



EDUCATION EVALUATION MODELS

MODEL-MODEL EVALUASI PENDIDIKAN

Lilik Anirowati^{1*}, Siti Rahma², Nur Komariah³

Universitas Islam Indragiri Riau^{1,2,3}

E-mail: lilikanirowati14@gmail.com¹, mutiaraamuti87@gmail.com², nurkomariah7179@gmail.com³

*email koresponden: lilikanirowati14@gmail.com

DOI: <https://doi.org/10.62567/ijosse.v1i3.1257>

Abstract

This study aims to examine and analyze various educational evaluation models used to measure and assess the effectiveness of learning programs and educational policies. Using a qualitative approach through library research, this study explores the concepts, characteristics, strengths, and weaknesses of several well-known evaluation models, including the CIPP model, Tyler model, Countenance model, Stake model, and Kirkpatrick model. The findings indicate that each model has distinct orientations and approaches, depending on the objectives, context, and type of program being evaluated. The CIPP model is commonly applied in comprehensive evaluations focused on program improvement, while the Tyler model emphasizes the achievement of instructional objectives. The study also reveals that no single evaluation model is universally applicable; therefore, selecting an appropriate model must be aligned with the specific needs and characteristics of the educational institution. These findings reinforce the importance of a deep understanding of evaluation theory and flexibility in its implementation. Thus, this research is expected to serve as a reference for educators, education planners, and researchers in determining suitable evaluation approaches to improve the quality of education.

Keywords: *educational evaluation, evaluation models, CIPP, Tyler, program evaluation.*

Abstrak

Penelitian ini bertujuan untuk mengkaji dan menganalisis berbagai model evaluasi pendidikan yang digunakan dalam proses pengukuran dan penilaian efektivitas program pembelajaran maupun kebijakan pendidikan. Melalui pendekatan kualitatif dengan metode studi kepustakaan, penelitian ini menelusuri konsep, karakteristik, serta kelebihan dan kekurangan dari beberapa model evaluasi pendidikan yang paling dikenal, seperti model CIPP, model Tyler, model Countenance, model Stake, dan model Kirkpatrick. Hasil kajian menunjukkan bahwa setiap model memiliki orientasi dan pendekatan yang berbeda-beda, tergantung pada tujuan, konteks, serta jenis program yang



dievaluasi. Model CIPP cenderung digunakan untuk evaluasi yang bersifat komprehensif dan berorientasi pada perbaikan program, sementara model Tyler lebih fokus pada pencapaian tujuan instruksional. Penelitian ini juga menemukan bahwa tidak ada satu model evaluasi yang bersifat universal, sehingga pemilihan model evaluasi harus disesuaikan dengan kebutuhan dan karakteristik lembaga pendidikan. Temuan ini memperkuat pentingnya pemahaman mendalam terhadap teori evaluasi serta fleksibilitas dalam implementasinya. Dengan demikian, penelitian ini diharapkan dapat menjadi referensi bagi pendidik, perencana pendidikan, dan peneliti dalam menentukan pendekatan evaluasi yang tepat guna meningkatkan mutu pendidikan.

Kata Kunci: Evaluasi Pendidikan, Model Evaluasi, CIPP, Tyler, Evaluasi Program.

1. PENDAHULUAN

Pendidikan merupakan salah satu aspek fundamental dalam pembangunan sumber daya manusia dan kemajuan suatu bangsa (Faridah et al., 2023). Dalam proses pendidikan, berbagai aktivitas dilakukan untuk mencapai tujuan pembelajaran, seperti perencanaan, pelaksanaan, dan penilaian hasil belajar. Salah satu komponen penting yang menentukan keberhasilan sistem pendidikan adalah evaluasi. Evaluasi pendidikan tidak hanya berfungsi untuk mengukur hasil belajar peserta didik, tetapi juga menjadi alat untuk menilai efektivitas proses pembelajaran, kinerja pendidik, dan mutu lembaga pendidikan secara keseluruhan (Wirati et al., 2024).

Secara umum, evaluasi dalam pendidikan bertujuan untuk mengumpulkan informasi secara sistematis guna membuat keputusan yang tepat tentang program, proses, dan hasil pembelajaran (Nugraha, 2023a). Evaluasi bukan hanya menjadi alat ukur semata, melainkan juga menjadi landasan pengambilan keputusan dalam merancang program pendidikan yang lebih efektif, efisien, dan relevan dengan kebutuhan peserta didik serta tuntutan masyarakat (Aulita et al., 2024). Oleh karena itu, pemahaman yang mendalam tentang model-model evaluasi pendidikan sangatlah penting bagi para pendidik, pengelola pendidikan, maupun perancang kebijakan.

Dalam konteks implementasinya, berbagai model evaluasi pendidikan telah dikembangkan oleh para ahli guna memberikan pendekatan sistematis dalam proses penilaian (Ginanjari et al., 2023). Setiap model memiliki tujuan, pendekatan, dan kerangka kerja yang berbeda sesuai dengan kebutuhan dan konteks evaluasi (Joko & Nugraha, 2023). Misalnya, model Tyler yang berorientasi pada tujuan (*objective-oriented evaluation*), model CIPP (*Context, Input, Process, Product*) yang dikembangkan oleh Stufflebeam dan berfokus pada pengambilan keputusan, hingga model Kirkpatrick yang banyak digunakan untuk mengevaluasi pelatihan dan pengembangan sumber daya manusia. Selain itu, terdapat pula model responsif, model *goal-free*, model *discrepancy*, serta pendekatan evaluasi berbasis nilai dan *stakeholder*.

