

## **BAB II**

### **MODEL-MODEL PENDEKATAN EVALUASI PROGRAM PENDIDIKAN**

Woolfolk dan Nicolich mengemukakan bahwa penilaian atau evaluasi merupakan suatu proses membandingkan informasi dengan kriteria, kemudian membuat pertimbangan; yakni membuat keputusan berdasarkan nilai-nilai. Sejalan dengan pengertian tersebut, Raka Joni mengemukakan bahwa penilaian adalah “peneta-pan baik-buruk terhadap sesuatu berdasarkan kriteria tertentu”. Gronlund dan Linn mengemukakan bahwa penilaian merupakan proses pengumpulan informasi, ana-lisis dan interpretasi informasi yang sistematis untuk menentukan sejauhmana murid mencapai tujuan pembelajaran. Secara lebih rinci, *Phi Delta Kappa National Study Committee of Evaluation* menguraikan pengertian evaluasi sebagai proses pen-carian, perolehan dan penyediaan informasi yang berguna bagi pertimbangan alternatif-alternatif keputusan. Pengertian ini ber-kaitan dengan tiga hal mendasar, yaitu: 1) evaluasi merupakan suatu proses sistematis yang berkelanjutan; 2) proses tersebut me-liputi tiga langkah, yakni: (1) menyusun pertanyaan yang memer-lukan jawaban dan informasi spesifik yang ingin diperoleh, (2) me-ngumpulkan data yang relevan, (3) menyajikan informasi yang di-hasilkan kepada pengambil keputusan yang akan mempertim-bangkan dan menginterpretasikannya berkaitan dengan alternatif keputusan yang akan diambil; 3) evaluasi mendukung proses pem-buatan keputusan dengan menyediakan alternatif-alternatif yang terseleksi serta menidaklanjuti konsekuensi-konsekuensinya.

Menyimak berbagai pengertian yang diberikan para ahli tentang penilaian, Stufflebeam et.al mengelompokkan adanya tiga sudut pandang definisi penilaian, yaitu: 1) pengertian yang mengidentikkan penilaian dengan pengukuran, 2) pengertian penilaian sebagai sebuah

proses melihat kongruensi antara tujuan dengan apa yang dilaksanakan, dan 3) penilaian sebagai sebuah pertimbangan (*judgment*) profesional. Perbedaan-perbedaan pengertian diatas nampak menunjukkan adanya perbedaan pendekatan yang dilakukan dalam melakukan penilaian.

Tiga pendekatan evaluasi program pendidikan yang banyak dikenal dan sering dijadikan rujukan dalam pelaksanaan evaluasi program pendidikan, yakni : a) *Objective-Oriented Approach*, b) *Management-Oriented Approach*, dan c) *Naturalistic-Participant Approach*. dengan uraian-uraian yang disertai contoh penggunaannya.

Tiga pendekatan evaluasi program pendidikan yang akan dibahas dalam bab ini, yaitu diharapkan dapat:

1. memberikan contoh model objective-oriented evaluation approach;
2. menjelaskan penggunaan model objective-oriented evaluation approach;
3. menjelaskan kekuatan model objective-oriented evaluation approach;
4. menjelaskan kelemahan model objective-oriented evaluation approach;
5. memberikan contoh model management-oriented evaluation approach;
6. menjelaskan penggunaan model management-oriented evaluation approach;
7. menjelaskan kekuatan model management-oriented evaluation approach;
8. menjelaskan kelemahan model management-oriented evaluation approach;
9. memberikan contoh model naturalistic and participant evaluation approach;
10. menjelaskan penggunaan model naturalistic and participant evaluation approach;

11. menjelaskan kekuatan model naturalistic and participant evaluation approach;
12. menjelaskan kelemahan model naturalistic and participant evaluation approach;

## **A. OBJECTIVE- ORIENTED EVALUATION APPROACH**

Pernahkah anda mendengar istilah *objective-oriented approach*? Model *Objective-Oriented Approach* (pendekatan penilaian berorientasi tujuan) adalah pendekatan dalam melakukan evaluasi program yang menitik beratkan pada penilaian ketercapaian tujuan. Oleh karena itu, pandangan ini mempersyaratkan bahwa suatu program pendidikan harus menetapkan atau merumuskan tujuan-tujuan spesifiknya secara jelas. Terhadap tujuan-tujuan program yang sudah ditetapkan tersebut lah evaluasi program difokuskan. Tujuan program yang dimaksud bisa saja hanya tujuan dari sebuah program pembelajaran di kelas dalam satu mata pelajaran, atau juga tujuan program dalam pengertian yang lebih luas, misalnya tujuan program sekolah dalam satu tahun, tujuan program pembangunan pendidikan tahun X di Kabupaten/ Kota Anu, dsb.

