

---

## Menganalisis Pemilihan Model Evaluasi Program Pendidikan

Widiastuti<sup>1</sup>, Linora<sup>2</sup>, Hafiz Hidayat<sup>3)\*</sup>

<sup>1,2,3</sup>Universitas Adzкия

Email : [widistutizakia@gmail.com](mailto:widistutizakia@gmail.com),  
[linoranora017@gmail.com](mailto:linoranora017@gmail.com),  
[hidayathafiz90@adzкисло.ac.id](mailto:hidayathafiz90@adzкия.ac.id).

---

### Abstrak

Tujuan dari penelitian ini adalah untuk menganalisis pemilihan model evaluasi program pendidikan.. Penelitian ini menggunakan pendekatan kualitatif. Hasil penelitian menunjukkan bahwa Model Scriven Scriven merancang dua model evaluasi yaitu: (1) Goal-Free Evaluation Approach, dan (b) Formative and Summative model, 2) Contentance Evaluation Model (Stake Model). Model evaluasi program yang diperkenalkan oleh Stake dikenal dengan model Contentance (keseluruhan). 3) Model Alkin. Batasan konstruk evaluasi sebagai suatu proses penentuan area yang akan di evaluasi. 4) CSE-UCLA Evaluation Model. langkah-langkah dari model CSE-UCLA menjadi empat tahap yaitu: need assessment, program planning, formative evaluation, dan summative evaluation., 5) Discrepancy Model. Langkah-langkah dalam model kesenjangan ini adalah: penyusunan desain, pemasangan instalasi (installation), proses (process), d) Pengukuran tujuan (product), e) perbandingan (programme comparison), 6) Model Brinkerhoff et-al terdiri dari: (1) fixed vs emergent evaluation design, (2) formative vs summative evaluation, dan (3) experimental and quasi experimental design vs natural/unobtrusive inquiry.

**Kata kunci:** Analisis, model, evaluasi, pendidikan.

### Abstract

The aim of this research is to analyze the selection of educational program evaluation models. This research uses a qualitative approach. The research results show that the Scriven Model Scriven designed two evaluation models, namely: (1) Goal-Free Evaluation Approach, and (b) Formative and Summative model, 2) Contentance Evaluation Model (Stake Model). The program evaluation model introduced by Stake is known as the Contentance (overall) model. 3) Alkin Model. Limitations of the evaluation construct as a process of determining the area to be evaluated. 4) CSE-UCLA Evaluation Model. The steps of the CSE-UCLA model are divided into four stages, namely: needs assessment, program planning, formative evaluation, and summative evaluation. 5) Discrepancy Model. The steps in this gap model are: preparation of design, installation, process, d) Measuring the goal (product), e) comparison (program comparison), 6) Brinkerhoff et-al's model consists of: (1) fixed vs emergent evaluation design, (2) formative vs summative evaluation, and (3) experimental and quasi experimental design vs natural/unobtrusive inquiry.

**Keywords:** Analysis, model, evaluation, education

---

## PENDAHULUAN

Evaluasi program dimaksud untuk memperoleh informasi yang akurat dan obyektif tentang suatu program. Informasi tersebut dapat berupa proses pelaksanaan program, dampak/hasil yang dicapai, efisiensi, serta pemanfaatan hasil evaluasi yang difokuskan untuk program itu sendiri, yaitu untuk mengambil keputusan apakah dilanjutkan, diperbaiki, atau dihentikan. Selain itu, juga dipergunakan untuk kepentingan penyusunan program berikutnya maupun penyusunan kebijakan yang terkait dengan program.

Dalam setiap perencanaan dan kegiatan selalu membutuhkan evaluasi, tujuannya adalah untuk mengetahui kegagalan atau keberhasilan sebuah perencanaan atau kegiatan tersebut. Tetapi, tidak semua orang menyadari bahwa setiap kita selalu melakukan pekerjaan evaluasi.

Adapun pengertian evaluasi adalah pengambilan keputusan berdasarkan hasil pengukuran dan standar kriteria. Evaluasi juga diartikan sebagai kegiatan atau proses untuk menilai sesuatu. Sedangkan menurut Marrison evaluasi adalah perbuatan pertimbangan berdasarkan seperangkat kriteria yang disepakati dan dapat dipertanggungjawabkan.

Dalam Pendidikan Islam, evaluasi merupakan salah satu komponen dari sistem pendidikan Islam yang harus dilakukan secara sistematis dan terencana sebagai alat untuk mengukur keberhasilan atau target yang akan dicapai dalam proses pendidikan atau proses pembelajaran.

Dalam ilmu evaluasi program pendidikan, ada banyak model yang bisa digunakan untuk mengevaluasi suatu program. Meskipun antara satu dengan yang lainnya berbeda, namun maksudnya sama yaitu melakukan kegiatan pengumpulan data atau informasi yang berkenaan dengan objek yang dievaluasi, yang tujuannya menyediakan bahan bagi pengambil keputusan dalam menentukan tindak lanjut suatu program. Adapun tujuan dalam penulisan ini adalah untuk mengetahui model-model evaluasi dan pengembangan instrumen evaluasi program pendidikan dasar.