Keragaman model evaluasi ini mencerminkan bahwa tidak ada satu pendekatan evaluasi yang cocok untuk semua situasi (Joko et al., 2024). Dalam praktiknya, pemilihan model evaluasi harus disesuaikan dengan tujuan evaluasi, jenis program yang dievaluasi, serta sumber daya yang tersedia (Nugraha, 2023b). Penguasaan terhadap model-model evaluasi akan membantu pendidik



dan evaluator dalam memilih pendekatan yang paling relevan dan tepat guna untuk memastikan bahwa program atau kebijakan pendidikan yang dijalankan benar-benar memberikan dampak positif dan berkelanjutan.

Namun, dalam kenyataannya masih banyak pendidik dan pelaku pendidikan yang belum memahami perbedaan, kelebihan, kekurangan, serta konteks penggunaan dari masing-masing model evaluasi tersebut. Hal ini menyebabkan proses evaluasi seringkali dilakukan secara terbatas, hanya berfokus pada hasil akhir tanpa memperhatikan proses dan faktor lain yang memengaruhi keberhasilan pembelajaran. Padahal, evaluasi yang holistik dan komprehensif sangat dibutuhkan untuk memperbaiki kualitas pendidikan secara menyeluruh.

Berdasarkan latar belakang tersebut, penulis merasa perlu untuk mengkaji dan menguraikan secara mendalam mengenai berbagai model evaluasi pendidikan yang berkembang, baik dari sisi konsep, prinsip kerja, hingga penerapannya dalam konteks pendidikan di Indonesia. Dengan demikian, makalah ini diharapkan dapat memberikan pemahaman yang lebih luas dan mendalam kepada pembaca mengenai pentingnya model-model evaluasi dalam mendukung tercapainya tujuan pendidikan yang berkualitas.

2. METODOLOGI

Penelitian ini menggunakan pendekatan kualitatif dengan metode studi kepustakaan (*library research*). Pendekatan ini dipilih karena fokus utama penelitian adalah untuk mengkaji, menganalisis, dan mendeskripsikan berbagai model evaluasi pendidikan dari sudut pandang teori dan praktik yang berkembang dalam literatur akademik. Sumber data utama berasal dari buku-buku ilmiah, artikel jurnal, hasil penelitian sebelumnya, dan dokumen-dokumen resmi yang relevan dengan tema evaluasi pendidikan (Amelia et al., 2023).

Teknik pengumpulan data dilakukan melalui penelusuran dan analisis literatur dari berbagai sumber baik cetak maupun digital (Muslimin et al., 2024). Data yang terkumpul kemudian dianalisis secara deskriptif-kualitatif dengan menekankan pada identifikasi konsep, karakteristik, prinsip dasar, kelebihan, kekurangan, serta penerapan masing-masing model evaluasi pendidikan, seperti model CIPP, model Countenance, model Stake, model Kirkpatrick, dan model Tyler.

Validitas data diperkuat melalui teknik triangulasi sumber, yaitu dengan membandingkan informasi dari berbagai referensi yang memiliki otoritas akademik untuk memastikan keakuratan dan kelengkapan data (Nugraha, 2025). Peneliti juga melakukan telaah kritis terhadap setiap model evaluasi yang dikaji, untuk memperoleh pemahaman yang utuh dan kontekstual mengenai penerapannya dalam praktik pendidikan di berbagai jenjang. Dengan pendekatan ini, diharapkan penelitian mampu memberikan kontribusi teoretis dalam memperkaya kajian tentang evaluasi pendidikan serta menjadi referensi bagi praktisi pendidikan dalam memilih model evaluasi yang sesuai dengan konteks dan tujuan pembelajaran.



3. PEMBAHASAN

a. Model Tyler (*Objective-Oriented Evaluation Model*)

Model ini dikembangkan oleh Ralph Tyler pada tahun 1949 dan menjadi salah satu model evaluasi yang paling awal dan mendasar. Model Tyler menekankan bahwa evaluasi harus berorientasi pada tujuan pembelajaran (Agung Wibowo et al., 2022). Artinya, evaluasi dilakukan untuk mengukur sejauh mana tujuan pembelajaran yang telah ditetapkan berhasil dicapai. Langkah-langkah dalam model Tyler:

- 1) Menentukan tujuan pembelajaran.
- 2) Mengidentifikasi pengalaman belajar yang mendukung pencapaian tujuan.
- 3) Mengorganisasikan pengalaman belajar secara efektif.
- 4) Menilai apakah tujuan telah tercapai.

Kelebihan model ini adalah mudah dipahami dan diterapkan. Namun, kelemahannya terletak pada pendekatannya yang terlalu fokus pada hasil akhir (*output*) dan mengabaikan proses.

b. Model CIPP (*Context, Input, Process, Product*)

Model ini dikembangkan oleh Daniel Stufflebeam dan bertujuan untuk membantu pengambilan keputusan dalam program pendidikan (Ali & Sukardi, 2021). CIPP merupakan model evaluasi yang komprehensif karena mengevaluasi dari tahap perencanaan hingga hasil. Empat komponen dalam model CIPP:

- 1) *Context Evaluation*: Menilai kebutuhan dan latar belakang program.
- 2) *Input Evaluation*: Menilai strategi, sumber daya, dan perencanaan.
- 3) *Process Evaluation*: Menilai pelaksanaan program.
- 4) *Product Evaluation*: Menilai hasil dan dampak program.

