

# **MODEL-MODEL EVALUASI PROGRAM**

**M A K A L A H**

**DISUSUN OLEH :**  
**DRS.ZAINAL ARIFIN, M.PD**  
**NIP.19610501.1986011003**

**JURUSAN KURIKULUM DAN TEKNOLOGI PENDIDIKAN**  
**FAKULTAS ILMU PENDIDIKAN**  
**UNIVERSITAS PENDIDIKAN INDONESIA**  
**B A N D U N G**  
**2010**

## **EVALUASI PROGRAM : MODEL-MODEL EVALUASI**

**By : Zainal Arifin**

Perkembangan model evaluasi termasuk suatu fenomena yang menarik. Setelah Tyler mengemukakan model *black box* tahun 1949, belum terlihat ada model lain yang muncul ke permukaan. Lebih kurang 10 tahun lamanya, orang-orang yang melakukan kegiatan evaluasi hanya menggunakan model evaluasi tersebut. Hal ini mungkin disebabkan evaluasi belum menjadi studi tersendiri. Ketika itu, orang banyak mempelajari evaluasi dari psikometrik dengan kajian utamanya adalah tes dan pengukuran. Evaluasi lebih banyak diarahkan kepada dimensi hasil, belum masuk ke dimensi-dimensi lainnya. Oleh sebab itu, janganlah heran bila evaluasi banyak dilakukan oleh orang-orang yang “terbentuk” dalam tes dan pengukuran. Studi tentang evaluasi belum begitu menarik perhatian orang banyak, karena kurang memiliki nilai praktis. Baru sekitar tahun 1960-an studi evaluasi mulai berdiri sendiri menjadi salah satu program studi di perguruan tinggi, tidak hanya di jenjang sarjana (S.1) dan magister (S.2) tetapi juga pada jenjang doktor (S.3).

Selanjutnya, sekitar tahun 1972, model evaluasi mulai berkembang. Taylor dan Cowley, misalnya, berhasil mengumpulkan berbagai pemikiran tentang model evaluasi dan menerbitkannya dalam suatu buku. Model evaluasi yang dikembangkan lebih banyak menggunakan pendekatan positivisme yang berakar pada teori psikometrik. Dalam model tersebut, nuansa tes dan pengukuran masih sangat kental, sekalipun tidak lagi diidentikkan dengan evaluasi. Penggunaan desain eksperimen seperti yang dikemukakan Campbell dan Stanley (1963) menjadi ciri utama dari model evaluasi. Berkembangnya model evaluasi pada tahun 70-an tersebut diawali dengan adanya pandangan alternatif dari para *expert*. Pandangan alternatif yang dilandasi sebuah paradigma fenomenologi banyak menampilkan model evaluasi.

Dari sekian banyak model-model evaluasi yang dikemukakan, tes dan pengukuran tidak lagi menempati posisi yang menentukan. Penggunaannya hanya untuk tujuan-tujuan tertentu saja, bukan lagi menjadi suatu keharusan, seperti ketika model pertama ditampilkan. Tes dan pengukuran tidak lagi menjadi parameter kualitas suatu studi evaluasi yang dilakukan. Perkembangan lain yang menarik dalam model evaluasi ini adalah adanya suatu upaya untuk bersikap eklektik dalam penggunaan pendekatan positivisme maupun fenomenologi yang oleh Patton (1980)

disebut *paradigm of choice*. Walaupun usaha ini tidak melahirkan model dalam pengertian terbatas tetapi memberikan alternatif baru dalam melakukan evaluasi.

Dalam studi tentang evaluasi, banyak sekali dijumpai model-model evaluasi dengan format atau sistematika yang berbeda, sekalipun dalam beberapa model ada juga yang sama. Misalnya saja, Said Hamid Hasan (1988 : 83 – 136) mengelompokkan model evaluasi sebagai berikut :

1. Model evaluasi kuantitatif, yang meliputi : model *Tyler*, model teoritik Taylor dan Maguire, model pendekatan sistem Alkin, model *Countenance Stake*, model CIPP, model ekonomi mikro.
2. Model evaluasi kualitatif, yang meliputi : model studi kasus, model iluminatif, dan model responsif.

