuntut pendekatan yang holistik agar proses penilaian mampu merefleksikan keterpaduan antara pengetahuan, sikap, dan pengamalan ajaran Islam dalam kehidupan sehari-hari.⁴

Dalam perspektif Al-Qur'an dan hadis, proses evaluasi pembelajaran harus mencerminkan nilai-nilai *ḥisab* (pertanggungjawaban amal) dan *mubasabah* (evaluasi diri), sehingga aspek evaluasi yang bersifat spiritual dan etis menjadi bagian tak terpisahkan dari sistem penilaian pendidikan Islam.⁵ Dengan demikian, evaluasi PAI merupakan proses yang sistematis dan berkesinambungan untuk menilai ketercapaian tujuan pembelajaran yang mencakup aspek kognitif, afektif, dan psikomotorik, sekaligus berlandaskan nilai-nilai ajaran Islam.

2. Fungsi dan Tujuan Evaluasi Pendidikan Agama Islam

Evaluasi dalam pendidikan memiliki fungsi utama sebagai alat untuk mengetahui tingkat ketercapaian tujuan pembelajaran serta sebagai dasar pengambilan keputusan pendidikan. Melalui Model CIPP, evaluasi tidak hanya berfungsi menilai hasil akhir, tetapi juga menilai konteks kebutuhan, kesiapan sumber daya, proses pelaksanaan, serta capaian program sebagai dasar penetapan kebijakan pendidikan.⁶

¹ R. Hidayat, *Ilmu Pendidikan Islam: Menuntun arab Pendidikan Islam Indonesia* (LPPPI, 2016).

² D. L. Stufflebeam, "The CIPP Model for Evaluation," in *International Handbook of Educational Evaluation*, eds. T. Kellaghan and D. L. Stufflebeam (Kluwer Academic Publishers, 2003).

³ P. Black and D. Wiliam, "Assessment and Classroom Learning," *Assessment in Education: Principles, Policy & Practice* 5, no. 1 (1998).

⁴ M. Fikri, F. Firdayeni, dan J. Jamilus, "Transformasi evaluasi Pendidikan Islam dari penilaian kognitif ke pembentukan karakter peserta didik," *YASIN* 5, no. 6 (2025), <https://doi.org/10.58578/yasin.v5i6.8504>.

⁵ O. R. P. Pitriani dkk., "Evaluasi pendidikan dalam perspektif Al-Qur'an dan Hadist," *Pendas: Jurnal Ilmiah Pendidikan Dasar* 10, no. 4 (2025).

⁶ D. L. Stufflebeam and A. J. Shinkfield, *Evaluation Theory, Models, and Applications* (Jossey-Bass, 2007).

Dalam konteks Pendidikan Islam, fungsi evaluasi mencakup tiga dimensi: (1) fungsi diagnostik, yakni membantu guru mengidentifikasi kelemahan dan kebutuhan belajar peserta didik; (2) fungsi formatif, yakni memperbaiki proses pembelajaran agar lebih efektif; dan (3) fungsi sumatif, yakni menilai keberhasilan program secara keseluruhan.⁷ Selain itu, tujuan evaluasi PAI mencakup dua dimensi utama: dimensi akademik untuk mengukur pemahaman peserta didik terhadap materi ajaran Islam, dan dimensi moral-spiritual untuk menilai internalisasi nilai serta pengamalan ajaran tersebut dalam kehidupan sehari-hari.

3. Prinsip-Prinsip Evaluasi Pendidikan Islam

Dalam perspektif teori evaluasi Barat, prinsip utama evaluasi menekankan pada validitas, reliabilitas, objektivitas, dan kebermanfaatan (*utility*). Menurut Tyler, evaluasi harus didasarkan pada kejelasan tujuan pendidikan.⁸ Sementara itu, Stufflebeam menegaskan bahwa evaluasi harus memenuhi standar kegunaan (*utility*), kelayakan (*feasibility*), kepatutan (*propriety*), dan ketepatan (*accuracy*).⁹

Dalam perspektif pendidikan Islam, prinsip evaluasi tidak hanya bertumpu pada kaidah ilmiah, tetapi juga berlandaskan nilai-nilai syariat. Evaluasi harus dilaksanakan secara adil (*al-'adl*), objektif, menyeluruh (*syumul*), dan berkesinambungan (*istimrar*), karena tujuan pendidikan Islam bukan hanya transfer pengetahuan, melainkan pembinaan akhlak dan spiritualitas peserta didik.¹⁰ Evaluasi PAI yang efektif harus mengintegrasikan standar ilmiah modern dengan nilai-nilai normatif ajaran Islam.¹¹