Kelebihan model CIPP adalah mampu memberikan informasi yang lengkap untuk pengambilan keputusan. Namun, model ini cukup kompleks dan memerlukan waktu serta biaya yang besar.

c. Model Stake (*Responsive Evaluation Model*)

Dikembangkan oleh Robert Stake, model ini lebih menekankan pada responsivitas terhadap kebutuhan informasi dari pihak-pihak terkait (*stakeholders*) (Bhakti, 2017). Evaluasi tidak hanya berdasarkan tujuan awal, tetapi juga memperhatikan aspek-aspek yang muncul dalam pelaksanaan program. Model Stake menekankan pada:

- 1) Fokus pada isu-isu yang relevan di lapangan.
- 2) Dialog dengan para pengguna hasil evaluasi.
- 3) Penggunaan data kualitatif dan naratif.

Kelebihan dari model ini adalah fleksibel dan responsif terhadap perubahan. Kekurangannya adalah kurang terstandar sehingga rawan bias subjektivitas. Model evaluasi responsif yang dikembangkan oleh Robert Stake merupakan pendekatan evaluasi yang berorientasi pada kebutuhan nyata dan informasi yang relevan bagi para pemangku kepentingan (*stakeholders*) dalam pelaksanaan suatu program. Tidak seperti model-model evaluasi tradisional yang terfokus secara ketat pada tujuan formal yang telah dirumuskan di



awal, model ini menekankan pentingnya responsivitas terhadap dinamika dan kompleksitas yang terjadi selama program berlangsung. Evaluasi dalam model ini bersifat lebih terbuka, fleksibel, dan kontekstual, dengan memberi perhatian khusus pada isu-isu yang muncul di lapangan serta pada harapan dan kebutuhan dari para pengguna hasil evaluasi.

Salah satu karakteristik utama dari model ini adalah fokus pada isu-isu yang dianggap relevan dan penting oleh para *stakeholder*. Artinya, evaluator tidak hanya berpedoman pada indikator yang telah ditentukan sebelumnya, tetapi juga menangkap masalah, kekuatan, kelemahan, dan dampak yang tampak secara nyata dalam proses pelaksanaan program. Evaluasi tidak dipandang sebagai kegiatan teknis semata, tetapi sebagai bagian dari komunikasi dua arah antara evaluator dan pihak-pihak yang berkepentingan terhadap program tersebut.

Model ini juga menekankan pentingnya dialog dan komunikasi dengan para pengguna hasil evaluasi, seperti pengelola program, guru, siswa, orang tua, maupun pihak lain yang terlibat langsung atau tidak langsung. Dengan begitu, proses evaluasi menjadi lebih partisipatif dan hasilnya lebih relevan serta dapat digunakan secara langsung dalam pengambilan keputusan atau perbaikan program.

Dari sisi metode, model Stake cenderung mengandalkan data kualitatif dan naratif, seperti wawancara, observasi, dan dokumentasi, untuk menangkap dinamika sosial dan perilaku yang tidak selalu dapat dijelaskan dengan angka. Evaluasi disusun dalam bentuk laporan deskriptif yang menyajikan konteks dan interpretasi yang kaya, dengan harapan mampu mencerminkan kenyataan di lapangan secara utuh.

Model ini memiliki kelebihan utama berupa fleksibilitas dalam pelaksanaannya. Evaluasi dapat disesuaikan dengan perkembangan situasi, tidak kaku terhadap instrumen awal, dan terbuka terhadap temuan baru yang sebelumnya tidak direncanakan. Pendekatan ini sangat cocok untuk mengevaluasi program-program yang bersifat kompleks, kontekstual, dan melibatkan banyak pihak dengan berbagai kepentingan yang berbeda.

Namun, model ini juga memiliki kelemahan yang perlu diwaspadai, yakni kurangnya struktur dan standar baku dalam pelaksanaan evaluasi. Hal ini dapat menyebabkan proses evaluasi menjadi terlalu bergantung pada interpretasi evaluator, sehingga berisiko tinggi terhadap subjektivitas dan bias, terutama jika evaluator tidak memiliki pengalaman atau kepekaan terhadap konteks sosial dan budaya di mana program dijalankan.

Dalam praktiknya, model Stake sangat sesuai untuk digunakan dalam evaluasi program pendidikan berbasis masyarakat, program pengembangan sekolah, atau inovasi pembelajaran yang melibatkan interaksi sosial yang kuat. Dengan pendekatan yang menekankan pada kenyataan dan kebutuhan di lapangan, model ini memungkinkan hasil evaluasi yang lebih bermakna dan langsung dapat dimanfaatkan oleh pengguna akhir.

d. Model Kirkpatrick

Model ini populer digunakan dalam evaluasi pelatihan dan dikembangkan oleh Donald Kirkpatrick (Damanik et al., 2020). Model ini terdiri dari empat level evaluasi:

- 1) *Reaction* – Bagaimana respon peserta terhadap program.



- 2) *Learning* – Sejauh mana pengetahuan atau keterampilan peserta meningkat.
- 3) *Behavior* – Apakah ada perubahan perilaku setelah mengikuti program.
- 4) *Result* – Dampak akhir dari program terhadap organisasi atau institusi.

Model ini cocok digunakan dalam konteks pendidikan nonformal atau pelatihan. Kelemahannya adalah sulit untuk mengukur level yang lebih tinggi seperti perubahan perilaku dan hasil dalam jangka panjang. Model evaluasi Kirkpatrick merupakan pendekatan yang sangat populer dan banyak digunakan dalam dunia pelatihan dan pendidikan nonformal. Model ini dikembangkan oleh Donald Kirkpatrick pada akhir tahun 1950-an dan hingga kini tetap menjadi acuan utama dalam mengevaluasi efektivitas suatu program pembelajaran atau pelatihan. Ciri khas utama dari model ini adalah pendekatan bertingkat yang dimulai dari evaluasi awal mengenai reaksi peserta hingga ke tingkat evaluasi yang lebih kompleks terkait dampak program terhadap organisasi (Turmuzy et al., 2022).