Agar anda tidak menganggap konsep ini sesuatu yang sangat baru bagi anda, coba anda ingat-ingat apa yang sering dilakukan di lingkungan sekolah. Apakah anda ingat apa yang dilakukan seorang guru ketika melakukan penilaian keberhasilan pembelajaran yang dilakukannya? Kebanyakan guru melakukan penilaian dengan memberikan tes kepada siswa. Bentuk kegiatan tes bagi siswa sebagai penilaian terhadap keberhasilan pembelajaran dilandasi asumsi bahwa suatu pembelajaran dianggap berhasil baik jika tujuan-tujuan belajar tercapai. Ketercapaian tujuan belajar tersebut tercermin dari hasil tes siswa. Oleh karena itu, tes sebagai alat (*instrument*) untuk melakukan penilaian selalu dibuat berdasarkan pada tujuan-tujuan belajar yang telah ditetapkan. Kalau anda pernah menjadi seorang guru, anda tentu masih ingat bagaimana membuat kisi-kisi penyusunan soal yang selalu didasarkan pada ranah-anah hasil belajar yang sudah ditetapkan sebagai tujuan pembelajaran. Kegiatan penilaian seperti yang dilakukan

guru itu adalah salah satu contoh penerapan pendekatan penilaian program yang berorientasi tujuan (*objective-oriented approach*).

### **Konsep Dasar dan Sejarah Perkembangan**

Pendekatan penilaian berorientasi tujuan (*objective-oriented approach*) berkembang sejak diperkenalkan tahun 1930an, dimana penerapannya dalam dunia pendidikan dimulai dan sangat kental dipengaruhi pemikiran Tyler. Anda masih ingat nama tokoh ini? Coba ingat-ingat ketika anda mempelajari tentang pengembangan kurikulum (*curriculum development*). Dalam perkembangan selanjutnya tradisi Tyler ini juga dikembangkan oleh beberapa ahli, seperti: Metfessel dan Michael (1967), Hammond (1973), dan Provus (1973). Pada bagian berikut secara ringkas akan diuraikan pokok-pokok pikiran dari masing-masing ahli tersebut.

Tyler mendefinisikan penilaian pendidikan sebagai suatu proses untuk menentukan sejauhmana tujuan-tujuan pendidikan dari program sekolah atau kurikulum tercapai. Pendekatan penilaian yang dikemukakan Tyler ini meliputi langkah-langkah sebagai berikut:

- 1) Menentukan tujuan secara jelas
- 2) Mengklasifikasikan tujuan-tujuan tersebut
- 3) Mendefinisikan tujuan-tujuan dalam istilah perilaku terukur
- 4) Temukan situasi dimana prestasi atau tujuan dapat diperlihatkan
- 5) Mengembangkan atau memilih teknik-teknik pengukuran
- 6) Mengumpulkan data
- 7) Membandingkan data kinerja dengan tujuan-tujuan yang dinyatakan dalam perilaku terukur.

Langkah-langkah sebagaimana diuraikan di atas merupakan suatu siklus, artinya bahwa jika dari hasil membandingkan data kinerja dengan tujuan sudah diperoleh berupa kesenjangan-kesenjangan, maka perlu dilakukan perumusan/ penentuan ulang tujuan program yang telah dievaluasi tersebut.

Kalau kita simak secara seksama, langkah-langkah di atas terdiri dari dua bagian pokok, yaitu: 1) bagian yang terkait dengan kegiatan perencanaan program (langkah satu sampai tiga), 2) bagian yang secara langsung memang merupakan kegiatan dalam tahap evaluasi program (langkah empat dan selanjutnya). Dengan demikian, siklus kegiatan yang dimaksud sebenarnya lebih merupakan siklus kegiatan pengelolaan dan pengembangan program. Hal ini bisa dimaklumi oleh karena pemikiran ini dilahirkan dalam rangka pengembangan kurikulum.

Pola pikir yang ditawarkan Tyler ini sangat logis dan dapat diterima secara ilmiah, bahkan mudah untuk ditiru/ dilakukan oleh para pelaksana penilaian pendidikan (*evaluator*). Salah satu penerapan model ini oleh Tyler adalah bagaimana melakukan pengukuran tes kemampuan awal siswa (*pre-test*) dibandingkan dengan hasil pengukuran paska kegiatan pembelajaran (*post-test*). Kegiatan ini menjadi salah satu teknik yang banyak berpengaruh terhadap cara-cara penilaian program pembelajaran di dunia pendidikan. Contoh yang dilakukan Tyler ini pula lah yang banyak dilakukan oleh guru-guru kita dalam melakukan penilaian keberhasilan program pembelajaran di kelas selama ini. Secara praktis, pendekatan ini memang tidak terlalu menyita waktu karena hanya dilakukan pada akhir kegiatan pembelajaran. Di samping itu, dengan pendekatan seperti ini sangat sejalan dengan tradisi pemikiran manajemen/ pengelolaan yang menempatkan kegiatan evaluasi sebagai kegiatan terakhir.