4. Model CIPP (Context, Input, Process, Product)

a. Landasan Filosofis Model CIPP

Model CIPP dikembangkan oleh Daniel L. Stufflebeam pada tahun 1966 sebagai respons terhadap keterbatasan model evaluasi Tyler yang hanya berfokus pada pengukuran hasil belajar. Model ini berlandaskan teori sistem dan pengambilan keputusan, serta dipengaruhi oleh paradigma evaluasi yang menempatkan evaluasi sebagai bagian integral dari manajemen mutu pendidikan.¹² Dalam paradigma ini, evaluasi dipahami sebagai proses reflektif yang mendukung peningkatan kualitas berkelanjutan (*continuous quality improvement*).

b. Komponen Model CIPP

Evaluasi Konteks (Context Evaluation) bertujuan mengidentifikasi dan menilai kebutuhan, tujuan, serta masalah yang menjadi dasar perencanaan program. Evaluasi ini tidak hanya mengidentifikasi kebutuhan, tetapi juga menganalisis kesenjangan antara kondisi aktual dan kondisi ideal (*gap analysis*).¹³ Dalam pembelajaran PAI, evaluasi konteks dapat mencakup

⁷ E. F. Azmi, I. Hidayat, dan A. Azhari, "Evaluation of Education in Islamic Education," *Jurnal Pendidikan Islam* 1, no. 3 (2024): <https://doi.org/10.70938/judikis.v1i3.49>.

⁸ R. W. Tyler, *Basic Principles of Curriculum and Instruction* (University of Chicago Press, 1949).

⁹ Stufflebeam and Shinkfield, *Evaluation Theory*.

¹⁰ A. Khoirunnisa dan Syamsudin, "Evaluasi pendidikan menurut perspektif filsafat Islam," *Jurnal Pendidikan Nusantara* 3, no. 2 (2024).

¹¹ D. H. Parahita dan H. Setiyatna, "Prinsip-prinsip dan teknik evaluasi dalam Pendidikan Agama Islam," *Media Penelitian Pendidikan* 19, no. 2 (2025).

¹² Stufflebeam and Shinkfield, *Evaluation Theory*.

¹³ J. L. Fitzpatrick, J. R. Sanders, and B. R. Worthen, *Program Evaluation: Alternative Approaches and Practical Guidelines*, 4th ed. (Pearson, 2011).

analisis karakteristik religius peserta didik, latar belakang sosial budaya, serta tuntutan kurikulum nasional.

Evaluasi Input (Input Evaluation) berfungsi untuk memilih strategi terbaik berdasarkan alternatif yang tersedia. Analisis ini melibatkan evaluasi kompetensi guru, ketersediaan sarana prasarana, kesiapan kurikulum, serta alokasi anggaran.¹⁴

Evaluasi Proses (Process Evaluation) bersifat formatif dan berfungsi sebagai mekanisme kontrol mutu selama program berlangsung.¹⁵ Dalam konteks PAI, evaluasi proses dapat dilakukan melalui observasi pembelajaran, supervisi akademik, serta refleksi guru terhadap metode internalisasi nilai.

Evaluasi Produk (Product Evaluation) tidak hanya menilai hasil langsung (*output*), tetapi juga dampak jangka panjang (*outcome* dan *impact*).¹⁶ Dalam PAI, evaluasi produk dapat mencakup perubahan sikap religius, pembiasaan ibadah, serta perilaku moral peserta didik. Stufflebeam menegaskan bahwa tujuan utama evaluasi bukan sekadar membuktikan (*to prove*), tetapi untuk memperbaiki (*to improve*).¹⁷

c. Kelebihan dan Kelemahan Model CIPP

Model CIPP unggul karena memberikan kerangka evaluasi yang menyeluruh dan terstruktur. Model ini adaptif terhadap berbagai konteks pendidikan, baik formal maupun nonformal, serta mampu mengintegrasikan evaluasi formatif dan sumatif secara simultan. Namun demikian, Model CIPP dapat menghasilkan data yang sangat luas sehingga memerlukan kapasitas analisis yang tinggi. Selain itu, jika tidak dirancang secara fokus, evaluasi bisa menjadi terlalu administratif dan kehilangan kedalaman reflektif.¹⁸