Pada tingkat pertama, evaluasi difokuskan pada reaksi peserta terhadap program. Ini mencakup bagaimana peserta merespons materi, metode penyampaian, fasilitator, dan suasana pelatihan. Evaluasi ini penting karena memberikan gambaran awal mengenai kepuasan peserta dan potensi keberhasilan program dari sisi penerimaan. Jika peserta merasa tidak nyaman atau tidak puas, kemungkinan besar proses belajar tidak akan optimal. Tingkat selanjutnya mengevaluasi sejauh mana peserta telah memperoleh pengetahuan, keterampilan, atau perubahan sikap setelah mengikuti program. Ini biasanya dilakukan dengan membandingkan kondisi sebelum dan sesudah pelatihan melalui tes atau tugas tertentu. Evaluasi ini bertujuan untuk memastikan bahwa ada peningkatan kapasitas individu yang relevan dengan tujuan pelatihan. Selanjutnya, model Kirkpatrick mengevaluasi perubahan perilaku peserta setelah kembali ke lingkungan kerja atau kehidupan sehari-hari. Evaluasi pada tahap ini bertujuan untuk menilai apakah peserta benar-benar menerapkan apa yang telah mereka pelajari dalam praktik nyata. Penilaian ini sering memerlukan waktu yang lebih lama dan melibatkan pihak ketiga, seperti atasan atau rekan kerja, untuk memastikan bahwa perubahan yang terjadi bukan hanya sementara.

Tahapan terakhir dari model ini mengevaluasi dampak akhir dari program pelatihan terhadap organisasi atau institusi secara keseluruhan. Hal ini bisa berupa peningkatan produktivitas, efisiensi kerja, kepuasan pelanggan, atau pencapaian target organisasi. Evaluasi ini sangat penting karena menunjukkan sejauh mana pelatihan memberikan nilai nyata bagi organisasi dalam jangka panjang. Kelebihan dari model evaluasi Kirkpatrick terletak pada kerangka berpikirnya yang sistematis dan menyeluruh. Dengan mengikuti tahapan-tahapan yang runtut, evaluator dapat memperoleh gambaran lengkap mulai dari tanggapan peserta hingga dampak strategis pelatihan. Namun, model ini juga memiliki kelemahan, terutama pada tingkat ketiga dan keempat yang sulit diukur secara objektif. Evaluasi perubahan perilaku dan dampak jangka panjang seringkali terpengaruh oleh berbagai faktor eksternal yang tidak bisa dikendalikan sepenuhnya oleh evaluator. Selain itu, pelaksanaan evaluasi pada level ini memerlukan waktu, biaya, dan sumber daya yang besar.



Dalam konteks pendidikan nonformal seperti pelatihan keterampilan, pelatihan guru, atau program pengembangan masyarakat, model Kirkpatrick tetap sangat relevan. Model ini memungkinkan penyelenggara program untuk melihat efektivitas pelatihan tidak hanya dari sudut pandang proses belajar, tetapi juga dari dampaknya terhadap individu dan organisasi. Dengan demikian, pendekatan ini dapat mendorong peningkatan kualitas program pelatihan secara berkelanjutan dan berbasis data.

e. Model *Countenance* (Robert Stake)

Model ini menyajikan dua dimensi evaluasi, yaitu deskripsi (*descriptive*) dan pertimbangan (*judgment*) yang masing-masing terdiri dari tiga aspek: *antecedents*, *transactions*, dan *outcomes* (Darodjat & Zuchdi, 2016). Model ini mencoba menyatukan pendekatan kualitatif dan kuantitatif dalam evaluasi. Kelebihan: Mampu menyajikan informasi secara holistik. Kekurangan: Kompleks dan memerlukan pemahaman mendalam dalam penggunaannya.

Model *Countenance* Evaluation dikembangkan oleh Robert E. Stake sebagai upaya untuk menghadirkan evaluasi yang lebih komprehensif dan menyeluruh, baik dari segi proses maupun hasil. Model ini memadukan dua dimensi penting dalam evaluasi, yaitu deskripsi (*descriptive*) dan pertimbangan atau penilaian (*judgment*). Setiap dimensi tersebut dianalisis melalui tiga aspek utama, yaitu *antecedents* (kondisi awal), *transactions* (proses), dan *outcomes* (hasil). Dengan struktur ini, model *countenance* mencoba menjembatani antara pendekatan kualitatif dan kuantitatif dalam proses evaluasi.

Dimensi deskriptif bertujuan untuk menggambarkan secara faktual apa yang terjadi dalam program, mulai dari konteks, pelaksanaan, hingga hasil. Evaluator dalam hal ini mengamati dan mencatat data secara sistematis tanpa memberikan penilaian terlebih dahulu. Sementara itu, dimensi pertimbangan memberikan penilaian terhadap temuan deskriptif tersebut, dengan membandingkannya terhadap standar atau kriteria tertentu, sehingga evaluator dapat memberikan simpulan mengenai baik atau tidaknya suatu program.

Model ini menekankan pada pentingnya memahami konteks secara menyeluruh sebelum melakukan penilaian, serta menghargai variasi dan kompleksitas dalam pelaksanaan program. Dengan mencakup aspek *antecedents*, evaluator menilai latar belakang, sumber daya, dan rencana program. Melalui aspek *transactions*, evaluator menilai bagaimana proses pelaksanaan berlangsung, dan melalui aspek *outcomes*, evaluator melihat apa hasil atau dampak dari pelaksanaan program tersebut (Saputra et al., 2014).