Sementara itu, Kaufman dan Thomas dalam Suharsimi Arikunto dan Cepi Safruddin AJ (2007 : 24) membedakan model evaluasi menjadi delapan, yaitu :

1. *Goal Oriented Evaluation Model*, dikembangkan oleh Tyler.
2. *Goal Free Evaluation Model*, dikembangkan oleh Scriven.
3. *Formatif-Sumatif Evaluation Model*, dikembangkan oleh Michael Scriven
4. *Countenance Evaluation Model*, dikembangkan oleh Stake.
5. *Responsive Evaluation Model*, dikembangkan oleh Stake.
6. *CSE-UCLA Evaluation Model*, menekankan pada “kapan” evaluasi dilakukan.
7. *CIPP Evaluation Model*, yang dikembangkan oleh Stufflebeam.
8. *Discrepancy Model*, yang dikembangkan oleh Provus.

Ada juga model evaluasi yang dikelompokkan Nana Sudjana dan R.Ibrahim (2007 : 234) yang membagi model evaluasi menjadi empat model utama, yaitu “*measurement, congruence, educational system, dan illumination*”. Dari beberapa model evaluasi di atas, beberapa diantaranya akan dikemukakan secara singkat sebagai berikut :

### **1. Model Tyler**

Nama model ini diambil dari nama pengembangnya yaitu Tyler. Dalam buku *Basic Principles of Curriculum and Instruction*, Tyler banyak mengemukakan ide dan gagasannya tentang evaluasi. Salah satu bab dari buku tersebut diberinya judul *how can the the effectiveness of learning experience be evaluated ?* Model ini dibangun atas dua dasar pemikiran. *Pertama*, evaluasi ditujukan kepada tingkah laku peserta didik. *Kedua*, evaluasi harus dilakukan pada tingkah laku awal peserta didik sebelum melaksanakan kegiatan pembelajaran dan sesudah melaksanakan kegiatan

pembelajaran (hasil). Dasar pemikiran yang kedua ini menunjukkan bahwa seorang evaluator harus dapat menentukan perubahan tingkah laku apa yang terjadi setelah peserta didik mengikuti pengalaman belajar tertentu, dan menegaskan bahwa perubahan yang terjadi merupakan perubahan yang disebabkan oleh pembelajaran.

Penggunaan model Tyler memerlukan informasi perubahan tingkah laku terutama pada saat sebelum dan sesudah terjadinya pembelajaran. Istilah yang populer dikalangan guru adalah tes awal (*pre-test*) dan tes akhir (*post-test*). Model ini mensyaratkan validitas informasi pada tes akhir. Untuk menjamin validitas ini maka perlu adanya kontrol dengan menggunakan desain eksperimen. Model Tyler disebut juga model *black box* karena model ini sangat menekankan adanya tes awal dan tes akhir. Dengan demikian, apa yang terjadi dalam proses tidak perlu diperhatikan. Dimensi proses ini dianggap sebagai “kotak hitam” yang menyimpan segala macam teka-teki. Menurut Tyler, ada tiga langkah pokok yang harus dilakukan, yaitu :

- a. Menentukan tujuan pembelajaran yang akan dievaluasi.
- b. Menentukan situasi dimana peserta didik memperoleh kesempatan untuk menunjukkan tingkah laku yang berhubungan dengan tujuan.
- c. Menentukan alat evaluasi yang akan dipergunakan untuk mengukur tingkah laku peserta didik.

## **2. Model yang Berorientasi pada Tujuan (*goal oriented evaluation model*)**

Dalam mendisain suatu program tentu tidak terlepas dari tujuan. Begitu pula dalam pendidikan, kurikulum dan pembelajaran, kita mengenal adanya hirarki tujuan pendidikan, yaitu tujuan pendidikan nasional, tujuan institusional, tujuan kurikuler, tujuan pembelajaran umum dan tujuan pembelajaran khusus. Model evaluasi ini menggunakan tujuan-tujuan tersebut sebagai kriteria untuk menentukan keberhasilan. Evaluasi diartikan sebagai proses pengukuran hinggamana tujuan program telah tercapai. Model ini dianggap lebih praktis untuk mendisain dan mengembangkan suatu program, karena menentukan hasil yang diinginkan dengan rumusan yang dapat diukur. Dengan demikian dapat dikatakan bahwa terdapat hubungan yang logis antara kegiatan, hasil dan prosedur pengukuran hasil. Tujuan model ini adalah membantu guru merumuskan tujuan dan menjelaskan hubungan antara tujuan dengan kegiatan. Jika rumusan tujuan program dapat diobservasi (*observable*) dan dapat diukur (*measurable*), maka kegiatan evaluasi pembelajaran akan menjadi lebih

praktis dan simpel. Di samping itu, model ini dapat membantu guru menjelaskan rencana pelaksanaan kegiatan suatu program dengan proses pencapaian tujuan. Instrumen yang digunakan bergantung kepada tujuan yang ingin diukur. Hasil evaluasi akan menggambarkan tingkat keberhasilan tujuan program berdasarkan kriteria program khusus. Kelebihan model ini terletak pada hubungan antara tujuan dengan kegiatan dan menekankan pada peserta didik sebagai aspek penting dalam program. Kekurangannya adalah memungkinkan terjadinya proses evaluasi melebihi konsekuensi yang tidak diharapkan.