Briekerhoff et-al (1983:2) mendefinisikan evaluasi program adalah suatu proses menemukan sejauh mana tujuan dan sasaran program atau proyek telah terealisasi, memberikan informasi untuk pengambilan keputusan, membandingkan kinerja dengan standar atau patokan untuk mengetahui adanya kesenjangan, penilaian harga dan kualitas dan penyelidikan sistematis tentang nilai atau kualitas suatu objek.

Evaluasi program menurut Tyler adalah proses untuk mengetahui apakah tujuan sudah dapat terealisasi (Arikunto dan Jabar, 2009:5). Menurut Arikunto (2005:291) evaluasi program adalah kegiatan yang dimaksudkan untuk mengetahui seberapa tinggi tingkat keberhasilan dari kegiatan yang direncanakan.

Tujuan evaluasi adalah untuk memperoleh informasi yang akurat dan obyektif tentang suatu program. Informasi itu dapat berupa proses pelaksanaan program, dampak/hasil yang dicapai, efisiensi, serta pemanfaatan hasil evaluasi yang difokuskan untuk program itu sendiri, yaitu untuk mengambil keputusan apakah dilanjutkan, diperbaiki, atau dihentikan. Selain itu, juga dipergunakan untuk kepentingan penyusunan program berikutnya maupun penyusunan kebijakan yang terkait dengan program.

Evaluasi program bertujuan untuk mengetahui pencapaian tujuan program yang telah dilaksanakan. Selanjutnya, hasil evaluasi program dimanfaatkan sebagai dasar untuk melaksanakan kegiatan tindak lanjut atau untuk melakukan pengambilan keputusan berikutnya.

Berdasarkan pemaparan di atas maka dapatlah dimakna bahwa evaluasi program adalah suatu unit atau kesatuan kegiatan yang bertujuan mengumpulkan informasi tentang realisasi atau implementasi dari suatu kebijakan, berlangsung dalam proses yang berkesinambungan, dan terjadi dalam suatu organisasi yang melibatkan sekelompok orang guna pengambilan keputusan.

## METODE PENELITIAN

Dalam penelitian ini menggunakan pendekatan kualitatif (Sugiyono, 2008). Sedangkan metode deskriptif yang dipakai dalam analisis datanya. (Moleong, J, Lexy, 2006: 248). Adapun tujuan penelitian ini yaitu untuk menyampaikan serta membagikan poin-poin yang mempunyai kaitan dengan kejadian yang diteliti, dan juga memaparkan problem yang ditemukan. (Margono, 2002).

## HASIL DAN PEMBAHASAN

### 1. Model Scriven

Scriven merancang dua model evaluasi yaitu: (1) *Goal-Free Evaluation Approach*, dan (2) *Formative and Summative model*.

#### a. *Goal-Free Evaluation Approach*

Scriven mengemukakan bahwa dalam melaksanakan evaluasi program, evaluator tidak perlu memperhatikan apa yang menjadi tujuan program. Yang perlu diperhatikan dalam program tersebut adalah bagaimanakerjanya (kinerja) suatu program, dengan jalan mengidentifikasi penampilanpenampilanyang terjadi (pengaruh) baik hal-hal yang positif (yaitu hal yang diharapkan) maupun hal-hal yang negatif (yang tidak diharapkan).

Evaluasi model *goal free evaluation*, fokus pada adanya perubahan perilaku yang terjadi sebagai dampak dari program yang diimplementasikan, melihat dampak sampingan baik yang diharapkan maupun yang tidakdiharapkan, dan membandingkan dengan sebelum program dilakukan. Evaluasi juga membandingkan antara hasil yang dicapai dengan besarnya biaya yang dikeluarkan untuk program tersebut atau melakukan *costbenefit analysis*.

Model *goal free evaluation* merupakan titik evaluasi program, di mana objek yang dievaluasi tidak perlu terkait dengan tujuan dari objek atau subjek tersebut, tetapi langsung kepada implikasi keberadaan program apakah bermanfaat atau tidak objek tersebut atas dasar penilaian kebutuhanyang ada.

Tujuan program tidak perlu diperhatikan karena kemungkinan evaluator terlalu rinci mengamati tiap-tiap tujuan khusus. Jika masing-masing tujuan khusus tercapai, artinya terpenuhi dalam penampilan tetapi evaluator lupa memperhatikan sejauh mana masing-masing penampilan tersebut mendukung penampilan terakhir yang diharapkan oleh tujuan umum maka akibatnya jumlah penampilan khusus ini tidak banyak bermanfaat. Dapat disimpulkan bahwa, dalam model ini bukan berarti lepas dari tujuan tetapi hanya lepas dari tujuan khusus. Model ini hanya mempertimbangkan tujuan umum yang akan dicapai oleh program, bukan secara rinci perkomponen yang ada.