Kelebihan utama dari model *countenance* adalah kemampuannya dalam menyajikan informasi secara holistik dan menyeluruh, baik dari segi proses maupun hasil, serta dari sudut pandang objektif maupun subjektif. Pendekatan ini memungkinkan evaluator untuk memahami tidak hanya apa yang terjadi, tetapi juga mengapa dan bagaimana hal itu terjadi, serta apa maknanya bagi pemangku kepentingan.

Namun demikian, kelemahan dari model ini terletak pada tingkat kompleksitasnya yang tinggi. Evaluator dituntut untuk memiliki pemahaman teoritis dan metodologis yang mendalam, terutama dalam hal integrasi data kualitatif dan kuantitatif. Proses evaluasi dengan



model ini juga memerlukan waktu dan sumber daya yang cukup besar, karena harus mengumpulkan dan menganalisis data dari berbagai sudut pandang secara simultan.

Dalam praktiknya, model countenance cocok digunakan dalam evaluasi program pendidikan yang bersifat luas dan berdampak besar, seperti evaluasi kurikulum, program peningkatan mutu guru, atau evaluasi kebijakan pendidikan berbasis masyarakat. Dengan struktur yang sistematis namun fleksibel, model ini memberikan ruang bagi evaluasi yang tidak hanya akurat dari sisi data, tetapi juga bermakna secara kontekstual.

f. Model Goal-Free Evaluation (Michael Scriven)

Model ini berusaha mengevaluasi program tanpa mengetahui tujuan awalnya. Tujuannya adalah untuk melihat semua efek program, termasuk efek tak terduga (Gaspersz et al., 2023). Kelebihan: Dapat menangkap dampak yang tidak direncanakan. Kekurangan: Rentan terhadap bias evaluator dan sulit diterapkan dalam program yang sangat terstruktur.

Model *Goal-Free Evaluation* merupakan pendekatan evaluasi yang dikembangkan oleh Michael Scriven sebagai alternatif dari model-model evaluasi yang berorientasi pada tujuan (*goal-oriented*). Ciri khas utama dari model ini adalah bahwa evaluator secara sengaja tidak diberi tahu mengenai tujuan formal atau tujuan awal dari suatu program yang akan dievaluasi. Tujuan dari pendekatan ini adalah agar evaluator dapat mengamati dan menilai seluruh efek atau dampak dari program secara menyeluruh dan objektif, tanpa terpengaruh oleh ekspektasi atau bias dari tujuan yang telah ditetapkan sebelumnya.

Model ini berangkat dari pemikiran bahwa dalam praktiknya, banyak program menghasilkan dampak-dampak yang tidak terduga, baik yang positif maupun negatif, yang sering kali tidak tercakup dalam tujuan resmi. Jika evaluator hanya fokus pada tujuan yang telah dirumuskan, maka mereka mungkin akan mengabaikan dampak lain yang justru lebih signifikan atau relevan bagi peserta atau masyarakat. Dengan tidak mengetahui tujuan awal, evaluator didorong untuk lebih terbuka, netral, dan jeli dalam melihat semua hasil dari program.

Pendekatan ini sangat cocok digunakan dalam konteks program-program inovatif, eksperimen pendidikan, atau intervensi sosial di mana hasil akhirnya belum sepenuhnya dapat diprediksi. Salah satu kelebihan utama dari model *goal-free evaluation* adalah kemampuannya dalam mengungkap dampak-dampak tersembunyi atau tidak direncanakan, baik yang bersifat positif maupun negatif. Hal ini menjadikan evaluasi lebih kaya dan komprehensif karena tidak terbatas hanya pada parameter formal yang ditetapkan sebelumnya.

Namun demikian, model ini juga memiliki beberapa kekurangan. Salah satu tantangan utamanya adalah potensi munculnya bias dari evaluator, mengingat ia harus mengandalkan penilaiannya sendiri tanpa adanya panduan tujuan. Evaluator yang kurang berpengalaman atau tidak netral dapat terjebak dalam persepsi subjektif terhadap apa yang dianggap sebagai "hasil" yang penting. Selain itu, model ini sulit diterapkan dalam program yang sangat terstruktur dan memiliki tujuan yang ketat, seperti program berskala nasional atau kebijakan pemerintah yang memiliki indikator keberhasilan yang terukur dan harus dipertanggungjawabkan (Rasyid et al., 2022).



Di samping itu, karena evaluator tidak diberi tahu tujuan awal, maka proses pengumpulan data menjadi lebih luas dan terbuka, yang sering kali memerlukan lebih banyak waktu, tenaga, dan biaya. Oleh sebab itu, model ini tidak selalu cocok diterapkan dalam semua konteks evaluasi, terutama ketika keterbatasan sumber daya menjadi kendala.

Dalam konteks pendidikan, model *goal-free* dapat digunakan untuk mengevaluasi program-program pengembangan karakter, pelatihan soft skill, atau pendekatan pembelajaran baru, di mana hasil-hasil yang diharapkan belum dapat dirumuskan secara jelas sejak awal. Dengan demikian, model ini menjadi pendekatan yang relevan untuk mendukung inovasi pendidikan dan memperluas wawasan terhadap efektivitas suatu program dari berbagai sudut pandang yang lebih netral dan terbuka.

g. Model *Discrepancy Evaluation* (Malcolm Provus)

Model ini menilai kesenjangan (*discrepancy*) antara standar yang telah ditetapkan dan kondisi aktual yang terjadi (Mandasari et al., 2025). Evaluasi dilakukan secara berkelanjutan dari tahap input hingga hasil. Kelebihan: Memberikan informasi akurat tentang aspek yang perlu diperbaiki. Kekurangan: Bergantung pada kejelasan standar yang telah ditentukan sebelumnya.¹²

Model *discrepancy* adalah salah satu pendekatan evaluasi pendidikan yang menekankan pada identifikasi kesenjangan antara standar atau kriteria ideal yang telah ditetapkan sebelumnya dengan kondisi nyata atau aktual yang terjadi di lapangan. Model ini pertama kali diperkenalkan oleh Malcolm Provus, dan bertujuan untuk mengungkap sejauh mana suatu program, proses, atau kegiatan pendidikan telah sesuai dengan harapan dan rencana yang telah dirancang.

