

ANALISIS KOMPARATIF TEORI EVALUASI PENDIDIKAN: MODEL TYLER, SCRIVEN, STUFFLEBEAM (CIPP), DAN STAKE

Andrini Rosa Putri Siregar¹, Atikah Rahmah Nasution², Mariam Nasution³

^{1, 2, 3}UIN Syekh Ali Hasan Ahmad Addary Padangsidempuan

¹ andrinirosaputri@gmail.com, ² atikahrahmahnst1008@gmail.com,

³ mariam@uinsyahada.ac.id

ABSTRACT

Evaluation is a strategic component in the education system that serves as a basis for decision-making to improve the quality of learning and educational programs. Various evaluation models have been developed to address diverse evaluation needs according to the objectives, context, and characteristics of the program being assessed. This article aims to comprehensively analyze the Tyler, Scriven, Stufflebeam (CIPP), and Stake evaluation models and identify their strengths, limitations, and relevance in modern educational practice. The research uses a literature review method with a descriptive-analytical approach through a review of various relevant scientific sources. The results of the study indicate that the Tyler Model emphasizes the achievement of educational objectives as an indicator of program success. The Scriven Model develops formative and summative evaluations as well as goal-free evaluations that enable a broader identification of program impacts. The CIPP Model offers a comprehensive evaluation approach through the evaluation of context, input, process, and product, thus supporting sustainable decision-making. Meanwhile, the Stake Model emphasizes evaluation that is responsive to the needs of stakeholders and real-world conditions. The findings indicate that no evaluation model is completely superior to the others because each has a different focus, strengths, and limitations. Therefore, the selection of an evaluation model needs to be adjusted to the objectives, characteristics of the program, and the information needs to be obtained.

Keywords: Educational evaluation, Tyler model, Scriven model, CIPP model, Stake model.

ABSTRAK

Evaluasi merupakan komponen strategis dalam sistem pendidikan yang berfungsi sebagai dasar pengambilan keputusan untuk meningkatkan mutu pembelajaran dan program pendidikan. Berbagai model evaluasi telah dikembangkan untuk menjawab kebutuhan evaluasi yang beragam sesuai dengan tujuan, konteks, dan

karakteristik program yang dinilai. Artikel ini bertujuan menganalisis secara komprehensif model evaluasi Tyler, Scriven, Stufflebeam (CIPP), dan Stake serta mengidentifikasi kelebihan, keterbatasan, dan relevansinya dalam praktik pendidikan modern. Penelitian menggunakan metode studi pustaka dengan pendekatan deskriptif-analitis melalui penelaahan berbagai sumber ilmiah yang relevan. Hasil kajian menunjukkan bahwa Model Tyler menitikberatkan pada pencapaian tujuan pendidikan sebagai indikator keberhasilan program. Model Scriven mengembangkan evaluasi formatif dan sumatif serta evaluasi bebas tujuan yang memungkinkan identifikasi dampak program secara lebih luas. Model CIPP menawarkan pendekatan evaluasi yang komprehensif melalui evaluasi konteks, masukan, proses, dan produk sehingga mendukung pengambilan keputusan secara berkelanjutan. Sementara itu, Model Stake menekankan evaluasi yang responsif terhadap kebutuhan para pemangku kepentingan dan kondisi nyata di lapangan. Temuan menunjukkan bahwa tidak ada model evaluasi yang sepenuhnya unggul dibandingkan model lainnya karena masing-masing memiliki fokus, kelebihan, dan keterbatasan yang berbeda. Oleh sebab itu, pemilihan model evaluasi perlu disesuaikan dengan tujuan, karakteristik program, serta kebutuhan informasi yang ingin diperoleh.

Kata Kunci: Evaluasi pendidikan, model *Tyler*, model *Scriven*, model *CIPP*, model *Stake*.