### **3. Model Pengukuran**

Model pengukuran (*measurement model*) banyak mengemukakan pemikiran-pemikiran dari R.Thorndike dan R.L.Ebel. Sesuai dengan namanya, model ini sangat menitikberatkan pada kegiatan pengukuran. Pengukuran digunakan untuk menentukan kuantitas suatu sifat (*attribute*) tertentu yang dimiliki oleh objek, orang maupun peristiwa, dalam bentuk unit ukuran tertentu. Dalam bidang pendidikan, model ini telah diterapkan untuk mengungkap perbedaan-perbedaan individual maupun kelompok dalam hal kemampuan, minat dan sikap. Hasil evaluasi digunakan untuk keperluan seleksi peserta didik, bimbingan, dan perencanaan pendidikan. Objek evaluasi dalam model ini adalah tingkah laku peserta didik, mencakup hasil belajar (kognitif), pembawaan, sikap, minat, bakat, dan juga aspek-aspek kepribadian peserta didik. Instrumen yang digunakan pada umumnya adalah tes tertulis (*paper and pencil test*) dalam bentuk tes objektif, yang cenderung dibakukan. Oleh sebab itu, dalam menganalisis soal sangat memperhatikan *difficulty index* dan *index of discrimination*. Model ini menggunakan pendekatan Penilaian Acuan Norma (*norm-referenced assessment*).

### **4. Model Kesesuaian (Ralph W.Tyler, John B.Carrol, and Lee J.Cronbach)**

Menurut model ini, evaluasi adalah suatu kegiatan untuk melihat kesesuaian (*congruence*) antara tujuan dengan hasil belajar yang telah dicapai. Hasil evaluasi digunakan untuk menyempurnakan sistem bimbingan peserta didik dan untuk memberikan informasi kepada pihak-pihak yang memerlukan. Objek evaluasi adalah tingkah laku peserta didik, yaitu perubahan tingkah laku yang diinginkan (*intended behaviour*) pada akhir kegiatan pendidikan, baik yang menyangkut aspek kognitif, afektif maupun psikomotor. Untuk itu, teknik evaluasi yang digunakan tidak hanya

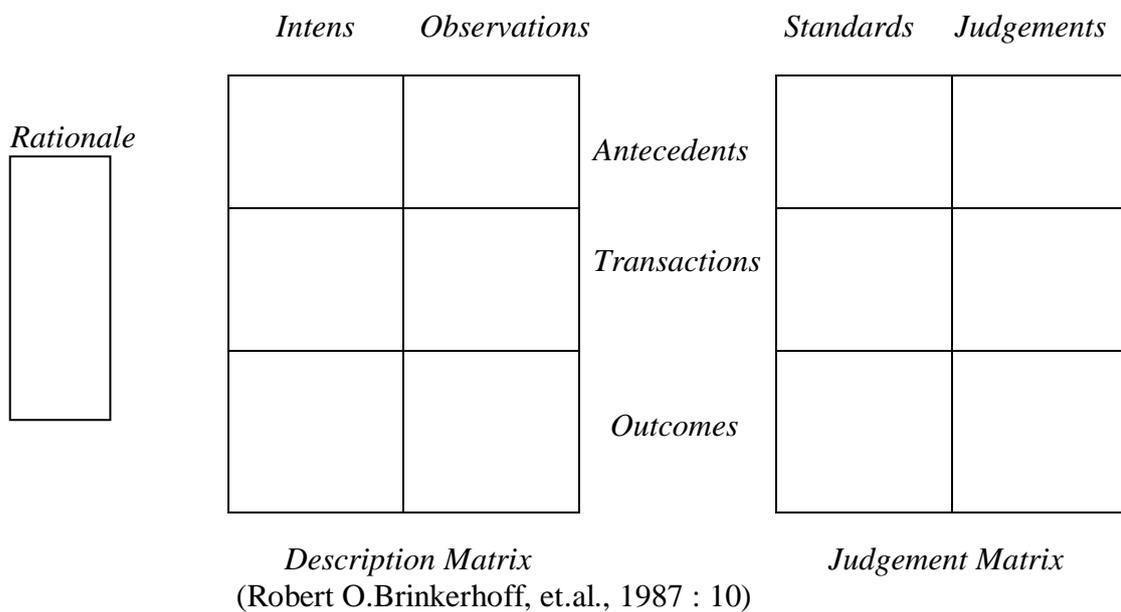
tes (tulisan, lisan, dan perbuatan), tetapi juga non-tes (observasi, wawancara, skala sikap, dan sebagainya). Model evaluasi ini memerlukan informasi perubahan tingkah laku pada dua tahap, yaitu sebelum dan sesudah kegiatan pembelajaran. Berdasarkan konsep ini, maka guru perlu melakukan *pre and post-test*. Adapun langkah-langkah yang harus ditempuh dalam model evaluasi ini adalah merumuskan tujuan tingkah laku (*behavioural objectives*), menentukan situasi dimana peserta didik dapat memperlihatkan tingkah laku yang akan dievaluasi, menyusun alat evaluasi, dan menggunakan hasil evaluasi. Oleh sebab itu, model ini menekankan pada pendekatan penilaian acuan patokan (*criterion-referenced assessment*).