Scriven menekankan bahwa evaluasi itu adalah interpretasi *judgement* ataupun *explanation* dan evaluator yang merupakan pengambil keputusan sekaligus penyedia informasi. Ciri-ciri evaluasi bebas tujuan yaitu: a) Evaluator sengaja menghindari untuk mengetahui tujuan program, b) Tujuan yang telah dirumuskan terlebih dahulu tidak dibenarkan menyempitkan fokus evaluasi, c) Evaluasi bebas tujuan berfokus pada hasil yang sebenarnya, bukan pada hasil yang direncanakan, d) Hubungan evaluator dan manajer atau dengan karyawan proyek dibuat seminimal mungkin, e) Evaluasi menambah kemungkinan ditemukannya dampak yang tidak diramalkan (Tayib naps, 2000:35).

Mungkin akan lebih baik apabila evaluasi yang berorientasi pada tujuan dan evaluasi bebas tujuan dipadukan, karena akan saling mengisi dan melengkapi. Evaluator internal biasanya melakukan evaluasi yang berorientasi pada tujuan, karena ia sulit menghindari atau mau tidak mau ia akan mengetahui tujuan program, akan tidak pantas apabila ia tidak acuh. Manajer program jelas ingin mengetahui sampai seberapa jauh program telah dicapai, dan evaluator internal akan dan harus menyediakan informasi untuk manajernya.

Di samping itu, perlu diketahui bagaimana orang luar menilai program bukan hanya untuk mengetahui apa yang seharusnya dilakukan, apa yang dilakukan di semua bagian, pada semua yang telah dihasilkan, secara sengaja atau tidak sengaja. Yang belakangan ini merupakan tugas evaluator bebas tujuan yang tidak mengetahui tujuan program. Jadi, evaluasi yang berorientasi pada tujuan dan evaluasi bebas tujuan dapat bekerja sama dengan baik. Scriven menjelaskan fokus pada program atau tujuan kegiatan ini dapat menjadi tempat awal yang

penting untuk evaluator bekerja dalam domain evaluasi. Tujuan program tertentu tidak harus diambil sebagai yang diberikan, tetapi diperiksa dan dievaluasi juga.

Model evaluasi *Goal Free Evaluation* ini mempunyai kekurangan dan kelebihan. Kelebihan dari model bebas tujuan di antaranya adalah: a) Evaluator tidak perlu memperhatikan secara rinci setiap komponen, tetapi hanya menekankan pada bagaimana mengurangi prasangka (bias), b) Model ini menganggap pengguna sebagai audiens utama. Melalui model ini, Scriven ingin evaluator mengukur kesan yang didapat dari sesuatu program dibandingkan dengan kebutuhan pengguna dan tidak membandingkannya dengan pihak penganjur, c) Pengaruh konsep pada masyarakat, bahwa tanpa mengetahui tujuan dari kegiatan yang telah dilakukan, seorang penilai bisa melakukan evaluasi, d) mendorong pertimbangan setiap kemungkinan pengaruh tidak saja yang direncanakan, tetapi juga dapat diperhatikan sampingan lain yang muncul dari produk.

Selanjutnya terkait dengan kelemahan model ini adalah: a) Model bebas tujuan ini pada umumnya bebas menjawab pertanyaan penting, seperti apa pengaruh yang telah diperhitungkan dalam suatu peristiwa dan bagaimana mengidentifikasi pengaruh tersebut, b) Pada tingkatan praktis model ini tidak terlalu berhasil dalam menggambarkan bagaimana evaluasi sebaiknya benar-benar dilaksanakan, c) Tidak merekomendasikan bagaimana menghasilkan penilaian kebutuhan walau pada akhirnya mengarah pada penilaian kebutuhan, d) Diperlukan evaluator yang benar-benar kompeten untuk dapat melaksanakan evaluasi model ini.

## **2. *Contentance Evaluation Model (Stake Model)***

Model evaluasi program yang diperkenalkan oleh Stake dikenal dengan model *Countenance* (keseluruhan). Model ini juga disebut model evaluasi pertimbangan. Maksudnya evaluator mempertimbangkan program dengan membandingkan kondisi hasil evaluasi program dengan yang terjadi di program lain, dengan objek sasaran yang sama dan membandingkan kondisi hasil pelaksanaan program dengan standar yang ditentukan oleh program tersebut.

Tujuan dari model *Countenance Stake* adalah melengkapi kerangka untuk pengembangan suatu rencana penilaian kurikulum. Perhatian utama Stake adalah hubungan antara tujuan penilaian dengan keputusan berikutnya berdasarkan sifat data yang dikumpulkan. Hal tersebut, karena Stake melihat adanya ketidaksesuaian antara harapan penilai dan guru.