A. Pendahuluan

Evaluasi pendidikan merupakan proses sistematis untuk memperoleh, menganalisis, dan menafsirkan informasi yang digunakan sebagai dasar dalam pengambilan keputusan terkait efektivitas suatu program pendidikan. Dalam konteks

pendidikan modern, evaluasi tidak hanya berfungsi untuk mengukur hasil belajar peserta didik, tetapi juga

untuk menilai kualitas program, efektivitas kurikulum, kompetensi pendidik, serta keberhasilan kebijakan pendidikan yang diterapkan.

Perkembangan ilmu evaluasi menunjukkan adanya perubahan paradigma dari evaluasi yang hanya berorientasi pada hasil menjadi evaluasi yang memperhatikan proses, konteks, serta dampak program secara menyeluruh. Perubahan

tersebut melahirkan berbagai model evaluasi yang dikembangkan oleh para ahli dengan perspektif yang berbeda. Di antara model evaluasi yang paling berpengaruh dalam dunia pendidikan adalah Model Tyler, Model Scriven, Model CIPP yang dikembangkan oleh Stufflebeam, dan Model Stake.

Model Tyler dikenal sebagai model evaluasi klasik yang berorientasi pada tujuan (*objective-oriented evaluation*). Sebaliknya, Scriven mengkritik pendekatan yang terlalu berfokus pada tujuan dan mengembangkan konsep evaluasi formatif, sumatif, serta *goal-free evaluation*. Stufflebeam kemudian memperluas cakupan evaluasi melalui model CIPP yang menekankan pentingnya evaluasi dalam pengambilan keputusan. Sementara itu, Stake mengembangkan model evaluasi responsif yang lebih memperhatikan kebutuhan pemangku kepentingan dan realitas yang terjadi di lapangan.

Keberagaman model evaluasi tersebut menunjukkan bahwa evaluasi pendidikan tidak dapat dilakukan dengan satu pendekatan tunggal. Oleh karena itu, diperlukan pemahaman yang mendalam

mengenai karakteristik masing-masing model agar evaluator dapat memilih pendekatan yang paling sesuai dengan kebutuhan evaluasi yang dilakukan.

B. Metode Penelitian

Artikel ini menggunakan metode penelitian kepustakaan (*library research*). Data diperoleh dari berbagai sumber ilmiah berupa buku, jurnal nasional dan internasional, artikel penelitian, serta dokumen akademik yang membahas teori dan model evaluasi pendidikan. Analisis data dilakukan secara deskriptif dan komparatif dengan mengidentifikasi konsep dasar, karakteristik, kelebihan, kelemahan, serta relevansi masing-masing model evaluasi dalam konteks pendidikan.

C. Hasil Penelitian dan Pembahasan

1 Model Evaluasi Tyler

Model evaluasi Tyler merupakan salah satu model evaluasi pendidikan yang paling awal dan memiliki pengaruh besar terhadap perkembangan teori evaluasi modern. Model ini dikembangkan oleh Ralph W. Tyler yang dikenal sebagai tokoh penting dalam bidang kurikulum dan evaluasi pendidikan. Tyler

memandang evaluasi sebagai suatu proses yang digunakan untuk mengetahui sejauh mana tujuan pendidikan yang telah dirumuskan dapat dicapai melalui proses pembelajaran. Oleh karena itu, model ini sering disebut sebagai *objective-oriented evaluation* atau evaluasi yang berorientasi pada tujuan. Dalam pandangan Tyler, seluruh kegiatan pendidikan harus dirancang berdasarkan tujuan yang jelas sehingga keberhasilan suatu program dapat diukur secara sistematis melalui tingkat ketercapaian tujuan tersebut. Pendekatan ini memberikan landasan yang kuat bagi pengembangan kurikulum karena tujuan pendidikan menjadi pedoman utama dalam menentukan isi kurikulum, strategi pembelajaran, serta bentuk evaluasi yang digunakan.