5. Model Stake (Countenance Model)

a. Landasan Paradigmatik Model Stake

Model Stake dikembangkan oleh Robert E. Stake pada tahun 1967 dalam tradisi evaluasi responsif (*responsive evaluation*) yang menekankan pentingnya pengalaman nyata para pemangku kepentingan. Berbeda dengan model yang sangat terstruktur, Model Stake lebih fleksibel dan kontekstual. Menurut Stake, evaluator harus memahami program secara naturalistik, yakni dengan mengamati fenomena sebagaimana adanya dalam konteks aslinya.¹⁹ Pendekatan ini memiliki kedekatan dengan paradigma penelitian kualitatif.²⁰

b. Komponen Description dan Judgment

Model Stake menekankan dua dimensi utama. Pertama, komponen deskripsi (*description*) yang menuntut evaluator menyajikan gambaran faktual tentang program tanpa intervensi nilai terlebih dahulu, dengan membandingkan antara rencana (*intents*) dan kenyataan (*observations*). Kedua, komponen pertimbangan (*judgment*) yang melibatkan pemberian nilai berdasarkan standar atau kriteria tertentu. Pendekatan ini memungkinkan

¹⁴ Stufflebeam and Shinkfield, *Evaluation Theory*.

¹⁵ M. Scriven, *Evaluation Thesaurus*, 4th ed. (SAGE Publications, 1991).

¹⁶ Stufflebeam and Shinkfield, *Evaluation Theory*.

¹⁷ Stufflebeam, "CIPP Model,"

¹⁸ Wirawan, *Evaluasi: Teori, model, standar, aplikasi, dan profesi* (Rajawali Pers, 2011).

¹⁹ R. E. Stake, "The Countenance of Educational Evaluation," *Teachers College Record* 68, no. 7 (1967).

²⁰ J. W. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 4th ed. (SAGE Publications, 2014).

evaluasi yang lebih humanistik dan mempertimbangkan dimensi sosial serta budaya dalam program pendidikan.²¹

c. Kelebihan dan Kelemahan Model Stake

Model Stake unggul dalam menangkap dinamika sosial dan pengalaman subjektif peserta program. Dalam PAI, pendekatan ini relevan untuk memahami proses internalisasi nilai yang tidak selalu terukur secara kuantitatif. Kelemahan Model Stake terletak pada potensi subjektivitas evaluator serta kurangnya struktur sistematis dibanding Model CIPP.²² Selain itu, model ini kurang menekankan analisis manajerial dalam pengambilan keputusan kebijakan.

6. Analisis Komparatif Model CIPP dan Model Stake dalam Pembelajaran PAI

Secara filosofis, Model CIPP berakar pada paradigma manajerial dan rasional,²³ sedangkan Model Stake berakar pada paradigma naturalistik dan responsif.²⁴ Perbedaan ini berimplikasi pada pendekatan metodologis yang digunakan: CIPP cenderung menggunakan pendekatan campuran (*mixed methods*), sementara Stake lebih dominan pada pendekatan kualitatif.²⁵

Dalam evaluasi pembelajaran PAI, Model CIPP lebih efektif untuk mengevaluasi kurikulum dan sistem pembelajaran secara menyeluruh, mulai dari analisis kebutuhan (context), kesiapan guru dan sarana (input), pelaksanaan pembelajaran (process), hingga capaian hasil belajar dan perubahan perilaku peserta didik (product). Sebaliknya, Model Stake lebih efektif untuk memahami pengalaman religius dan dinamika pembentukan karakter peserta didik melalui perbandingan antara perencanaan dan realitas di lapangan.

Secara keseluruhan, Model CIPP dan Model Stake tidak saling bertentangan, melainkan saling melengkapi. Kombinasi keduanya dapat menghasilkan evaluasi pembelajaran PAI yang komprehensif, sistematis, sekaligus kontekstual dan humanistik. Integrasi ini sejalan dengan pandangan bahwa evaluasi PAI yang ideal harus mampu mengukur penguasaan materi (kognitif), menilai sikap religius (afektif), sekaligus memahami proses penghayatan nilai-nilai Islam dalam kehidupan sehari-hari (psikomotorik).

KESIMPULAN

Evaluasi pembelajaran Pendidikan Agama Islam memiliki karakteristik yang lebih kompleks dibandingkan penilaian yang hanya berorientasi pada pencapaian kognitif. Evaluasi PAI perlu mencakup penguasaan pengetahuan keagamaan, perkembangan sikap dan nilai religius, serta kemampuan peserta didik mengaktualisasikan ajaran Islam dalam praktik kehidupan. Oleh karena itu, sistem evaluasi PAI memerlukan model yang mampu menilai proses dan hasil pembelajaran secara menyeluruh sekaligus mempertimbangkan dimensi moral dan spiritual.