Dari pengalamannya melakukan penilaian program pendidikan, Tyler mengadvokasikan tujuan-tujuan umum pendidikan yang perlu menjadi criteria dalam melakukan penilaian program pendidikan. Untuk pendidikan di Amerika, Tyler merekomendasikan 6 (enam) tujuan umum pendidikan, yaitu:

- 1) memperoleh informasi,
- 2) *mengembangkan* kebiasaan bekerja dan ketrampilan belajar,
- 3) mengembangkan cara berfikir yang efektif,

- 4) menginternalisasikan sikap social, minat, apresiasi, dan sensitifitas,
- 5) memelihara kesehatan fisik, dan
- 6) mengembangkan filsafat hidup.

Pemikiran-pemikiran Tyler tentang tujuan umum pendidikan ternyata terus merangsang pemikiran para ahli untuk mereformulasi dan mendefinisikan tujuan-tujuan pendidikan secara beragam. Setidaknya ada dua karya besar yang sangat terkait dengan pola pikir Tyler ini. *Pertama*, Pusat Evaluasi di *Western Michigan University* bekerjasama dengan *Toledo, Ohio, Public Schools*, mengembangkan *Handbook of Educational Variables*. *Handbook* ini membagi perkembangan siswa sekolah dasar dan menengah ke dalam 7 (tujuh) kategori, yaitu:

- 1) intelektual,
- 2) emosional,
- 3) Fisik dan Rekreasi,
- 4) Estetika dan Budaya,
- 5) Moral,
- 6) Vokasional,
- 7) Sosial.

*Kedua*, karya Goodlad's (1979) yang menguraikan 12 (duabelas) kategori yang menjadi area tujuan sekolah, yaitu:

- 1) Penuntasan ketrampilan dasar atau proses fundamental,
- 2) Pendidikan karier/ pendidikan vokasional,
- 3) Pengembangan intelektual,
- 4) Enkulturasi,
- 5) Hubungan interpersonal,
- 6) Otonomi,
- 7) Kewarganegaraan,
- 8) Kreatifitas dan persepsi estetika,
- 9) Konsep diri,
- 10) Kesehatan fisik dan emosional

- 11) Moral dan karakter etika
- 12) Realisasi diri.

Tujuan sekolah sebagaimana digambarkan di atas sejalan dengan berbagai perkembangan pemikiran yang berkaitan dengan potensi manusia. Apa yang sekarang banyak dihebohkan dengan istilah *Intellectual Intelequences*, *Emotional Intelequences*, dan *Spiritual Intelequences*, sebenarnya bukan merupakan hal baru di dunia pendidikan. Kalaupun ada hal baru adalah semangatnya untuk makin menegaskan dan menegakkan hakekat pendidikan yang sesungguhnya. Demikianpun, pandangan Gardner (1996) tentang *Multiple Intelequence* yang ia kembangkan atas kritiknya terhadap Intelekuensi yang dikemukakan para pendahulunya, juga dalam koridor semangat yang sama. Oleh karena itu, dua daftar tujuan sekolah di atas hanyalah contoh dari keragaman pendapat para ahli tentang tujuan umum sekolah. Anda bisa membaca lebih banyak lagi daftar serupa.

Dalam rangka menjelaskan secara spesifik tentang bagaimana penilaian terhadap ketercapaian tujuan dilakukan, Sanders dan Cunningham (1974), mengemukakan metode yang diklasifikasikannya menjadi dua, yaitu metode-metode logis dan metode-metode empiris. Adapun yang termasuk pada kategori metode logis, antara lain:

- 1) Menguji cogency tentang argument-argumen atau rasional yang melandasi setiap tujuan pendidikan yang ditetapkan,
- 2) Menguji konsekuensi dari pencapaian tujuan. Dengan memperkirakan secara logis konsekuensi-konsekuensi dari pencapaian tujuan, baik kekuatan dan kelemahan dalam mencapai tujuan tersebut mungkin **revealed**.
- 3) Mempertimbangkan apakah nilai-nilai luhur seperti hukum dan kebijakan itu relevan dengan praktek-praktek kehidupan yang ada, dengan prinsip-prinsip moral, dengan cita-cita masyarakat, ataupun konstitusi yang ada.



Sedangkan metode-metode empiric, antara lain:

- 1) Mengumpulkan data untuk menggambarkan pertimbangan-pertimbangan tentang nilai dari sebuah tujuan,
- 2) Menyelenggarakan pertemuan pakar, dengar pendapat, atau panel untuk mengkaji dan mengevaluasi tujuan-tujuan yang potensial,
- 3) Melaksanakan studi analisis isi (content) terhadap arsip-arsip, seperti bahan pidato, catatan rapat, editorial, Koran, dll.