5. ***Educational System Evaluation Model*** (Daniel L.Stufflebeam, Michael Scriven, Robert E.Stake, dan Malcolm M.Provus)

Menurut model ini, evaluasi berarti membandingkan *performance* dari berbagai dimensi (tidak hanya dimensi hasil saja) dengan sejumlah *criterion*, baik yang bersifat mutlak/intern maupun relatif/ekstern. Model yang menekankan sistem sebagai suatu keseluruhan ini sebenarnya merupakan penggabungan dari beberapa model, sehingga objek evaluasinya pun diambil dari beberapa model, yaitu (1) model *countenance* dari Stake, yang meliputi : keadaan sebelum kegiatan berlangsung (*antecedents*), kegiatan yang terjadi dan saling mempengaruhi (*transactions*), hasil yang diperoleh (*outcomes*) (2) model CIPP dari Stufflebeam, yang meliputi *Context, Input, Process, dan Product* (3) model Scriven yang meliputi *instrumental evaluation and consequential evaluation* (4) model Provus yang meliputi : *design, operation program, interim products, dan terminal products*. Dari keempat model yang tergabung dalam *educational system model*, pada kesempatan ini akan dijelaskan secara singkat tentang dua model, yaitu model *countenance* dan model CIPP.

Model Stake menitikberatkan evaluasi pada dua hal pokok, yaitu *description* dan *judgement*. Setiap hal tersebut terdiri atas tiga dimensi, seperti telah dijelaskan di atas, yaitu *antecedents (context)*, *transaction (process)*, dan *outcomes (output)*. *Description* terdiri atas dua aspek, yaitu *intents (goals)* dan *observation (effects)* atau yang sebenarnya terjadi. Sedangkan *judgement* terdiri atas dua aspek, yaitu *standard* dan *judgement*. Dalam model ini, evaluasi dilakukan dengan membandingkan antara satu program dengan program lain yang dianggap standar. Stake mengatakan

*description* berbeda dengan *judgement* atau menilai. Dalam ketiga dimensi di atas (*antecedents, transactions, outcomes*), data dibandingkan tidak hanya untuk menentukan apakah ada perbedaan tujuan dengan keadaan yang sebenarnya tetapi juga dibandingkan dengan standar yang absolut untuk menilai manfaat program. Menurut Stake, suatu hasil penelitian tidak dapat diandalkan jika tidak dilakukan evaluasi. Secara keseluruhan, model *Countenance* dapat digambarkan dalam matriks berikut ini :



Jika ingin menggunakan model *Countenance* dalam program pelatihan (misalnya), maka kita dapat menjelaskan hal-hal sebagai berikut :

- a. *Rationale*, yaitu menjelaskan pentingnya suatu program pelatihan.
- b. *Antecedents*, yaitu kondisi-kondisi yang diharapkan sebelum kegiatan pelatihan berlangsung, seperti motivasi, tingkat keterampilan, dan minat.
- c. *Transactions*, yaitu proses atau kegiatan-kegiatan yang saling mempengaruhi selama pelatihan.
- d. *Outcomes*, yaitu hasil yang diperoleh dari pelatihan, seperti pengetahuan, keterampilan, sikap dan nilai-nilai.
- e. *Judgements*, yaitu menilai pendekatan dan prosedur yang digunakan dalam pelatihan, para pelatih/instruktur, dan bahan-bahan.
- f. *Intents*, yaitu tujuan apa yang diharapkan dari suatu program pelatihan.
- g. *Observations*, yaitu apa yang dilihat oleh para pengamat tentang pelaksanaan pelatihan.