Dalam pelaksanaannya, model *discrepancy* dilakukan secara berkelanjutan dan menyeluruh mulai dari tahap input (sumber daya, perencanaan, tenaga pengajar), proses (pelaksanaan kegiatan pembelajaran, manajemen kelas), hingga hasil (capaian belajar peserta didik, keberhasilan program). Evaluator akan membandingkan kondisi aktual di setiap tahap tersebut dengan standar atau indikator yang telah ditentukan sebelumnya. Jika ditemukan perbedaan atau ketidaksesuaian (*discrepancy*), maka hal tersebut menjadi fokus utama untuk diperbaiki, diperkuat, atau ditinjau ulang.

Kekuatan utama dari model ini terletak pada kemampuannya dalam memberikan informasi yang sangat spesifik dan akurat mengenai bagian mana dari program pendidikan yang mengalami deviasi dari standar. Dengan demikian, model *discrepancy* sangat berguna dalam perencanaan tindakan perbaikan (*corrective action*) yang berbasis data dan bukti nyata. Evaluasi ini juga mendorong praktik reflektif yang mendalam terhadap pelaksanaan program, serta menumbuhkan budaya peningkatan mutu secara berkelanjutan di lingkungan pendidikan (Mardiah & Syarifudin, 2019).

Namun demikian, model *discrepancy* juga memiliki kelemahan yang perlu diperhatikan. Keberhasilannya sangat bergantung pada sejauh mana standar atau kriteria penilaian dirumuskan dengan jelas, terukur, dan realistis sejak awal. Jika standar yang digunakan bersifat terlalu umum, tidak kontekstual, atau bahkan belum disusun dengan baik, maka



evaluasi yang dilakukan akan menjadi tidak efektif, karena tidak ada acuan pasti yang dapat digunakan untuk menilai deviasi yang terjadi. Di samping itu, model ini memerlukan keterampilan tinggi dari evaluator untuk dapat melakukan analisis secara kritis dan objektif terhadap kesenjangan yang ditemukan.

Dalam konteks pendidikan di Indonesia, penerapan model discrepancy sangat relevan untuk mengevaluasi kebijakan, program sekolah, atau pelaksanaan kurikulum nasional. Misalnya, evaluasi pelaksanaan Kurikulum Merdeka di sekolah-sekolah dapat menggunakan model ini untuk melihat sejauh mana implementasi di lapangan sesuai dengan pedoman resmi dari pemerintah. Dengan pendekatan ini, sekolah dan pemangku kebijakan dapat mengetahui bagian mana yang sudah berjalan dengan baik dan mana yang masih memerlukan pembenahan, sehingga perbaikan dapat dilakukan secara terarah dan berkelanjutan.

h. Relevansi Model Evaluasi dalam Konteks Pendidikan Indonesia

Dalam konteks pendidikan di Indonesia, keberadaan berbagai model evaluasi ini sangat penting untuk mendukung kebijakan pendidikan yang berbasis bukti. Misalnya, evaluasi berbasis CIPP dapat digunakan untuk menilai efektivitas kurikulum Merdeka, sedangkan model Kirkpatrick dapat diterapkan dalam mengevaluasi pelatihan guru dan peningkatan kompetensi. Akan tetapi, perlu dicatat bahwa pelaksanaan evaluasi di lapangan masih menghadapi berbagai tantangan, seperti keterbatasan sumber daya, kurangnya pemahaman evaluator, dan resistensi terhadap hasil evaluasi. Oleh karena itu, pemilihan model evaluasi harus mempertimbangkan tujuan evaluasi, jenis program, serta kondisi nyata di lapangan.¹⁴

i. Diskusi

Hasil penelitian menunjukkan bahwa penerapan model-model evaluasi pendidikan tidak hanya bergantung pada pendekatan teoritis yang digunakan, tetapi juga pada konteks implementasi dan kebutuhan spesifik lembaga pendidikan. Temuan ini sejalan dengan pandangan Stufflebeam (2003) yang menyatakan bahwa model evaluasi harus dipilih berdasarkan relevansi terhadap tujuan dan karakteristik program pendidikan. Misalnya, dalam konteks evaluasi program yang bersifat berkelanjutan dan komprehensif, model CIPP (*Context, Input, Process, Product*) lebih sering digunakan karena mampu memberikan gambaran menyeluruh mengenai pelaksanaan program dari awal hingga hasil akhir. Temuan ini juga didukung oleh studi sebelumnya oleh Arikunto dan Jabar (2010) yang menyatakan bahwa model CIPP sangat cocok digunakan untuk mengevaluasi program berbasis institusi.

Di sisi lain, terdapat kecenderungan bahwa model evaluasi seperti Tyler lebih menekankan pada pencapaian tujuan instruksional, sehingga lebih cocok digunakan dalam evaluasi pembelajaran di kelas. Hal ini menunjukkan bahwa perbedaan orientasi dari tiap model evaluasi perlu dipahami secara kritis oleh para pendidik agar tidak terjadi kekeliruan dalam menentukan model yang sesuai. Temuan ini memperkuat teori yang dikemukakan oleh Worthen dan Sanders (1987) yang menegaskan pentingnya mempertimbangkan tujuan dan hasil akhir pendidikan sebelum memilih model evaluasi yang akan diterapkan.