Dalam model ini Stake menekankan peran evaluator dalam mengembangkan tujuan kurikulum menjadi tujuan khusus yang terukur. Model *countenance* terdiri atas dua matriks yaitu *description* (gambaran) dan *judgement* (pertimbangan). Matriks pertimbangan baru dapat dikerjakan oleh evaluator setelah matriks deskripsi diselesaikan. Matriks Deskripsi terdiri atas kategori rencana (*intent*) dan observasi. Matriks Pertimbangan terdiri atas kategori standar dan pertimbangan. Pada setiap kategori terdapat tiga fokus yaitu:

- a. *Antecedents* yaitu sebuah kondisi yang ada sebelum instruksi yang mungkin berhubungan dengan hasil, contohnya: latar belakang guru, kurikulum yang sesuai, ketersediaan sumber daya.
- b. *Transaction* yaitu pertemuan dinamis yang merupakan proses instruksi (kegiatan, proses, dan lainnya), contohnya: interaksi guru dan siswa.
- c. *Outcomes* yaitu efek dari pengalaman pembelajaran (pengamatan dan hasil tenaga kerja), contohnya *performance* guru, peningkatan kinerja.

Model evaluasi Stake dapat membawa dampak yang cukup besar dalam penilaian, dan merupakan konsep yang cukup kuat untuk perkembangannya yang lebih jauh dalam bidang evaluasi. Dalam model ini, evaluasi dilakukan dengan membandingkan antara satu program dengan program lain yang dianggap standar. Stake berpendapat menilai suatu program pendidikan harus melakukan perbandingan yang relatif antara program satu dan program yang lain, atau perbandingan yang absolut yaitu membandingkan suatu program dengan standar tertentu (Tayibnaps, 2000:19).

Kelebihan evaluasi model Stake adalah:

- a. Dalam evaluasi memasukkan data tentang latar belakang program, proses dan hasil yang merupakan perluasan ruang lingkup evaluasi.
- b. Evaluator memegang kendali dalam evaluasi dan juga memutuskan cara yang paling tepat untuk hadir dan menggambarkan hasil.
- c. Memiliki potensi besar untuk memperoleh wawasan baru dan teori tentang lapangan dan program yang akan di evaluasi.

Selanjutnya terkait dengan kelemahan evaluasi model Stake ini adalah:

- a. Pendekatan yang dilakukan terlalu subjektif.
- b. Terjadinya kemungkinan dalam meminimalkan pentingnya instrumen pengumpulan data dan evaluasi kuantitatif.
- c. Kemungkinan biaya yang terlalu besar.

### 3. Model Alkin

Dalam merumuskan model evaluasi program yang disusunnya, Alkin membuat batasan konstruk evaluasi sebagai suatu proses penentuan area yang akan di evaluasi, pemilihan informasi yang cocok untuk dievaluasi, pengumpulan dan analisis informasi serta penyusunan laporan atau ringkasan data yang berguna bagi pengambil keputusan dalam memilih alternatif yang berguna yang tepat dari berbagai alternatif yang ada.

Model Alkin dikembangkan berdasarkan 4 (empat) asumsi. Apabila keempat asumsi ini sudah dipenuhi maka model ini dapat digunakan. Adapun keempat asumsi itu yaitu:

- a. Variabel perantara adalah satu-satunya variabel yang dimanipulasi.
- b. Sistem eksternal tidak langsung dipengaruhi oleh keluaran sistem (persekolahan).
- c. Para pengambil keputusan sekolah tidak memiliki kontrol mengenai pengaruh yang diberikan sistem eksternal terhadap sekolah.
- d. Faktor masukan mempengaruhi aktivitas faktor perantara dan pada gilirannya faktor perantara berpengaruh terhadap hasil.

Adapun yang dimaksud dengan proses di sini meliputi faktor perantara yang merupakan kelompok variabel yang secara langsung mempengaruhi keluaran. Adapun yang masuk dalam variabel perantara ini diantaranya adalah ratio jumlah guru dengan siswa, jumlah siswa dalam kelas, pengaturan administrasi, penyediaan buku bacaan, prosedur pengajaran dan sebagainya. Adapun keluaran siswa adalah setiap perubahan yang terjadi pada diri siswa sebagai akibat dari pengalaman belajar yang diperolehnya. Perubahan ini harus diikuti sejak siswa masuk sistem hingga keluar sistem. Perubahan harus diukur meliputi setiap aspek perubahan yang mungkin terjadi termasuk di dalamnya kemampuan siswa dalam melanjutkan pelajaran di

tingkat pendidikan yang lebih tinggi, pada waktu memasuki lapangan kerja, dalam melakukan pekerjaan bahkan termasuk aktivitas dalam kehidupan di masyarakat.

Kelemahan dari model Alkin adalah keterbatasannya dalam fokus kajian yaitu yang hanya fokus pada kegiatan persekolahan. Sehingga model ini hanya dapat digunakan untuk mengevaluasi kurikulum yang sudah siap dilaksanakan di sekolah.