- h. *Standards*, yaitu apa yang diharapkan dari para *stakeholders*.
- i. *Judgements*, yaitu menilai suatu program, baik yang dilakukan oleh penilai itu sendiri maupun dari pihak-pihak lain.

Model CIPP berorientasi kepada suatu keputusan (*a decision oriented evaluation approach structured*). Tujuannya adalah untuk membantu administrator (kepala sekolah dan guru) di dalam membuat keputusan. Evaluasi diartikan sebagai suatu proses mendeskripsikan, memperoleh dan menyediakan informasi yang berguna untuk menilai alternatif keputusan (Stufflebeam, 1973 : 127). Sesuai dengan nama modelnya, model ini membagi empat jenis kegiatan evaluasi, yaitu :

- a. *Context evaluation to serve planning decision*, yaitu konteks evaluasi untuk membantu administrator merencanakan keputusan, menentukan kebutuhan program, dan merumuskan tujuan program.
- b. *Input evaluation, structuring decision*. Kegiatan evaluasi bertujuan untuk membantu mengatur keputusan, menentukan sumber-sumber, alternatif apa yang akan diambil, apa rencana dan strategi untuk mencapai kebutuhan, dan bagaimana prosedur kerja untuk mencapainya.
- c. *Process evaluation, to serve implementing decision*. Kegiatan evaluasi ini bertujuan untuk membantu melaksanakan keputusan. Pertanyaan yang harus Anda jawab adalah hinggamana suatu rencana telah dilaksanakan, apakah rencana tersebut sesuai dengan prosedur kerja, dan apa yang harus diperbaiki.
- d. *Product evaluation, to serve recycling decision*. Kegiatan evaluasi ini bertujuan untuk membantu keputusan selanjutnya. Pertanyaan yang harus Anda jawab adalah hasil apa yang telah dicapai dan apa yang dilakukan setelah program berjalan.

Proses evaluasi tidak hanya berakhir dengan suatu *description* mengenai keadaan sistem yang bersangkutan, tetapi harus sampai pada *judgment* sebagai kesimpulan dari hasil evaluasi. Model ini menuntut agar hasil evaluasi digunakan sebagai *input* untuk *decision making* dalam rangka penyempurnaan sistem secara keseluruhan. Pendekatan yang digunakan adalah penilaian acuan norma (PAN) dan penilaian acuan patokan (PAP). Untuk memahami lebih jauh tentang model CIPP ini, kita dapat melihat rincian penjelasan keempat dimensi tersebut dari segi tujuan, metode, dan hubungannya dengan pembuatan keputusan. Adapun rinciannya adalah :

|   | <b><i>Context Evaluation</i></b>   | <b><i>Input Evaluation</i></b>  | <b><i>Process Evaluation</i></b>   | <b><i>Product Evaluation</i></b>  |
|---|--|---|--|---|
| <b><i>Objective</i></b>   | To define the institutional context, to identify the target population and assess their needs, to identify opportunities for addressing the needs, to diagnose problems underlying the needs & to judge whether proposed objectives are sufficiently responsive to the assessed needs. | To identify & assess system capabilities, alternative program strategies, procedural designs for implementing the strategies, budgets, schedules, and program.  | To identify or predict, in process, defects in the procedural design or its implementation, to provide information for the preprogrammed decisions, and to record & judge procedural events & activities.  | To collect descriptions & judgements of outcomes & to relate them to objectives & to context, input & process information & to interpret their worth & merit.             |
| <b><i>Method</i></b>  | By using such methods as system analysis, survey, document review, hearings, interviews, diagnostic tests, & the Delplir technique.  | By inventorying & analyzing available human & material resources, solution strategies, & procedural designs for relevance, feasibility & economy. And by using such methods as literature search, visits to “misicle workers”, advocate teams & pilot trials. | By monitoring the activity’s potential procedural barriers & remaining alert to unanticipated ones, by obtaining specified information for programmed decisions, by describing the actual process & by continually interacting with & observing the activities of project staff. | By defining operationally & measuring outcomes criteria, by collecting judgements of outcomes from stakeholders, & by performing both qualitative & quantitative analyses |
| <b><i>Relation to decision making in the change process</i></b> | For deciding upon the setting to be served, the goals associated with meeting needs or using opportunities, & the objectives associated with solving problems, i.e., for planning needed changes. And to provide a basis for judging outcomes.   | For selecting sources of support, solution strategies & procedural designs, i.e., for structuring change activities. And to provide a basis for judging implementation  | For implementing and refining the program design and procedure, i.e., for effecting process control. And to provide a log of the actual process for later use in interpreting outcomes   | For deciding to continue, terminate, modify, or refocus a change activity, & present a clear record of effects (intended, positive & negative)                            |