Kedua cara di atas bukan merupakan suatu perbedaan, tetapi lebih merupakan saling melengkapi. Oleh karena itu, dalam praktek melakukan penilaian berorientasi tujuan kedua cara ini lebih banyak dimanfaatkan sebagai suatu urutan kegiatan yang dilakukan. Artinya, pada awal penilaian dilakukan melalui analisis logis, kemudian dilanjutkan dengan analisis empiric. Kecenderungan ini terutama banyak dipengaruhi oleh pemikiran ilmiah (*scientific methods*) yang mempersyaratkan adanya kegiatan deduktif-induktif dalam melakukan suatu kajian.

Paradigma penilaian lain yang banyak dipengaruhi pemikiran Tyler adalah yang dikembangkan Metfessel dan Michael (1967). Kontribusi yang paling banyak dari pemikiran ini dalam evaluasi berorientasi tujuan adalah berkembangnya visi instrument alternative untuk pegumpulan data. Adapun proses penilaian yang dikembangkan pemikiran ini meliputi langkah-langkah sebagai berikut:

- 1) melibatkan seluruh komunitas sekolah sebagai fasilitator dalam penilaian program,
- 2) memformulasikan tujuan yang spesifik dan kohesif
- 3) menterjemahkan tujuan yang spesifik tersebut kedalam format-format yang komunikatif dan dapat diaplikasikan untuk memfasilitasi pembelajaran di lingkungan sekolah yang bersangkutan,
- 4) memilih atau mengkonstruksi instrumen-instrumen yang akan dipergunakan dalam pengukuran untuk dapat menyimpulkan efektifitas suatu program,

- 5) Melakukan observasi secara periodic dengan menggunakan instrumen-instrumen pengukuran perilaku yang valid (tes, skala, dll.),
- 6) Menganalisis data menggunakan statistika yang tepat,
- 7) Menginterpretasikan data dengan menggunakan standar tingkat kinerja yang diharapkan,
- 8) Meyusun rekomendasi untk implementasi, modifikasi, atau revisi tujuan-tujuan program selanjutnya

Kalau kita simak pemikiran di atas, dilihat dari langkah-langkah yang dilakukan hampir keseluruhan sama dengan apa yang dikembangkan Tyler. Hal yang menonjol adalah proses pelibatan komunitas yang terkait dengan program (*stakeholders*). Langkah ini dianggap strategis dalam rangka menjamin obyektifitas penilaian yang dilakukan.

Pemikiran lain yang termasuk pendekatan penilaian berorientasi tujuan adalah paradigma peniaian Hammond's (1973). Penilaian dalam pemikiran ini tidak hanya memusatkan perhatian pada tercapai tidaknya tujuan, tetapi juga melakukan kajian terhadap persoalan: mengapa suatu inovasi gagal dan inovasi lainnya sukses? Untuk mengidentifikasi factor-faktor yang mempengaruhi keberhasilan suatu program pendidikan, Hammond mengembangkan kubus tiga dimensi, terdiri dari: 1) dimensi pembelajaran, 2) dimensi kelembagaan, dan 3) dimensi tujuan. Kubus ini digunakan untuk menggambarkan program pendidikan dan mengorganisasikan variable-variabel yang dievaluasi. Kubus ini dinamakan Hammond sebagai '*Structure of Evaluation*'.

- 1) *Dimensi Pembelajaran*, menggambarkan karakteristik aktifitas pendidikan yang akan dievaluasi, terdiri dari:
  - a) Organisasi, termasuk didalamnya: waktu, jadwal kegiatan, urutan mata peajaran, dan juga termasuk organisasi sekolah, baik vertical maupun horizontal.
  - b) Konten, meliputi topic-topik yang termasuk akan dievaluasi;
  - c) Metode, mencakup seluruh aktifitas pembelajaran, tipe-tipe iteraksi guru-murid, dan teori-teori pembelajaran;
  - d) Fasilitas, meliputi: ruangan, peralatan, dan sarana-prasarana lainnya.
  - e) Biaya, mencakup anggaran yang diperlukan untuk fasilitas, pemeliharaan, dan personil.
- 2) *Dimensi Kelembagaan*, menggambarkan karakteristik individu atau kelompok yang terlibat dalam aktifitas pendidikan yang akan dievaluasi meliputi:
  - a) Siswa, mencakup: umur, jenis kelamin, tingkat/ kelas, latar belakang keluarga, kelas social, kesehatan, kemampuan, minat, dan prestasi belajar;
  - b) Guru, administrator, dan tenaga kependidikan lainnya, mencakup: umur, jenis kelamin, ras/ suku, agama, kesehatan, latar belakang pendidikan, pengalaman, kebiasaan kerja, dl.
  - c) Keluarga, mencakup tingkat keterlibatannya dalam pelaksanaan kegiatan yang akan dinilai, serta karakteristik umum, seperti: budaya, bahasa, ukuran keluarga, status perkawinan, tingkat pendidikan orangtua, afiliasi politik, agama, dll.
- 3) *Dimensi Tujuan*, menggambarkan ranah tujuan aktifitas pendidikan yang akan dievaluasi, terdiri dari:
  - a) Tujuan kognitif, mencakup pengetahuan dan ketrampilan intelektual;
  - b) Tujuan Afektif, meliputi: minat, sikap, perasaan, dan emosi;

- c) Tujuan Psikomotorik, termasuk didalamnya ketrampilan fisik dan koordinasi.