#### 4. CSE-UCLA Evaluation Model

CSE-UCLA adalah akronim dari *Center for the Study of Evaluation-University of California in Los Angeles*. Pada awalnya, karakteristik dari model CSE-UCLA adalah adanya 5 (lima) tahap yang dilakukan dalam evaluasi yaitu: perencanaan, pengembangan, implementasi, hasil dan dampak. Seiring dengan perkembangannya, Fernandes sebagaimana dikutip Arikunto dan Jabar (2009:44) memaparkan bahwa langkah-langkah dari model CSE-UCLA menjadi empat tahap yaitu:

a. *Need assessment.*

Pada tahap pertama ini yaitu analisis kebutuhan, evaluator memusatkan perhatian pada penentuan masalah pertanyaan yang dapat diajukan yaitu:

- 1) Hal-hal apakah yang perlu dipertimbangkan sehubungan dengan keberadaan program?
- 2) Kebutuhan apakah yang terpenuhi sehubungan dengan adanya pelaksanaan program ini?
- 3) Tujuan jangka panjang apakah yang dapat dicapai melalui program ini?

b. *Program planning.*

Pada tahap kedua ini yaitu perencanaan program, evaluator mengumpulkan data yang terkait langsung dengan program dan mengarahkan pada pemenuhan kebutuhan yang telah diidentifikasi pada tahap pertama. Dalam tahap perencanaan ini program yang di evaluasi dengan cermat untuk mengetahui apakah rencana program yang telah disusun berdasarkan analisis kebutuhan. Evaluasi tahap ini tidak lepas dari tujuan yang telah dirumuskan sebelumnya.

c. *Formative evaluation.*

Dalam tahap ketiga ini yaitu evaluasi formatif, evaluator memusatkan perhatian pada keterlaksanaan program. Dengan demikian, evaluator diharapkan betul-betul terlibat dalam program karena harus mengumpulkan data dan berbagai informasi dari pengembang program.

d. *Summative evaluation.*

Dalam tahap keempat yaitu evaluasi sumatif, evaluator diharapkan dapat mengumpulkan semua data tentang hasil dan dampak dari program. Melalui evaluasi sumatif ini diharapkan dapat diketahui apakah tujuan yang dirumuskan untuk program sudah tercapai dan jika belum dicari bagian mana yang dan apa faktor-faktor penyebabnya.

#### 5. Discrepancy Model

*Discrepancy model* atau model kesenjangan digagas oleh Malcolm Probus yaitu evaluasi yang dilakukan dengan maksud untuk mengetahui tingkat kesesuaian antara standar yang telah ditentukan dalam program dengan penampilan aktual dari program tersebut. Standar adalah kriteria yang telah dikembangkan dan ditetapkan dengan hasil yang efektif, sedangkan penampilan adalah sumber, prosedur, manajemen, dan hasil nyata yang tampak ketika program dilaksanakan.

Langkah-langkah dalam model kesenjangan ini adalah:

a. Penyusunan desain.

Dalam tahap ini dilakukan kegiatan sebagai berikut:

- 1) Merumuskan tujuan program.
- 2) Menyiapkan klien, staf dan kelengkapan lain.
- 3) Merumuskan standar dalam bentuk rumusan yang menunjukpada sesuatu yang dapat diukur. Biasanya di dalam langkah inievaluator berkonsultasi dengan pengembangan program.

b. Pemasangan instalasi (*installation*).

Tahap ini melihat apakah kelengkapan yang tersedia sudah sesuaidengan yang diperlukan atau belum. Dalam tahap ini dilakukan kegiatan sebagai berikut:

- 1) Meninjau kembali penetapan standar.
- 2) Meninjau program yang sedang berjalan.
- 3) Meneliti kesenjangan antara yang direncanakan dengan yang sudah dicapai.

c. Proses (*process*).

Dalam tahap ini adalah mengadakan penilaian tujuan-tujuan manakahyang sudah dicapai. Dalam hal ini adalah mengumpulkan data dari pelaksanaan program.

d. Pengukuran tujuan (*product*)

Tahap ini adalah mengadakan analisis data dan menetapkan tingkat *output* yang diperoleh. Pertanyaan yang diajukan dalam tahap ini adalah: apakah program sudah mencapai tujuan terminalnya?

e. Perbandingan (*programme comparison*).

Tahap ini adalah membandingkan hasil yang telah dicapai dengantujuan yang telah ditetapkan. Dalam tahap ini evaluator menuliskan semua penemuan tentang kesenjangan untuk disajikan kepada pengambil keputusan, agar dapat memutuskan kelanjutan dari program tersebut.

## 6. Model Brinkerhoff

Desain evaluasi program pada umumnya terdiri atas elemen-elemen yang sama, ada banyak cara untuk menggabungkan elemen tersebut, masing-masing ahli atau evaluator mempunyai konsep yang berbeda dalam hal ini Brinkerhoff et-al mengemukakan tiga golongan evaluasi yang disusun berdasarkan penggabungan elemen-elemen yang sama seperti model lainnya, namun dalam komposisi dan versi sendiri.

Model Brinkerhoff et-al terdiri dari: (1) *fixed vs emergent evaluation design*, (2) *formative vs summative evaluation*, dan (3) *experimental and quasi experimental design vs natural/unobtrusive inquiry* (Tayib naps, 2000:16).

a. *Fixed vs emergent evaluation design*.

Desain evaluasi yang tetap (*fixed*) ditentukan dan direncanakan secara sistematis sebelum implementasi dikerjakan. Desain dikembangkan berdasarkan tujuan program disertai seperangkat pertanyaan yang akan dijawab oleh informasi yang akan diperoleh dari sumber-sumber tertentu.