(Robert O.Brinkerhoff, et.al., 1987 : 11)

## 6. Model Alkin

Model ini diambil dari nama pengembangnya yaitu Marvin Alkin (1969). Menurut Alkin, evaluasi adalah suatu proses untuk meyakinkan keputusan, mengumpulkan informasi, memilih informasi yang tepat, dan menganalisis informasi, sehingga dapat disusun laporan bagi pembuat keputusan dalam memilih beberapa alternatif. Alkin mengemukakan ada lima jenis evaluasi, yaitu :

- a. Sistem *assessment*, yaitu untuk memberikan informasi tentang keadaan atau posisi dari suatu sistem.
- b. Program *planning*, yaitu untuk membantu pemilihan program tertentu yang mungkin akan berhasil memenuhi kebutuhan program.
- c. Program *implementation*, yaitu untuk menyiapkan informasi apakah suatu program sudah diperkenalkan kepada kelompok tertentu yang tepat sebagaimana yang direncanakan.
- d. Program *improvement*, yaitu memberikan informasi tentang bagaimana suatu program dapat berfungsi, bekerja atau berjalan. Apakah sesuai dengan pencapaian tujuan ? Adakah hal-hal atau masalah-masalah baru yang muncul secara tiba-tiba ?
- e. Program *certification*, yaitu memberikan informasi tentang nilai atau manfaat suatu program.

## 7. Model Brinkerhoff

Robert O.Brinkerhoff (1987) mengemukakan dilihat dari segi disain evaluasi, ada tiga jenis evaluasi yang disusun berdasarkan penggabungan elemen-elemen yang sama, yaitu :

### a. *Fixed vs Emergent Evaluation Design*

Disain evaluasi *fixed* (tetap) harus direncanakan dan disusun secara sistematis-terstruktur sebelum program dilaksanakan. Namun demikian, disain *fixed* dapat juga disesuaikan dengan kebutuhan yang sewaktu-waktu dapat berubah. Disain evaluasi ini dikembangkan berdasarkan tujuan program, kemudian disusun pertanyaan-pertanyaan untuk mengumpulkan berbagai informasi yang diperoleh dari sumber-sumber tertentu. Begitu juga dengan model analisis yang akan digunakan harus dibuat sebelum program dilaksanakan. Pihak pemakai (*user*) akan menerima informasi sebagai hasil evaluasi sesuai dengan tujuan yang telah ditetapkan. Pada

umumnya, evaluasi formal yang dibuat secara individual menggunakan disain *fixed*, karena tujuan program sudah ditetapkan sebelumnya. Begitu juga dengan anggaran biaya dan organisasi pelaksana, yang kesemuanya dituangkan dalam sebuah proposal evaluasi.

Kegiatan-kegiatan evaluasi yang dilakukan dalam disain *fixed* ini antara lain menyusun pertanyaan-pertanyaan, menyusun dan menyiapkan instrumen, menganalisis hasil evaluasi, dan melaporkan hasil evaluasi secara formal kepada pihak pemakai. Dalam menyusun pertanyaan-pertanyaan atau merumuskan masalah, seorang evaluator harus mengacu kepada tujuan program. Di samping itu, evaluator juga harus merangsang *audience* untuk memperbaiki pertanyaan-pertanyaan yang dianggap kurang relevan. Sesuai dengan kegiatan-kegiatan evaluasi ini, maka wajar bila disain *fixed* ini banyak memerlukan biaya. Belum lagi proses komunikasi yang harus dibangun secara teratur dan kontinu, baik secara langsung maupun tak langsung antara evaluator dengan *audience* atau klien.