Lebih lanjut, penggunaan model Stake dan model Countenance yang menitikberatkan pada perspektif respons partisipan terhadap proses pendidikan, juga mencerminkan adanya



pergeseran paradigma evaluasi dari yang bersifat objektif dan terstandar menjadi evaluasi yang lebih kualitatif dan partisipatif. Hal ini menunjukkan bahwa evaluasi pendidikan saat ini tidak hanya dilihat dari angka-angka capaian akademik, tetapi juga dari persepsi dan pengalaman peserta didik serta stakeholder pendidikan lainnya. Perubahan paradigma ini sejalan dengan pendekatan konstruktivistik dalam evaluasi yang menekankan peran aktif individu dalam memberikan makna terhadap proses pembelajaran.

Namun demikian, temuan ini juga menunjukkan bahwa tidak ada satu model evaluasi yang dianggap paling ideal atau superior dibandingkan yang lain. Setiap model memiliki keunggulan dan keterbatasannya masing-masing. Ini selaras dengan pernyataan Scriven (1991) yang menggarisbawahi pentingnya menggunakan pendekatan multimodel dalam evaluasi pendidikan agar hasilnya lebih kaya dan komprehensif. Dalam praktiknya, peneliti menemukan bahwa sebagian lembaga pendidikan bahkan menggabungkan elemen dari beberapa model sekaligus, seperti mengintegrasikan evaluasi proses dari model CIPP dengan evaluasi hasil dari model Tyler, demi memperoleh gambaran yang lebih utuh terhadap efektivitas program pendidikan.

Dari hasil analisis tersebut, dapat disimpulkan bahwa pemilihan model evaluasi pendidikan harus didasarkan pada pertimbangan logis yang mencakup tujuan evaluasi, karakteristik program, serta pendekatan pembelajaran yang digunakan. Peneliti memandang bahwa fleksibilitas dalam penggunaan model menjadi kunci utama agar proses evaluasi tidak bersifat kaku, melainkan adaptif terhadap kebutuhan kontekstual lembaga pendidikan. Selain itu, pendekatan kombinatorik dalam penggunaan model juga layak untuk dikembangkan lebih lanjut sebagai alternatif inovatif dalam evaluasi pendidikan kontemporer.

4. KESIMPULAN

Evaluasi pendidikan merupakan komponen penting dalam sistem pendidikan yang berfungsi untuk menilai keberhasilan proses dan hasil pembelajaran. Melalui evaluasi, pendidik dan pengambil kebijakan dapat mengetahui apakah tujuan pendidikan telah tercapai serta bagaimana proses dan input yang telah digunakan. Dalam pelaksanaannya, evaluasi memerlukan model atau pendekatan yang sistematis agar hasil yang diperoleh valid, akurat, dan bermanfaat bagi peningkatan mutu pendidikan. Berbagai model evaluasi telah dikembangkan untuk menjawab kebutuhan yang beragam, seperti model Tyler yang berorientasi pada tujuan, model CIPP yang menyeluruh dari konteks hingga produk, model Stake yang responsif terhadap *stakeholder*, model Kirkpatrick yang menilai hingga dampak, serta model-model lain seperti *goal-free*, *discrepancy*, dan *countenance*. Masing-masing model memiliki kelebihan dan kekurangan serta karakteristik tersendiri yang harus disesuaikan dengan konteks dan tujuan evaluasi. Pemahaman yang mendalam tentang model-model evaluasi ini akan membantu pendidik, pengelola pendidikan, dan evaluator dalam memilih pendekatan yang tepat guna meningkatkan kualitas pendidikan. Evaluasi bukan hanya tentang pengukuran hasil, tetapi juga sebagai alat refleksi untuk memperbaiki proses pendidikan secara menyeluruh.