Dari kubus tersebut terbentuk 90 (sembilan puluh) sel yang potensial dapat dipergunakan. Setiap sel memungkinkan untuk menentukan tipe-tipe pertanyaan evaluasi secara umum. Dalam mengembangkan pertanyaan-pertanyaan tersebut, seorang penilai dapat memanfaatkan 3 (tiga) sel dari tiga dimensi yang berbeda sebagai substansi permasalahan.

Contoh:

- 1) Apakah guru (dimensi 2) menggunakan materi pelajaran (dimensi 1) untuk mencapai tujuan-tujuan afektif (dimensi 3) ?
- 2) Apakah materi pembelajaran (dimensi 1) yang diajarkan guru (dimensi 2) cukup memadai untuk mencapai tujuan-tujuan psikomotorik (dimensi 3) ?

Anda bias bayangkan, bagaimana jika anda mencoba 'mengawinkan' antar sel tersebut untuk kepentingan perumusan focus evaluasi. Betapa banyaknya persoalan yang menjadi wilayah penilaian program pendidikan itu.

Mengadaptasi pendekatan Tyler, Hammond's mengemukakan langkah-langkah penilaian yang harus dilakukan adalah sebagai berikut:

1. Mendefinisikan program
2. Mendefinisikan variable-variabel deskriptif, dengan menggunakan kubus
3. Merumuskan tujuan-tujuan program
4. Mengukur kinerja program
5. Menganalisis hasil pengukuran
6. Membandingkan hasil yang diperoleh dengan tujuan

Pendekatan lain yang banyak dipengaruhi pemikiran Tyler dikembangkan Provus berdasarkan pada tugas-tugas evaluasi di sebuah sekolah umum di Pittsburgh, Pennsylvania. Provus (1973) memandang

penilaian sebagai proses pengelolaan informasi berkelanjutan yang dirancang memberi pelayanan sebagai *the watchdog of program management* dan *the handmaiden of administration in the management of program development through sound decision making*.

Walaupun nampak adanya pendekatan manajemen dalam pemikiran Provus, tetapi tradisi Tyler lebih dominan. Hal ini dapat dilihat dari definisi evaluasi yang ia kembangkan. Menurut Provus, evaluasi adalah proses: 1) menyetujui berdasarkan standar (istilah lain yang digunakan secara bergantian dengan istilah tujuan), 2) menentukan apakah ada kesenjangan antara kinerja aspek-aspek program dengan standar kinerja yang ditetapkan; 3) menggunakan informasi tentang kesenjangan-kesenjangan yang ditemukan sebagai bahan untuk meningkatkan mengelola, atau mengakhiri program atau salah satu aspek dari program tersebut.

Pendekatan yang diperkenalkan Provus ini dinamakan *Discrepancy Evaluation Model*. Pendekatan ini memperkenalkan pelaksanaan evaluasi dengan langkah-langkah yang perlu dilakukan, meliputi:

1. Definisi
2. Instalasi
3. Proses
4. Produk
5. Analisis Biaya-Manfaat (*Cost-Benefit Analysis*)

Dalam tahap *definisi*, focus kegiatan dilakukan untuk merumuskan tujuan, proses atau aktifitas, serta pengalokasian sumberdaya dan partisipan untuk melakukan aktifitas dan mencapai tujuan-tujuan yang telah ditetapkan. Menurut Provus, program pendidikan merupakan system dinamis yang meliputi inputs (*antecedent*), proses, dan outputs (juga *outcomes*). Standar atau harapan-harapan yang ingin dicapai ditentukan untk masing-masing

komponen tersebut. Standar ini merupakan tujuan program yang kemudian menjadi criteria dalam kegiatan penilaian yang dilakukan.

Selama tahap *instalasi*, rancangan program digunakan sebagai standar untuk mempertimbangkan langkah-langkah operasional program. Seorang evaluator perlu mengembangkan seperangkat tes kongruensi untuk mengidentifikasi tiap kesenjangan antara instalasi program atau aktifitas yang diharapkan dan yang actual. Hal ini perlu untuk meyakinkan bahwa program telah diinstal sesuai dengan rancangan yang ditetapkan. Pengalaman selama ini menunjukkan bahwa banyak rancangan program yang sama dioperasionalkan oleh guru-guru dengan aktifitas yang berbeda-beda.

Pada tahap *proses*, evaluasi difokuskan pada upaya bagaimana memperoleh data tentang kemajuan para peserta program, untuk menentukan apakah perilakunya berubah sesuai dengan yang diharapkan atau tidak. Jika ternyata tidak, maka perlu dilakukan perubahan terhadap aktifitas-aktiaitas yang diarahkan untuk mencapai tujuan perubahan perilaku tersebut.