Rencana analisis dibuat sebelumnya di mana si pemakai akan menerima informasi seperti yang telah ditentukan dalam tujuan. Desain biasanya dibicarakan dan dirundingkan dengan pemakai

utama atau pemesan, apabila ada perubahan biasanya hanya lebih memperlancar pencapaian tujuan dan rencana utama.

Walaupun desain *fixed* ini lebih terstruktur daripada desain *emergent*, desain *fixed* juga dapat disesuaikan dengan kebutuhan yang mungkin berubah. Kebanyakan evaluasi formal yang dibuat secara individu berdasarkan desain *fixed*, karena tujuan program telah ditentukan dengan jelas sebelumnya, dibiayai dan melalui usulan atau proposal evaluasi.

Desain *fixed* relatif memakai biaya yang cukup besar, kegiatan-kegiatan berkisar membuat pertanyaan-pertanyaan menyiapkan dan membuat instrumen, menganalisis hasil evaluasi dan melaporkan secara formal hasil evaluasi kepada pemakai. Komunikasi antara evaluator dan klien dilakukan secara teratur, formal atau tertulis.

Evaluator berpedoman pada tujuan program untuk merumuskan masalah atau pertanyaan-pertanyaan desain dan menstimulasi klien yang relevan untuk mengembangkan dan membetulkan pertanyaan tersebut. Strategi pengumpulan informasi khususnya menggunakan cara-cara formal juga metode penelitian, kriteria penelitian dan pengumpulan data biasanya dilakukan secara kuantitatif.

Selanjutnya desain evaluasi *emergent*, dibuat untuk beradaptasi dengan pengaruh dan situasi yang sedang berlangsung dan berkembang seperti menampung pendapat-pendapat klien, masalah-masalah, kegiatan program. Evaluasi ini menghabiskan banyak waktu dari permulaan sampai akhirnya mencari tujuan dan isu, karena semuanya pada dasarnya tidak dikhususkan dan ditentukan sebelumnya.

Biaya relatif lebih besar, sumber-sumber dalam desain ditentukan untuk mengamati program dan memfokuskan pertanyaan lebih lanjut. Evaluator tidak mendorong klien memikirkan tentang program atau isu evaluasi. Klien menentukan isu-isu penting dan informasi yang diperlukan desain. Komunikasi antar evaluator dengan klien terus berkesinambungan selama proses evaluasi.

Metode yang digunakan dalam desain ini adalah observasi, studi kasus dan laporan tim penyokong. Pengukuran yang tidak selalu berpedoman pada tujuan biasanya dilakukan dan evaluator sering kali mengorbankan ketepatan pengukuran untuk lebih berguna. Informasi yang bersifat kualitatif biasanya dikumpulkan. Desain terus berkembang, berubah dan bereaksi sesuai dengan situasi dan kondisi yang dapat dikatakan tak pernah berhenti.

#### **b. *Formative vs summative evaluation.***

Evaluasi formatif digunakan untuk memperoleh informasi yang dapat membantu memperbaiki program. Dibuat untuk karyawan, dan juga mengorbankan kepentingan orang luar untuk lebih bermanfaat bagi program. Ada yang mengatakan bahwa evaluasi yang paling melindungi yaitu evaluasi formatif.

Fokus evaluasi formatif berkisar pada kebutuhan yang dirumuskan oleh karyawan atau orang-orang yang terlibat dalam program. Evaluator sering merupakan bagian dari program dan bekerja sama dengan orang-orang program. Strategi pengumpulan informasi mungkin juga dipakai, tetapi penekanan pada usaha memberikan informasi yang berguna secepatnya bagi perbaikan program. Desain evaluasi dibuat bersama orang-orang program dan direvisi untuk mencapai kebutuhan mereka.

Selanjutnya evaluasi sumatif dibuat untuk menilai kegunaan suatu objek yang sering diminati atau dibiayai oleh pemakai/pemesan atau oleh sponsor. Evaluasi sumatif digunakan untuk menilai apakah suatu program akan diteruskan atau dihentikan saja. Evaluator harus dapat dipercaya oleh sejumlah klien yang akan dipengaruhi oleh keputusan tersebut. Untuk usaha-usaha yang dibiayai perorangan, evaluasi sumatif ini lebih populer daripada evaluasi formatif.

Pada evaluasi sumatif, evaluasi berfokus pada variabel-variabel yang dianggap penting oleh sponsor atau pembuat keputusan (pengambil kebijakan). Evaluator luar atau tim *review* sering dipakai, karena evaluator internal dapat mempunyai minat yang berbeda. Strategi pengumpulan informasi akan memaksimalkan validitas eksternal dan internal yang mungkin dikumpulkan dalam waktu yang cukup lama.

**c. *Experimental and quasi experimental design vs natural/unobtrusive inquiry.***

Desain eksperimen dan quasi eksperimen menggunakan metodologi penelitian yang subjek penelitiannya di *random* (acak), perlakuan diberikan dan pengukuran dampak dilakukan. Tujuan dari penelitian yaitu untuk menilai manfaat suatu objek, suatu program atau strategi yang dicobakan. Apabila subjek atau program dipilih secara acak, maka generalisasi dapat populasi.