5. REFERENSI

- Agung Wibowo, Armanto, D., & Wildansyah Lubis. (2022). Evaluasi Pembelajaran Berbasis Proyek Pada Materi Bangun Ruang Kelas V Sekolah Dasar Dengan Model CIPP. *Journal of Educational Analytics*, 1(1), 27–40. <https://doi.org/10.55927/jeda.v1i1.424>
- Ali, M. K., & Sukardi, S. (2021). Pengembangan Model Evaluasi Pembelajaran Daring di Sekolah Menengah Kejuruan. *JRTI (Jurnal Riset Tindakan Indonesia)*, 6(2), 161. <https://doi.org/10.29210/3003991000>
- Amelia, D., Setiaji, B., Primadewi, K., Habibah, U., Lounggina, T., Peny, L., Rajagukguk, K. P., Nugraha, D., Safitri, W., Wahab, A., Larisu, Z., Setiaji, B., & Dharta, F. Y. (2023). *Metode Penelitian Kuantitatif*. Yayasan Penerbit Muhammad Zaini.
- Aulita, D., Nurazizah, F., Meilinda, L., & Nugraha, D. (2024). Social Media As Source Study Generation Millennials. *Journal Economic and Economic Education*, 1(1), 36–40.
- Bhakti, Y. B. (2017). Evaluasi Program Model CIPP pada Proses Pembelajaran IPA. *JIPFRI (Jurnal Inovasi Pendidikan Fisika Dan Riset Ilmiah)*, 1(2), 75–82. <https://doi.org/10.30599/jipfri.v1i2.109>
- Damanik, S., Zuhdi, M., & Herlina, H. (2020). Model evaluasi pembelajaran AUD berbasis daring di RA Nurun Namirah Medan Marelan (Studi Kasus Selama Masa Pandemi Covid-19). *Al-Fatih: Jurnal Pendidikan Dan Keislaman*, 3(1), 155–172.
- Darodjat, D., & Zuchdi, D. (2016). MODEL EVALUASI PEMBELAJARAN AKIDAH DAN AKHLAK DI MADRASAH TSANAWIYAH (MTs). *Jurnal Penelitian Dan Evaluasi Pendidikan*, 20(1), 11–26. <https://doi.org/10.21831/pep.v20i1.7517>
- Faridah, E. S., Febrianti, R., Purnomo, Hajar, M., Dahlan, M. Z., Gaol, E. L., Maqbuloh, A., Nugraha, D., Nurjanah, Laelasari, E., Sayekti, S. P., & Wijaya, S. (2023). *Pengembangan Kurikulum dan Pembelajaran Abad 21* (1st ed.). Yayasan Penerbit Muhammad Zaini.
- Gaspersz, M., AW, S., & Gaspersz, N. (2023). Model Evaluasi Formatif-Sumatif Terhadap Hasil Belajar Matematika Melalui Pembelajaran Berdiferensiasi Pada Peserta Didik Sma. *Jurnal Magister Pendidikan Matematika (JUMADIKA)*, 5(1), 1–7. <https://doi.org/10.30598/jumadikavol5iss1year2023page1-7>
- Ginanjari, H., Nugraha, D., Noviar, N., & Rahmawati, R. (2023). PENGARUH MOTIVASI BELAJAR DAN LINGKUNGAN KELUARGA TERHADAP HASIL BELAJAR MAHASISWA STKIP PGRI SUKABUMI. 4(1), 22–27.
- Joko, & Nugraha, D. (2023). PENGARUH MOTIVASI BELAJAR DAN LINGKUNGAN KELUARGA TERHADAP HASIL BELAJAR MAHASISWA. *Jurnal Pena Edukasi*, 10(1), 27–34. <https://doi.org/10.31004/basicedu.v7i1.4385>
- Joko, Rachmadio, R. E., & Nugraha, D. (2024). IMPLEMENTASI SISTEM PEMBELAJARAN EFEKTIF SEBAGAI STRATEGI PENGUATAN PROFESIONALISME GURU DALAM ERA DIGITAL. *Indonesian Journal of Community Empowerment*, 6(1)(06), 2657-117X.
- Mandasari, K., N.A, L., & Adiyono, A. (2025). Implementasi Model Evaluasi-Refleksi Siklik Dalam Peningkatan Kualitas Pembelajaran Di Madrasah Aliyah. *JPG: Jurnal Pendidikan Guru*, 6(2), 303–317. <https://doi.org/10.32832/jpg.v6i2.19929>
- Mardiah, M., & Syarifudin, S. (2019). Model-Model Evaluasi Pendidikan. *MITRA ASH-*



- SHIBYAN: *Jurnal Pendidikan Dan Konseling*, 2(1), 38–50.
<https://doi.org/10.46963/mash.v2i1.24>
- Muslimin, D., Alamin, Z., Alizunna, D., Nur Ainia, R., Prakoso, F. A., Missouri, R., Masita, Allo, K. P., Nugraha, D., & Dian, H. (2024). *METODOLOGI PENELITIAN: PENDEKATAN KUANTITATIF, KUALITATIF, DAN CAMPURAN* (T. P. Wahyuni (ed.); 1st ed.). CV LAUK PUYU PRESS.
- Nugraha, D. (2023a). Meniti Sukses Akademis: Peran Fasilitas Sekolah dan Motivasi Prestasi pada Hasil Belajar Mahasiswa. *Jurnal Ilmu Sosial, Politik, Dan Hukum*, 1(1), 9–14.
- Nugraha, D. (2023b). Pengaruh metode simulasi demonstrasi terhadap hasil belajar mahasiswa. *Jurnal Pena Edukasi*, 10(1), 1–8.
- Nugraha, D. (2025). *METODOLOGI PENELITIAN : TEORI DAN PRAKTIK* (N. Mayasari (ed.)). Penerbit Widina Media Utama.
- Rasyid, A. H. A., Yunitasari, B., Susila, I. W., Dewanto, D., Yunus, Y., & Santoso, D. I. (2022). Pengembangan Model Evaluasi Pembelajaran Berbasis Obe. *Jurnal Pendidikan (Teori Dan Praktik)*, 7(1), 8–17. <https://doi.org/10.26740/jp.v7n1.p8-17>
- Saputra, D. I., Abdullah, A. G., & Hakim, D. L. (2014). Pengembangan Model Evaluasi Pembelajaran Project Based Learning Berbasis Logika Fuzzy. *Innovation of Vocational Technology Education*, 9(1), 13–34. <https://doi.org/10.17509/invotec.v9i1.5089>
- Turmuzi, M., Ratnaya, I. G., Al Idrus, S. W., Paraniti, A. A. I., & Nugraha, I. N. B. S. (2022). Literature Review: Evaluasi Keterlaksanaan Kurikulum 2013 Menggunakan Model Evaluasi CIPP (Context, Input, Process, dan Product). *Jurnal Basicedu*, 6(4), 7220–7232. <https://doi.org/10.31004/basicedu.v6i4.3428>
- Wirati, T. W., Vidyastuti, H. A., Utarsih, H., Kurniawan, G. I., Sugiharto, N. A., Hamdani, D., Annisawati, A. A., Mulyana, I., Nugraha, D., Wardhana, M. A., & Persada, A. R. (2024). *Berpikir Kreatif dan Kritis di Era VUCA* (R. Solihin (ed.)). Ekuitas Publisher.