Selama tahap *produk*, penilaian dilakukan untuk menentukan apakah tujuan akhir program tercapai atau tidak. Provus membedakan antara dampak terminal (*immediate outcomes*) dan dampak jangka panjang (*long term-outcomes*). Dengan pemikiran ini ia mendorong evaluator untuk tidak hanya mengevaluasi hasil berupa kinerja program, tetapi lebih dari itu perlu mengadakan studi lanjut sebagai bagian dari evaluasi.

Tahap lainnya yang ditawarkan Provus adalah analisis biaya-manfaat (*cost-benefit analysis*), dimana hasil-hasil yang diperoleh dibandingkan dengan biaya yang dikeluarkan. Analisis ini menjadi sangat urgen dalam keadaan sumber daya (khususnya biaya) pembangunan pendidikan yang sangat terbatas (*limited resources*).

Apapun kesenjangan yang ditemukan melalui evaluasi, Provus menganjurkan agar pemecahan masalah dilakukan secara kooperatif antara evaluator dengan staf pengelola program. Proses kerjasama yang dilakukan antara lain membicarakan tentang: 1) mengapa ada kesenjangan, 2) upaya perbaikan apa yang mungkin dilakukan, 3) upaya mana yang paling baik dilakukan untuk memecahkan masalah yang dihadapi.

Selama tahun 1950-an sampai awal 1960-an, pendekatan berorientasi tujuan sangat kuat digunakan dalam rangka evaluasi dan pengembangan kurikulum. Taba (1962) seorang ahli yang beraliran Tyler (Tylerian) mengemukakan langkah-langkah pengembangan kurikulum sebagai berikut: 1) mendiagnosis kebutuhan, 2) memformulasikan tujuan-tujuan, 3) memilih materi/ isi, 4) mengorganisasikan materi/ isi, 5) memilih pengalaman belajar, 6) megorganisasikan pengalaman belajar, 7) menentukan apa dan bagaimana evaluasi akan dilakukan. Beberapa ahli pendidikan (seperti: Gideonse, 1969; Popham, 1973) telah banyak memelopori penggunaan pendekatan berorientasi tujuan, sementara ahli lainnya (missal: Atkin, 1968) memandang bahwa spesifikasi tujuan berupa perilaku tidak banyak membantu pengembangan atau evaluasi kurikulum.

Pada awal uraian Kegiatan Belajar 1 ini ditegaskan bahwa contoh penerapan pendekatan penilaian berorientasi tujuan adalah apa yang selama ini terjadi dalam penilaian keberhasilan pendidikan di Negara kita melalui tes prestasi belajar. Untuk sekedar mengingatkan, berikut ini akan dipaparkan beberapa pendapat tentang prinsip dan langkah-langkah penting dalam mengelola penilaian keberhasilan program pendidikan dilihat dari ketercapaian pretasi belajar.

Gronlund (1977) mengemukakan beberapa prinsip dalam pengukuran prestasi belajar sebagai berikut:

1. Tes prestasi harus mengukur hasil belajar yang telah dibatasi secara jelas sesuai dengan tujuan instruksional;

2. Tes prestasi harus mengukur suatu sample yang representative dari hasil belajar dan dari materi yang dicakup oleh program instruksional atau pengajaran;
3. Tes prestasi belajar harus berisi butir-butir dengan tipe yang paling cocok guna mengukur hasil belajar yang diinginkan;
4. Tes prestasi belajar harus dirancang sedemikian rupa agar sesuai dengan tujuan penggunaan hasilnya;
5. Reliabilitas tes prestasi belajar harus diusahakan setinggi mungkin dan hasil ukurnya harus ditafsirkan dengan hati-hati;
6. Tes prestasi belajar harus dapat digunakan untuk meningkatkan belajar para anak didik;

Mohammad Noer (1987) mengemukakan langkah-langkah penyusunan tes sebagai berikut:

1. Menentukan tujuan utama penggunaan skor tes;
2. Menentukan tingkahlaku yang menggambarkan konstruk yang hendak diukur atau menentukan domain;
3. Menyiapkan spesifikasi tes, menetapkan proporsi butir yang harus terpusat pada setiap jenis tingkahlaku yang ditentukan pada langkah 2;
4. Menyusun pool awal butir;
5. Mengadakan penelaahan kembali terhadap butir-butir yang diperoleh pada langkah 4 dan melakukan revisi bila perlu;
6. Melaksanakan uji-coba butir pendahuluan dan melakukan revisi bila perlu;
7. Melaksanakan uji-lapangan terhadap butir-butir hasil langkah 6 pada sampel yang besar yang mewakili populasi untuk siapa tes ini dimaksudkan;
8. Menentukan ciri-ciri statistik skor butir, dan apabila perlu, sisihkan butir-butir yang tidak memenuhi kriteria yang ditentukan;
9. Merencanakan dan melaksanakan pengkajian reliabilitas dan validitas untuk bentuk akhir tes;



10. Mengembangkan panduan pengadministrasian, penskoran dan penafsiran skor tes (sebagai misal, siapkan tabel norma, prestasi standar, dan sebagainya).