Menurut Tyler, pengembangan kurikulum dan evaluasi pendidikan harus menjawab empat pertanyaan mendasar. Pertama, tujuan pendidikan apa yang ingin dicapai oleh sekolah atau lembaga pendidikan. Kedua, pengalaman belajar apa yang perlu diberikan kepada peserta didik agar tujuan tersebut dapat tercapai. Ketiga, bagaimana pengalaman belajar

tersebut diorganisasikan secara efektif sehingga mampu menghasilkan perubahan perilaku yang diharapkan. Keempat, bagaimana cara menentukan bahwa tujuan pendidikan tersebut telah berhasil dicapai. Keempat pertanyaan tersebut membentuk kerangka berpikir yang sistematis dan logis sehingga seluruh komponen pendidikan saling berkaitan satu sama lain. Dengan demikian, evaluasi bukan hanya dilakukan pada akhir pembelajaran, tetapi sudah menjadi bagian integral dari keseluruhan proses perencanaan pendidikan.

Tyler juga menjelaskan bahwa evaluasi harus mampu menunjukkan perubahan perilaku peserta didik sebagai hasil dari proses pembelajaran yang telah dilaksanakan. Oleh karena itu, evaluasi dilakukan dengan membandingkan kondisi peserta didik sebelum mengikuti program pendidikan dengan kondisi setelah program selesai dilaksanakan. Perubahan yang terjadi kemudian dianalisis untuk mengetahui apakah tujuan yang telah dirumuskan benar-benar tercapai. Dalam pelaksanaannya, Tyler menyarankan penggunaan berbagai instrumen

evaluasi agar data yang diperoleh lebih akurat dan objektif. Pendekatan ini menunjukkan bahwa evaluasi tidak hanya berfungsi sebagai alat pengukuran, tetapi juga sebagai sarana untuk memperbaiki kualitas program pendidikan secara keseluruhan.

Salah satu kelebihan utama model Tyler adalah kesederhanaannya sehingga mudah dipahami dan diterapkan oleh guru maupun pengelola pendidikan. Model ini memberikan arah yang jelas mengenai apa yang harus dicapai dan bagaimana keberhasilan program dapat diukur. Selain itu, adanya tujuan yang dirumuskan secara spesifik membantu evaluator dalam menentukan indikator keberhasilan serta instrumen evaluasi yang sesuai. Namun demikian, model Tyler juga memiliki beberapa keterbatasan. Fokus yang terlalu besar pada tujuan yang telah ditetapkan sering kali menyebabkan evaluator mengabaikan dampak lain yang muncul selama program berlangsung. Model ini juga cenderung lebih menekankan hasil akhir daripada proses pelaksanaan program. Akibatnya, berbagai faktor yang memengaruhi keberhasilan program

selama proses berlangsung sering kali kurang mendapat perhatian. Oleh karena itu, meskipun model Tyler masih relevan digunakan dalam evaluasi hasil belajar, penggunaannya perlu dilengkapi dengan pendekatan lain agar hasil evaluasi menjadi lebih komprehensif.

2 Model Evaluasi Scriven

Model evaluasi Scriven dikembangkan oleh Michael Scriven sebagai respons terhadap keterbatasan model evaluasi yang hanya berorientasi pada tujuan. Scriven berpendapat bahwa evaluasi seharusnya tidak hanya berfokus pada ketercapaian tujuan yang telah dirumuskan sebelumnya, tetapi juga harus mampu menilai nilai, manfaat, dan dampak nyata dari suatu program. Menurutnya, suatu program dapat saja berhasil mencapai tujuan yang telah ditetapkan, tetapi belum tentu memberikan manfaat yang optimal bagi peserta didik atau masyarakat. Oleh karena itu, evaluasi harus dilakukan secara lebih luas dengan memperhatikan berbagai aspek yang muncul selama pelaksanaan program.

Kontribusi terbesar Scriven dalam dunia evaluasi adalah pengembangan konsep evaluasi

formatif dan evaluasi sumatif. Evaluasi formatif dilakukan selama program berlangsung dengan tujuan memberikan umpan balik kepada pelaksana program sehingga perbaikan dapat dilakukan sebelum program selesai. Dalam konteks pendidikan, evaluasi formatif dapat berupa observasi pembelajaran, tugas harian, kuis, atau refleksi yang dilakukan guru terhadap proses pembelajaran. Informasi yang diperoleh dari evaluasi formatif digunakan untuk mengidentifikasi kelemahan program dan merancang strategi perbaikan agar tujuan yang diharapkan dapat tercapai secara lebih efektif. Dengan demikian, evaluasi formatif memiliki fungsi pengembangan karena membantu meningkatkan kualitas program selama proses berlangsung.