Model CIPP menawarkan kerangka evaluasi yang sistematis melalui empat komponen, yaitu *context*, *input*, *process*, dan *product*. Evaluasi konteks membantu

²¹ R. E. Stake, *Standards-Based and Responsive Evaluation* (SAGE Publications, 2004).

²² Fitzpatrick, Sanders, and Worthen, *Program Evaluation*.

²³ Stufflebeam and Shinkfield, *Evaluation Theory*.

²⁴ Stake, *Standards-Based*.

²⁵ Creswell, *Research Design*.

mengidentifikasi kebutuhan dan tujuan program, evaluasi input menilai kesiapan sumber daya dan strategi, evaluasi proses memantau pelaksanaan pembelajaran, sedangkan evaluasi produk menilai capaian serta dampak program. Dalam pembelajaran PAI, kerangka ini memungkinkan evaluasi dilakukan sejak tahap perencanaan hingga hasil dan dampak pembelajaran sehingga mendukung pengambilan keputusan dan perbaikan mutu secara berkelanjutan.

Model Stake memberikan perspektif yang berbeda dengan menekankan hubungan antara rencana program dan realitas pelaksanaannya serta memperhatikan pengalaman para pemangku kepentingan. Pendekatan ini relevan dalam PAI karena proses internalisasi nilai, pengalaman religius, perubahan sikap, dan pembentukan karakter tidak selalu dapat dipahami secara memadai hanya melalui ukuran kuantitatif. Dengan orientasi yang lebih kontekstual dan humanistik, Model Stake memberikan ruang yang lebih luas untuk memahami bagaimana nilai-nilai keagamaan dialami dan dimaknai peserta didik.

Perbandingan kedua model menunjukkan bahwa CIPP dan Stake memiliki orientasi dan kekuatan yang berbeda tetapi tidak harus diposisikan secara berlawanan. CIPP lebih kuat pada aspek sistematisasi evaluasi, manajemen program, dan pengambilan keputusan, sedangkan Stake lebih sensitif terhadap konteks, pengalaman, serta dinamika yang berlangsung selama pembelajaran. Dalam konteks PAI, perbedaan tersebut justru membuka peluang untuk mengembangkan evaluasi yang menggabungkan struktur evaluatif yang komprehensif dengan pemahaman kontekstual terhadap proses pembentukan nilai dan karakter.

Secara keseluruhan, integrasi Model CIPP dan Model Stake berpotensi membentuk kerangka evaluasi PAI yang lebih komprehensif, sistematis, kontekstual, dan humanistik. CIPP dapat digunakan untuk memastikan kualitas program pada tingkat konteks, sumber daya, proses, dan hasil, sedangkan Stake dapat memperdalam pemahaman terhadap kesesuaian perencanaan dengan realitas serta pengalaman peserta didik. Karena kajian ini bersifat kepustakaan, penelitian selanjutnya dapat menguji kerangka integratif tersebut secara empiris pada sekolah atau madrasah untuk mengetahui efektivitas dan keterterapannya dalam evaluasi pembelajaran PAI.

Kebaruan Penelitian

Kebaruan penelitian ini terletak pada upaya menempatkan Model CIPP dan Model Stake sebagai dua kerangka evaluasi yang saling melengkapi dalam pembelajaran Pendidikan Agama Islam, bukan sekadar membandingkan kelebihan dan kelemahan masing-masing model. CIPP memberikan struktur evaluasi yang sistematis dan berorientasi pada pengambilan keputusan, sedangkan Stake memberikan kedalaman kontekstual melalui perhatian terhadap perencanaan, realitas pelaksanaan, pengalaman, dan kebutuhan pemangku kepentingan.

Kontribusi lainnya adalah pengembangan perspektif evaluasi PAI yang bergerak melampaui evaluasi hasil belajar kognitif. Kajian ini menempatkan evaluasi sebagai proses yang harus mampu menangkap aspek pengetahuan, sikap religius, praktik keagamaan, internalisasi nilai, serta pengalaman belajar peserta didik. Dengan demikian, evaluasi tidak

hanya berfungsi sebagai alat pengukuran hasil, tetapi juga sebagai instrumen refleksi dan perbaikan proses pendidikan Islam.

Kajian ini juga menawarkan dasar konseptual bagi pengembangan model evaluasi PAI integratif, yaitu penggunaan struktur *context–input–process–product* untuk mengevaluasi sistem pembelajaran secara menyeluruh yang diperkaya dengan perhatian terhadap kesesuaian antara perencanaan dan realitas serta pengalaman pemangku kepentingan. Pendekatan ini dapat menjadi alternatif bagi evaluasi PAI yang selama ini masih dominan berorientasi pada tes dan capaian akademik.