Dalam beberapa hal, intervensi tidak mungkin dilakukan atau tidak dikehendaki. Apabila proses sudah terjadi, evaluator harus melihat dokumen-dokumen sejarah, mempelajari nilai tes, atau menganalisis penelitian yang dilakukan. Bila dianggap penting untuk mengevaluasi suatu lingkungan atau program agar dapat diperbaiki, evaluator mungkin memilih mengamati, bicara dengan orang-orang yang terlibat, dan selalu merendah (*low profile*) sehingga program yang dievaluasi tidak terancam dan mengubah diri karena kehadiran evaluator. Banyak metodologi termasuk observasi, survey, analisis meta evaluasi, studi kasus dan wawancara dapat dilakukan seperti itu untuk mengurangi dampak evaluasi pada orang dan proyek dan memaksimalkan laporan yang sebenarnya.

Desain penelitian ini menggunakan waktu dan biaya yang relatif banyak yang digunakan untuk mempersiapkan instrumen untuk menilai perlakuan, data kuantitatif biasanya dikumpulkan, dan kriteria statistik juga digunakan. Kriteria statistik berfokus pada hasil program dan dibuat sebelumnya. Interaksi dengan klien dilakukan untuk membuat rencana, mengumpulkan informasi dan melaporkan kembali.

Strategi pengumpulan data terutama menggunakan instrumen formal seperti tes, survey, kuesioner dan *skala rating* serta memakai metode penelitian yang standar. Kriteria penelitian seperti *internal* dan *external validity* dianggap penting. Data yang dikumpulkan kebanyakan kuantitatif. Desain penelitian biasanya dibuat bersama pemesan atau pemakai. Bila ada perubahan hanya untuk memperlancar pencapaian tujuan sesuai rencana.

Selanjutnya penelitian *natural inquiry*, evaluator menghabiskan banyak waktu untuk mengamati dan berbicara dengan klien yang relevan. Strategi yang *multiple* dan sumber-sumber dipakai untuk mempertinggi reliabilitas pengumpulan data. Evaluator merundingkan isu dengan klien, hal ini dilakukan sesuai dengan cara evaluator. Interaksi dengan klien berkesinambungan dan informal. Observasi, studi kasus, laporan tim penyokong, merupakan ciri-ciri desain *natural inquiry*.

## KESIMPULAN

Berdasarkan pembahasan diatas, kami mengambil kesimpulan-kesimpulan sebagai berikut :

Model-model evaluasi pendidikan adalah contoh atau pola acuan suatu proses penilaian dalam mengumpulkan dan menganalisis untuk menentukan taraf kemajuan suatu aktivitas di dalam pendidikan guna menetapkan pencapaian suatu tujuan baik untuk pendidik dan peserta didik.

Model Scriven Scriven merancang dua model evaluasi yaitu: (1) Goal-Free Evaluation Approach, dan (b) Formative and Summative model, 2) Contentance Evaluation Model (Stake Model). Model evaluasi program yang diperkenalkan oleh Stake dikenal dengan model Countenance (keseluruhan). 3) Model Alkin. Batasan konstruk evaluasi sebagai suatu proses penentuan area yang akan di evaluasi. 4) CSE-UCLA Evaluation Model. langkah-langkah dari model CSE-UCLA menjadi empat tahap yaitu: need assessment, program planning, formative evaluation, dan summative evaluation., 5) Discrepancy Model. Langkah-langkah dalam model kesenjangan ini adalah: penyusunan desain, pemasangan instalasi (installation), proses (process)., d) Pengukuran tujuan (product), e) perbandingan (programme comparison), 6) Model Brinkerhoff et-al terdiri dari: (1) fixed vs emergent evaluation design, (2) formative vs summative evaluation, dan (3) experimental and quasi experimental design vs natural/unobtrusive inquiry

## REFERENSI

- Alkin, M.C. (1985). *Guide For Evaluation Decision Makers*. Beverly Hills: Sage Publication.
- Arikunto, S. (1988). *Penilaian Program Pendidikan*. Jakarta: Bina Aksara.
- Arikunto, S. (2005). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Arikunto, S., dan Jabar, C.S.A. (2009). *Evaluasi Program Pendidikan. Pedoman Teoretis Praktis Bagi Mahasiswa Dan Praktisi Pendidikan*. Jakarta: Bumi Aksara.
- Brikerhoff, R.O., et-al. (1983). *Program Evaluation. A Source Book*. Boston: Kluwer Nijboff Publishing.
- Brikerhoff, R.O., et-al. (1986). *Program Evaluation, A Practitioner's Guide for Trainer and Educationer*, 4th Edition. Boston: Kluwer Nijboff Publishing.
- Briggs, L. (1977). *Instructional Design, Principles and Applications*. New Jersey: Educational Technology Publications.
- Djaali dan Muljono, P. (2004). *Pengukuran Dalam Bidang Pendidikan*. Jakarta: Universitas Negeri Jakarta.
- Djarwanto, PS., dan Subagya, P. (1998) *Statistik Induktif*. Yogyakarta: BPFE.
- Faisal, S. (1999). *Penelitian Kualitatif. Dasar-Dasar dan Aplikasi*. Malang: YA3.
- Fitzpatrick, J.L., Sanders, J.R., dan Worthen, B.R. (2004). *Program Evaluation: Alternative Approaches and Practical Guidelines*. Boston: Pearson Education Inc.