Berbeda dengan evaluasi formatif, evaluasi sumatif dilakukan setelah program selesai dilaksanakan. Tujuan utama evaluasi sumatif adalah untuk menentukan tingkat keberhasilan program secara keseluruhan. Hasil evaluasi sumatif biasanya digunakan sebagai dasar pengambilan keputusan mengenai apakah program perlu dilanjutkan, direvisi, diperluas, atau bahkan

dihentikan. Dalam dunia pendidikan, ujian akhir semester, ujian nasional, atau laporan hasil belajar merupakan contoh evaluasi sumatif yang sering digunakan. Melalui evaluasi sumatif, pengambil kebijakan dapat memperoleh informasi mengenai efektivitas dan efisiensi program yang telah dilaksanakan.

Selain evaluasi formatif dan sumatif, Scriven juga memperkenalkan konsep Goal-Free Evaluation atau evaluasi bebas tujuan. Dalam pendekatan ini, evaluator tidak perlu mengetahui tujuan program yang telah dirumuskan sebelumnya. Evaluator lebih difokuskan pada identifikasi dampak nyata yang dihasilkan oleh program, baik dampak yang direncanakan maupun yang tidak direncanakan. Pendekatan ini bertujuan untuk menghindari bias yang mungkin muncul ketika evaluator terlalu terfokus pada tujuan tertentu. Dengan mengevaluasi seluruh dampak program secara terbuka, evaluator dapat menemukan berbagai efek positif maupun negatif yang sebelumnya tidak diperkirakan oleh perancang program.

Meskipun memberikan perspektif yang lebih luas dalam

evaluasi, model Scriven juga memiliki tantangan tersendiri. Pelaksanaan evaluasi bebas tujuan memerlukan evaluator yang sangat kompeten karena harus mampu mengidentifikasi berbagai dampak program tanpa menggunakan tujuan sebagai acuan utama. Selain itu, proses pengumpulan data sering kali membutuhkan waktu dan sumber daya yang lebih besar dibandingkan model evaluasi tradisional. Namun demikian, konsep evaluasi formatif, sumatif, dan goal-free evaluation yang dikembangkan oleh Scriven hingga saat ini masih menjadi salah satu fondasi utama dalam praktik evaluasi pendidikan modern.

3 Model Evaluasi Stufflebeam (CIPP)

Model evaluasi CIPP dikembangkan oleh Daniel Stufflebeam sebagai suatu pendekatan evaluasi yang berorientasi pada pengambilan keputusan. Berbeda dengan model Tyler yang berfokus pada pencapaian tujuan, model CIPP memandang evaluasi sebagai alat untuk membantu pengelola program dalam membuat keputusan yang tepat pada setiap tahap pelaksanaan program. Oleh karena itu, model ini dirancang untuk

memberikan informasi yang lengkap dan berkelanjutan sejak tahap perencanaan hingga tahap penilaian hasil program.

Komponen pertama dalam model CIPP adalah Context Evaluation atau evaluasi konteks. Evaluasi konteks bertujuan untuk mengidentifikasi kebutuhan, masalah, peluang, dan kondisi lingkungan yang melatarbelakangi pelaksanaan suatu program. Dalam dunia pendidikan, evaluasi konteks dapat dilakukan melalui analisis kebutuhan peserta didik, kondisi sekolah, kebijakan pendidikan, serta tuntutan masyarakat. Informasi yang diperoleh dari evaluasi konteks sangat penting karena menjadi dasar dalam menentukan tujuan dan prioritas program yang akan dilaksanakan. Tanpa pemahaman yang baik mengenai konteks program, tujuan yang dirumuskan berpotensi tidak sesuai dengan kebutuhan nyata yang ada di lapangan.