### **Kelebihan dan Kekurangan Penilaian Berorientasi Tujuan**

Pendekatan penilaian yang berorientasi tujuan ini secara teknologis telah merangsang berkembangnya proses-proses perumusan tujuan secara spesifik serta pengembangan atau penemuan instrument-instrumen maupun prosedur pengukuran yang beragam. Dilihat dari kajian dan literature, pendekatan penilaian berorientasi tujuan sudah lebih banyak dan terarah kepada persoalan bagaimana pendekatan ini diaplikasikan dalam penilaian di kelas, penilaian sekolah, penilaian program sekolah di satu kabupaten, atau lainnya. Oleh karena itu, secara sederhana dapat dikatakan bahwa kelebihan pendekatan ini adalah mudah dipahami, mudah untuk diimplementasikan, dan disepakati banyak pendidik dapat menghasilkan informasi yang relevan dengan misi mereka.

Pendekatan ini juga telah menyebabkan para pendidik merefleksikan dan mengklarifikasi perhatian mereka terhadap pemikiran-pemikiran terdahulu berkaitan dengan ambiguitas tujuan-tujuan pendidikan. Diskusi-diskusi bersama masyarakat tentang tujuan pendidikan yang dianggap paling tepat, dijadikan ajang untuk meningkatkan validitas program pendidikan yang dilakukan. Dengan begitu, akuntabilitas dan legitimasi program yang sudah dirancang menjadi lebih kuat. Sebagai hasil dari perhatian berlebih para ahli terhadap pendekatan ini adalah berkembangnya tes (ujian) dan praktek-praktek pengukuran lainnya yang **broadened unobtrusive and non paper and pencil evidence**.

Disamping manfaat dan keunggulan sebagaimana dipaparkan di atas, pendekatan ini juga mendapatkan beberapa kritik yang sekaligus

menggambarkan sebagai kelemahan dari pendekatan tersebut. Beberapa kritik yang mengemuka adalah (Worten and Sander, 1987):

- 1) komponen penilaian kurang realistis ( lebih memfasilitasi pengukuran dan penilaian ketercapaian tujuan daripada menghasilkan pertimbangan-pertimbangan tentang kebenaran dan merit secara eksplisit)
- 2) kurangnya standar untuk memberi pertimbangan pentingnya diskrepansi yang nampak antara tujuan dan kinerja;
- 3) mengabaikan nilai (*value*) dari tujuan itu sendiri;
- 4) mengabaikan alternative penting yang harus dipertimbangkan dalam perencanaan suatu program pendidikan
- 5) mengabaikan transaksi yang terjadi selama proses atau aktifitas program yang dinilai
- 6) mengabaikan konteks dimana suatu penilaian dilakukan;
- 7) mengabaikan tujuan-tujuan penting lainnya diluar yang tujuan yang dirumuskan (termasuk tercapainya tujuan-tujuan yang tidak diharapkan);
- 8) omit fakta dari nilai suatu program tidak merefleksikan tujuan
- 9) mempromosikan penilaian yang linier dan tidak fleksibel

Dari kesembilan kelemahan tersebut, secara umum dapat disimpulkan bahwa kelemahan pendekatan penilaian berorientasi tujuan dapat menghasilkan suatu **tunnel vision** yang cenderung membatasi efektifitas dan potensi penilaian.

Untuk melihat lebih jauh kelemahan dan keterbatasan penilaian dengan pendekatan berorientasi tujuan, dibawah ini ada beberapa pertanyaan penting untuk direnungkan.

- 1) Siapa sesungguhnya yang menetapkan atau merumuskan tujuan pendidikan selama ini?
- 2) Apakah tujuan-tujuan pendidikan tersebut telah mencakup semua hal yang dianggap penting?

- 3) Apakah semua pihak yang berkepentingan (*stakeholders*) dalam pendidikan sudah sepakat dengan tujuan-tujuan yang ditetapkan tersebut?
- 4) Siapa yang menetapkan criteria keberhasilan atau ketercapaian/ketidakercapaian tujuan tersebut?
- 5) Dan pertanyaan-pertanyaan kritis lain yang perlu dipertimbangkan agar penilaian dengan pendekatan ini tetap memiliki makna.

Kalau kita simak pertanyaan-pertanyaan di atas, nampaknya sumber kelemahan pendekatan penilaian yang berorientasi tujuan bukan terletak pada prosedur pelaksanaan penilaiannya sendiri, tetapi lebih pada rancangan program yang akan dinilai, terutama pada saat penetapan tujuan-tujuan program. Oleh karena itu, penting untuk disadari seorang evaluator bahwa kegiatan penilaian program tidak bisa berdiri sendiri terlepas dari kegiatan perencanaan. Apalagi jika penilaian akan dilakukan dengan mempergunakan penekatan yang berorientasi tujuan.