Untuk mengumpulkan data dalam disain ini dapat menggunakan berbagai teknik, seperti tes, observasi, wawancara, kuesioner, dan skala penilaian. Untuk itu, syarat-syarat penyusunan instrumen yang baik, seperti validitas dan reliabilitas tetap harus diperhatikan, karena data yang dikumpulkan biasanya bersifat kuantitatif. Dalam penyusunan disain biasanya didiskusikan terlebih dahulu dengan pihak pemakai, sehingga jika terdapat kekurangan dapat segera diperbaiki.

Sementara itu, dalam disain evaluasi *emergent*, tujuan evaluasi adalah untuk beradaptasi dengan situasi yang sedang berlangsung dan berkembang, seperti menampung pendapat audiensi, masalah-masalah, dan kegiatan program. Proses adaptasi ini tentu memerlukan waktu yang cukup lama, mulai dari awal sampai dengan akhir kegiatan guna menetapkan dan merumuskan tujuan dan isu. Hal ini wajar karena hal tersebut tidak ditentukan sebelumnya. Disini, seorang evaluator tidak perlu mendorong audiensi untuk memikirkan tentang suatu program atau isu-isu evaluasi karena audiensi akan menentukan sendiri isu-isu dan informasi penting lainnya yang diperlukan dalam disain *emergent*.

Selama proses evaluasi, seorang evaluator harus tetap menjalin komunikasi yang kontinu dengan audiensi, sehingga data dan informasi yang dikumpulkan tidak terputus dan tetap utuh. Teknik pengumpulan data dapat menggunakan observasi,

studi kasus dan laporan tim pendukung. Pengukuran tidak selalu mengacu kepada tujuan program, seperti yang biasa dilakukan, bahkan seorang evaluator sering mengabaikan penggunaan teknik pengukuran karena informasi yang dibutuhkan lebih bersifat kualitatif-naturalistik. Hal ini dimaksudkan agar informasi yang dikumpulkan lebih banyak, mendalam dan bermanfaat. Dengan demikian, disain akan terus berkembang dan berubah sesuai dengan situasi dan kondisi di lapangan.

*b. Formative vs Summative Evaluation*

**8. *Illuminative Model*** (Malcolm Parlett dan Hamilton)

Jika model *measurement* dan *congruence* lebih berorientasi pada evaluasi kuantitatif-terstruktur, maka model ini lebih menekankan pada evaluasi kualitatif-terbuka (*open-ended*). Kegiatan evaluasi dihubungkan dengan *learning milieu*, dalam konteks sekolah sebagai lingkungan material dan psiko-sosial, dimana guru dan peserta didik dapat berinteraksi. Tujuan evaluasi adalah untuk mempelajari secara cermat dan hati-hati terhadap pelaksanaan sistem, faktor-faktor yang mempengaruhinya, kelebihan dan kekurangan sistem, dan pengaruh sistem terhadap pengalaman belajar peserta didik. Hasil evaluasi lebih bersifat deskriptif dan interpretasi, bukan pengukuran dan prediksi. Model ini lebih banyak menggunakan *judgment*. Fungsi evaluasi adalah sebagai input untuk kepentingan pengambilan keputusan dalam rangka penyesuaian dan penyempurnaan sistem yang sedang dikembangkan.

Objek evaluasi model ini mencakup latar belakang dan perkembangan sistem, proses pelaksanaan sistem, hasil belajar peserta didik, kesukaran-kesukaran yang dialami dari perencanaan sampai dengan pelaksanaan, termasuk efek samping dari sistem itu sendiri. Pendekatan yang digunakan lebih menyerupai pendekatan yang diterapkan dalam bidang antropologi sosial, psikiatri, dan sosiologi. Cara-cara yang digunakan tidak bersifat *standard*, melainkan bersifat fleksibel dan selektif. Berdasarkan tujuan dan pendekatan evaluasi dalam model ini, maka ada tiga fase evaluasi yang harus ditempuh, yaitu : *observe, inquiry further, dan seek to explain*.