Komponen kedua adalah Input Evaluation atau evaluasi masukan. Evaluasi ini berfokus pada penilaian terhadap berbagai sumber daya yang digunakan untuk mencapai tujuan program. Sumber daya tersebut meliputi tenaga pendidik, sarana dan

prasarana, anggaran, kurikulum, serta strategi pelaksanaan program. Melalui evaluasi input, pengelola program dapat menentukan apakah sumber daya yang tersedia sudah memadai atau masih memerlukan perbaikan sebelum program dijalankan. Dengan demikian, evaluasi input berfungsi untuk meningkatkan peluang keberhasilan program sejak tahap awal perencanaan.

Komponen ketiga adalah Process Evaluation atau evaluasi proses. Evaluasi ini dilakukan selama program berlangsung dengan tujuan memantau pelaksanaan program dan memastikan bahwa seluruh kegiatan berjalan sesuai dengan rencana yang telah ditetapkan. Evaluasi proses memungkinkan evaluator mengidentifikasi berbagai hambatan yang muncul selama pelaksanaan program sehingga tindakan korektif dapat segera dilakukan. Dalam konteks pembelajaran, evaluasi proses dapat dilakukan melalui observasi kelas, supervisi akademik, maupun wawancara dengan peserta didik dan guru.

Komponen terakhir adalah Product Evaluation atau evaluasi hasil. Evaluasi ini bertujuan untuk menilai efektivitas program

berdasarkan hasil yang dicapai serta dampak yang ditimbulkan. Penilaian tidak hanya dilakukan terhadap hasil jangka pendek, tetapi juga terhadap dampak jangka panjang yang mungkin muncul setelah program selesai dilaksanakan. Informasi yang diperoleh dari evaluasi produk kemudian digunakan sebagai dasar untuk menentukan apakah program perlu dilanjutkan, dimodifikasi, atau dihentikan.

Keunggulan utama model CIPP adalah kemampuannya memberikan gambaran yang menyeluruh mengenai suatu program. Namun, kompleksitas model ini sering kali menjadi tantangan karena membutuhkan waktu, biaya, serta tenaga evaluator yang cukup besar. Meskipun demikian, model CIPP tetap menjadi salah satu model evaluasi yang paling banyak digunakan dalam berbagai bidang pendidikan karena mampu menghasilkan informasi yang komprehensif untuk mendukung pengambilan keputusan.

4. Model Evaluasi Stake

Model evaluasi Stake atau yang dikenal sebagai Responsive Evaluation Model dikembangkan oleh Robert Stake sebagai alternatif terhadap model evaluasi yang terlalu

berorientasi pada tujuan dan prosedur yang kaku. Stake berpendapat bahwa evaluasi harus berfokus pada kebutuhan informasi yang benar-benar dibutuhkan oleh para pemangku kepentingan atau stakeholders. Oleh karena itu, model ini lebih menekankan pemahaman terhadap pengalaman nyata para pelaksana dan pengguna program dibandingkan sekadar mengukur ketercapaian tujuan yang telah dirumuskan sebelumnya.

Dalam model Stake, evaluator dituntut untuk memahami berbagai isu yang berkembang selama pelaksanaan program. Fokus evaluasi tidak ditentukan sepenuhnya oleh tujuan program, tetapi juga oleh kebutuhan, harapan, dan kepentingan para stakeholder. Dengan pendekatan ini, evaluasi menjadi lebih fleksibel dan mampu menangkap berbagai fenomena yang muncul di lapangan. Data yang dikumpulkan umumnya bersifat kualitatif, seperti hasil wawancara, observasi, dokumentasi, dan diskusi kelompok. Pendekatan tersebut memungkinkan evaluator memperoleh gambaran yang lebih mendalam mengenai kondisi nyata yang terjadi selama program berlangsung.

Model Stake memiliki tiga komponen utama yaitu antecedents, transactions, dan outcomes. Antecedents berkaitan dengan kondisi awal sebelum program dilaksanakan, transactions berkaitan dengan proses pelaksanaan program, sedangkan outcomes berkaitan dengan hasil dan dampak program. Ketiga komponen tersebut dianalisis secara terpadu untuk memberikan gambaran yang utuh mengenai kualitas program yang dievaluasi.