- Feuerstein, M.T. (1990). *Partners in Evaluation*. Alihbahasa: Farid Wadjidi. *Evaluasi Partisipatoris*. Jakarta: P3M.
- Guba, E.G., dan Lincoln, Y.S. (1985). *Effective Evaluation*. San Fransisco: Jossey Bass Publishing.
- Hardjodipuro, S. (1982). *System Planning*. Jakarta: Erlangga.
- Isaac, S., dan Michael, W.B., (1987). *Handbook in Research and Evaluation for Education and the Behavioral Science, 2nd Edition*. California: Edits Publishers.
- Kaufman, R., dan English, F.W. (1979) *Need Assessment, Concept And Application*. New Jersey: Educational Technology Publication, Englewood Cliffs.
- Kifer, E. dalam Anglin, G. (1995) *Instructional Technology, Past, Present, Future, 2nd Edition*, Englewood: Colorado Libraries Unlimited Inc.
- Kirkpatrick, J.L. (1998). *Evaluating Training Program, The Four Levels, 2nd Edition*, San Fransisco: Berret-Kohler Publisher, Inc.
- Madaus, G.F., Sriven, M.S., dan Stufflebeam. (1987). *Evaluation Models Viewpoints on Educational and Human Services Evaluation*. Boston: Kluwer Nijboff Publishing.
- Mehrens, W.A dan Lehmann, I.J. (1978). *Measurement and Evaluation in Education and Psychology. 2nd* New York: Holt, Rinehart and Winston.
- Mutrofin, (2010). *Evaluasi Program, Teks Pilihan Untuk Pemula*. Yogyakarta: Lakesbang Pressindo.
- Oliva, P. F. (1992). *Developing the Curriculum*. New York: Harper Collins Publishers.
- Purwanto dan Suparman, A. (1999). *Evaluasi Program Diklat*. Jakarta: STIA-LAN Press.
- Purwanto, M.N. (2001). *Prinsip-Prinsip Dan Teknik Evaluasi Pembelajaran*. Bandung: Remaja Rosdakarya.
- Riduwan. (2002). *Skala Pengukuran Variabel-Variabel Penelitian*. Bandung: Alfabeta.
- Sagala, S. 2010. *Manajemen Strategik Dalam Peningkatan Mutu Pendidikan*. Bandung: Alfabeta.
- Stark, J. S., dan Thomas, A. (1994). *Assessment and Program Evaluation*. Needham Heights: Simon & Schuster Custom Publishing.
- Stufflebeam, D. L. (2003). *The CIPP Model for Evaluation: the Article Presented at the 2003 Annual Conference of the Oregon Program Evaluators Network. (OPEN) 3 October 2003* (online). (<http://www.wmich.edu>).
- Stufflebeam, D.L dan Shinkfield, A.J. (2007). *Evaluation Theory, Models and Application*. San Francisco: Jossey Bass.
- Sudidjono, A. (2000). *Pengantar Statistik Pendidikan*, Jakarta: Raja Grafindo Persada.

- Sudjana, D. (2008). *Evaluasi Program Pendidikan Luar Sekolah*. Bandung: Remaja Rosdakarya.
- Sudjana, N., dan Ibrahim. (2004). *Penelitian dan Penilaian Pendidikan*. Bandung: Sinar Baru Algesindo.
- Sugiyono. (2004). *Statistika Untuk Penelitian*. Bandung: Alfabeta.
- Sugiyono. (2010). *Metode Penelitian Pendidikan, Pendekatan Kuantitatif, Kualitatif dan R&D*. Bandung: Alfabeta.
- Suherman, E. dan Sukjaya, Y. (1990). *Petunjuk Praktis Untuk Melakukan Evaluasi Pendidikan Matematika*. Bandung: Wijayakusumah.
- Sukmadinata, N.S. (2006). *Metode Penelitian Pendidikan*. Bandung: Remaja. Rosdakarya.
- Supardi U.S. (2013). *Aplikasi Statistika Dalam Penelitian Konsep Statistika Yang Lebih Komprehensif*. Jakarta: Change Publication.
- Tayibnapi, F.Y. (2000). *Evaluasi Program*. Jakarta: Rineka Cipta.
- Taylor, P. et-al. (1996). *Planning a Program Evaluation*. Wisconsin: University of Wisconsin-Extension.
- Weiss, C.H. (1972), *Evaluation Research, Methods for Assesing Program Effectivenesss*. Englewood: Prentice Hall.
- Worthern, B.R., dan Sander, J.R. dalam N.L. Smith (1996). *Evaluation Models and Approaches*. Japan: Pergamon Elsevies Science Ltd.