Implikasi Penelitian

Secara teoretis, temuan kajian ini mengimplikasikan bahwa evaluasi PAI perlu dikembangkan melalui pendekatan multidimensional. Evaluasi tidak cukup hanya menilai ketercapaian kompetensi kognitif, tetapi juga perlu mencakup dimensi afektif, psikomotorik, moral, dan spiritual. Integrasi CIPP dan Stake dapat memperluas kerangka evaluasi dari sekadar pengukuran hasil menuju pemahaman mengenai kualitas keseluruhan proses pembelajaran dan pengalaman religius peserta didik.

Secara pedagogis, guru PAI dapat menggunakan komponen CIPP untuk mengevaluasi kebutuhan peserta didik, kesiapan pembelajaran, pelaksanaan strategi, dan hasil belajar, kemudian memperkaya evaluasi tersebut dengan pendekatan Stake untuk memahami pengalaman peserta didik, kesesuaian antara rencana dan praktik, serta dinamika internalisasi nilai. Hal ini memungkinkan evaluasi digunakan sebagai dasar pemberian umpan balik dan perbaikan pembelajaran secara lebih berkelanjutan.

Secara institusional, sekolah dan madrasah dapat mengembangkan sistem evaluasi PAI yang menggabungkan data kuantitatif dan kualitatif, seperti hasil tes, observasi praktik ibadah, refleksi peserta didik, portofolio, wawancara, serta pengamatan perubahan perilaku. Pendekatan tersebut dapat memberikan informasi yang lebih kaya bagi guru, kepala sekolah, dan pemangku kebijakan dalam mengambil keputusan mengenai perbaikan kurikulum, strategi pembelajaran, kompetensi guru, dan program pembinaan karakter.

DAFTAR PUSTAKA

- Azmi, E. F., I. Hidayat, dan A. Azhari. "Evaluation of Education in Islamic Education." *Jurnal Pendidikan Islam* 1, no. 3 (2024): 149–158. <https://doi.org/10.70938/judikis.v1i3.49>.
- Black, P., and D. Wiliam. "Assessment and Classroom Learning." *Assessment in Education: Principles, Policy & Practice* 5, no. 1 (1998): 7–74.
- Creswell, J. W. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 4th ed. SAGE Publications, 2014.
- Fikri, M., F. Firdayeni, dan J. Jamilus. "Transformasi evaluasi Pendidikan Islam dari penilaian kognitif ke pembentukan karakter peserta didik." *YASIN* 5, no. 6 (2025): 7045–7057. <https://doi.org/10.58578/yasin.v5i6.8504>.
- Fitzpatrick, J. L., J. R. Sanders, and B. R. Worthen. *Program Evaluation: Alternative Approaches and Practical Guidelines*. 4th ed. Pearson, 2011.
- Hidayat, R. *Ilmu Pendidikan Islam: Menuntun arah Pendidikan Islam Indonesia*. LPPPI, 2016.
- Khoirunnisa, A., dan Syamsudin. "Evaluasi pendidikan menurut perspektif filsafat Islam." *Jurnal Pendidikan Nusantara* 3, no. 2 (2024): 105–115.
- Parahita, D. H., dan H. Setiyatna. "Prinsip-prinsip dan teknik evaluasi dalam Pendidikan Agama Islam." *Media Penelitian Pendidikan* 19, no. 2 (2025): 416–420.
- Pitriani, O. R. P., M. Agustina, A. Imron, dan M. Z. Arifin. "Evaluasi pendidikan dalam perspektif Al-Qur'an dan Hadist." *Pendas: Jurnal Ilmiah Pendidikan Dasar* 10, no. 4 (2025): 302–310.
- Scriven, M. *Evaluation Thesaurus*. 4th ed. SAGE Publications, 1991.
- Stake, R. E. "The Countenance of Educational Evaluation." *Teachers College Record* 68, no. 7 (1967): 523–540.
- Stake, R. E. *Standards-Based and Responsive Evaluation*. SAGE Publications, 2004.
- Stufflebeam, D. L. "The CIPP Model for Evaluation." In *International Handbook of Educational Evaluation*, edited by T. Kellaghan and D. L. Stufflebeam, 31–62. Kluwer Academic Publishers, 2003.
- Stufflebeam, D. L., and A. J. Shinkfield. *Evaluation Theory, Models, and Applications*. Jossey-Bass, 2007.
- Tyler, R. W. *Basic Principles of Curriculum and Instruction*. University of Chicago Press, 1949.
- Wirawan. *Evaluasi: Teori, model, standar, aplikasi, dan profesi*. Rajawali Pers, 2011.