Keunggulan model Stake terletak pada kemampuannya menghasilkan informasi yang kaya, mendalam, dan sesuai dengan kebutuhan stakeholder. Namun demikian, penggunaan data kualitatif yang dominan menyebabkan hasil evaluasi sering kali bersifat subjektif dan sulit digeneralisasikan. Oleh karena itu, evaluator harus memiliki kemampuan analisis yang baik agar informasi yang diperoleh tetap valid dan dapat dipertanggungjawabkan secara ilmiah.

D. Kesimpulan

Keempat model evaluasi yang dikembangkan oleh Tyler, Scriven, Stufflebeam, dan Stake memberikan kontribusi signifikan terhadap

perkembangan evaluasi pendidikan. Model Tyler menekankan pencapaian tujuan, Model Scriven berorientasi pada nilai dan manfaat program, Model CIPP berfokus pada pengambilan keputusan melalui evaluasi yang komprehensif, sedangkan Model Stake menekankan kebutuhan dan perspektif stakeholder.

Pemilihan model evaluasi tidak dapat dilakukan secara sembarangan karena harus mempertimbangkan tujuan evaluasi, karakteristik program, sumber daya yang tersedia, serta kebutuhan informasi yang ingin diperoleh. Dalam praktiknya, kombinasi beberapa model evaluasi sering kali menghasilkan informasi yang lebih komprehensif dibandingkan penggunaan satu model secara terpisah.

DAFTAR PUSTAKA

- Anirowati, Lilik, Siti Rahma, and Nur Komariah. "Education Evaluation Models Model-Model Evaluasi Pendidikan" 1, no. 3 (2025): 133–45.
- Hasbi, H., Ardiansyah, A., & Mulyan, A. (2025). "Model-Model Evaluasi Pendidikan: Telaah Mendalam Kirkpatrick, CIPP, Dan Outcome-Based Evaluation" 5 (2025): 1–8.
- Hidayat, Tatang, Endis Firdaus, and Momod Abdul Somad. "Model Pengembangan Kurikulum Tyler Dan Impikasinya Dalam Pembelajaran Pendidikan Agama Islam Di Sekolah." *Jurnal Kependidikan Islam* 5, no. 2 (2019): 197–218.
- Irfan, Muhammad Khairul, and Mohammad Sofwan Effendi. "Evaluasi Program Pengimbasan Menggunakan Model CIPP Sekolah Penggerak" 14, no. 2 (2025): 2519–30.
- Nukhbatillah, Isyfi Agni, Santi Setiawati, Uswatun Hasanah, and Neneng Nurmalasari. "Evaluasi Mutu Pendidikan Menggunakan Pendekatan Teori Stufflebeam" 2, no. 1 (2024): 34–43. <https://doi.org/10.59996/globalistika.v2i1.352>.
- Permata, Alvin Tassar, Angga Hadiapurwa, and Yayu Wulandari. "Evaluasi Program 'Pustakawan Terpelajar' Dengan Model Stake Di MAN 1 Bandung" 6, no. 1 (2025).
- "PERTEMUAN KE 4 SEMESTER 5 Model-Model Pengembangan Kurikulum," n.d.
- Rahmat, Zetri, Universita Islam, Negeri Sultan, Syarif Kasim, and Universitas Negeri Padang. "JKIP : Jurnal Kajian Ilmu Pendidikan The CIPP Evaluation Model in School Programs : A Systematic Literature Review Model Evaluasi CIPP Dalam Program Sekolah : Systematic

Literature Review” 5, no. 4
(2025): 911–19.

Wardani, Helda Kusuma, Fajarsih
Darusuprapti, and Mami Hajaroh.
“Model-Model Evaluasi
Pendidikan Dasar (Scriven
Model , Tyler Model , Dan Goal
Free Evaluation)” 6, no. 1 (2022):
36–49.