**9. Model Responsif**

Sebagaimana model illuminatif, model ini juga menekankan pada pendekatan kualitatif-naturalistik. Evaluasi tidak diartikan sebagai pengukuran melainkan

pemberian makna atau melukiskan sebuah realitas dari berbagai perspektif orang-orang yang terlibat, berminat dan berkepentingan dengan program. Tujuan evaluasi adalah untuk memahami semua komponen program melalui berbagai sudut pandangan yang berbeda. Sesuai dengan pendekatan yang digunakan, maka model ini kurang percaya terhadap hal-hal yang bersifat kuantitatif. Instrumen yang digunakan pada umumnya mengandalkan observasi langsung maupun tak langsung dengan interpretasi data yang impresionistik. Langkah-langkah kegiatan evaluasi meliputi observasi, merekam hasil wawancara, mengumpulkan data, mengecek pengetahuan awal (*preliminary understanding*) peserta didik dan mengembangkan disain atau model. Berdasarkan langkah-langkah ini, evaluator mencoba responsif terhadap orang-orang yang berkepentingan pada hasil evaluasi. Hal yang penting dalam model responsif adalah pengumpulan dan sintesis data.

Kelebihan model ini adalah peka terhadap berbagai pandangan dan kemampuannya mengakomodasi pendapat yang ambigu serta tidak fokus. Sedangkan kekurangannya antara lain (1) pembuat keputusan sulit menentukan prioritas atau penyederhanaan informasi (2) tidak mungkin menampung semua sudut pandangan dari berbagai kelompok (3) membutuhkan waktu dan tenaga. Evaluator harus dapat beradaptasi dengan lingkungan yang diamati. Untuk mempelajari lebih jauh tentang model ini, silahkan Anda membaca buku Stake (1975) atau Lincoln dan Guba (1985).

Setelah Anda mempelajari berbagai model evaluasi, model mana yang akan digunakan dalam menilai suatu program? Jawabannya tentu sangat bergantung kepada tujuan evaluasi yang ditetapkan. Namun demikian, perlu juga Anda pahami bahwa keberhasilan suatu evaluasi program secara keseluruhan bukan hanya dipengaruhi penggunaan yang tepat pada sebuah model evaluasi melainkan juga dipengaruhi oleh berbagai faktor. *Pertama*, tujuan program, baik tujuan umum maupun tujuan khusus. Seringkali kedua tujuan program ini saling bertentangan satu sama lain dilihat dari kebutuhan dan komponen-komponen program lainnya. Bahkan, kadang-kadang evaluator sendiri mempunyai tujuan sendiri-sendiri. Semuanya harus dipertimbangkan agar terdapat keseimbangan dan keserasian.

*Kedua*, sistem sekolah. Faktor ini perlu dipertimbangkan dengan matang dan hati-hati karena melibatkan berbagai komponen yang saling berinteraksi dan

ketergantungan. Mengingat kompleksnya sistem sekolah, maka fungsi sekolah juga menjadi ganda. Di satu pihak sekolah ingin mewariskan kebudayaan masa lampau dengan sistem norma, nilai dan adat yang dianggap terbaik untuk generasi muda. Di pihak lain, sekolah berkewajiban mempersiapkan peserta didik menghadapi masa depan, memperoleh keterampilan dan kemampuan untuk berinovasi, bahkan menghasilkan perubahan. Jadi, sekolah sekaligus bersikap konservatif-radikal serta reaksioner-progresif. Oleh sebab itu, peranan evaluasi menjadi sangat penting. Tujuannya adalah untuk melihat dan mempertimbangkan hal-hal apa yang perlu diberikan di sekolah. Begitu juga bentuk kurikulum dan silabus mata pelajaran sangat bergantung pada evaluasi yang dilaksanakan oleh guru-guru di sekolah, sehingga timbul masalah lainnya yaitu teknik evaluasi apa yang akan digunakan untuk mencapai tujuan itu.

*Ketiga*, program pembinaan. Banyak program pembinaan yang belum menyentuh secara langsung tentang evaluasi. Program pembinaan guru, misalnya, lebih banyak difokuskan kepada pengembangan kurikulum dan metodologi pembelajaran. Hal ini pula yang menyebabkan perbaikan sistem evaluasi pembelajaran menjadi kurang efektif. Di samping itu, guru juga sering dihadapkan dengan beragam kegiatan, seperti membuat persiapan mengajar, mengikuti kegiatan ekstra kurikuler, penyesuaian diri, dan kegiatan administratif lainnya. Artinya, bagaimana mungkin kualitas sistem evaluasi pembelajaran di sekolah dapat ditingkatkan, bila fokus pembinaan guru hanya menyentuh domain-domain tertentu saja, ditambah lagi dengan kesibukan-kesibukan guru di luar tugas pokoknya sebagai pengajar.