## **B. MANAGEMENT-ORIENTED APPROACH**

*Management-oriented approach* merupakan salah satu pendekatan dalam penilaian pendidikan yang memfokuskan pada kepentingan manajerial. Oleh karena itu, pendekatan penilaian berorientasi manajemen sangat berarti dalam membantu para pengambil keputusan. Hal ini mengingat pentingnya informasi hasil penilaian sebagai bagian dari pengambilan keputusan yang baik. Artinya, bahwa pengambilan keputusan akan tepat dan berguna jika didasarkan pada informasi-informasi hasil penilaian.

Dalam bidang pendidikan, dengan pendekatan ini dimungkinkan seorang evaluator memberi informasi yang bermanfaat kepada guru, komite sekolah, pengambil keputusan/ birokrasi pendidikan, administrator pendidikan, atau pihak lainnya; sesuai dengan tingkat kewenangan pengambilan keputusan masing-masing. Berdasarkan level kewenangan tersebut menjadi jelas siapa yang akan menjadi pengguna utama hasil-hasil penilaian, bagaimana mereka akan menggunakannya, dan pada aspek-aspek apa mereka akan mengambil keputusan. Aspek-aspek yang menjadi cakupan penilaian dalam pendekatan ini biasanya diklasifikasi berdasarkan komponen system, yakni: input, proses, dan out put.

Untuk mempelajari lebih dalam tentang *management-oriented approach* akan diperkenalkan pada konsep dasar, cara penggunaan, serta kelebihan dan kelemahan pendekatan ini, khususnya dalam penilaian bidang pendidikan.

### **Konsep dasar dan sejarah perkembangan**

Kontributor utama terhadap berkembangnya pendekatan penilaian berorientasi manajemen dalam pendidikan adalah Stufflebeam dan Alkin (1960). Mereka sama-sama menyadari banyaknya kelemahan dari pendekatan penilaian yang selama ini ada. Dengan berlandaskan pada pemikiran-pemikiran terdahulu, seperti Bernard, Mann, Harris, dan Wishburne, mereka memperluas cakrawala dan berfikir sistematis

tentang studi-studi administrative dan pengambilan keputusan pendidikan. Selama tahun 1960-1970-an mereka juga mengembangkan wacana berdasarkan pada teori manajemen (seperti, Braybrook dan Lindblom 1963). Baik Stufflebeam maupun Alkin menjadikan keputusan manajer program **the pivotal organizer** bagi penilaian.

Dalam pendekatan ini, tujuan program bukan menjadi perhatian utama, mereka lebih menekankan pada kebersamaan antara evaluator dan administrator secara erat dalam melakukan penilaian. Mereka bersama-sama mengidentifikasi keputusan-keputusan dimana administrator harus membuat dan kemudian mengumpulkan informasi yang cukup tentang keunggulan dan kelemahan dari setiap alternatif keputusan agar diperoleh keputusan dan pertimbangan yang adil berdasarkan criteria yang spesifik. Oleh karena itu, suksesnya penilaian sangat bergantung pada kualitas tim kerja antara evaluator dan para pengambil keputusan.

Stufflebeam dalam sejarah penilaian ini pada akhirnya mengembangkan apa yang disebut sebagai CIPP Model (*Context, Input, Process, dan Product Model*).

Dilihat dari namanya nampak jelas bahwa pendekatan ini merupakan kerangka kerja penilaian yang bertujuan memberi pelayanan kepada manajer dan administrator program untuk menghadapi empat macam keputusan pendidikan yang berbeda, yaitu:

- 1) *Context Evaluation*, melayani keputusan-keputusan pada level perencanaan. Pada level ini lebih menitikberatkan pada upaya menentukan kebutuhan yang akan dijadikan dasar dalam pengembangan program pendidikan, termasuk perumusan tujuan-tujuan program.

Pada era sentralisasi, dimana kurikulum bahkan sampai pada strategi pembelajaran ditentukan oleh pemerintah pusat, nampaknya evaluasi konteks hanya milik orang pusat. Bagi anda yang di era ini berpengalaman dalam bidang pendidikan luar

sekolah (atau yang disebut dalam Undang-undang No.20 Tahun 2003 tentang Sistem Pendidikan Nasional, disebut jalur non formal dan informal), mungkin penilaian konteks bukan masalah baru. Hal ini dikarenakan program-program pendidikan luar sekolah lebih banyak ditumbuh-kembangkan dari kebutuhan belajar masyarakat yang ingin mengikuti program. Walaupun istilah yang banyak dipergunakan bukan evaluasi konteks tapi sering disebut analisis atau penilaian kebutuhan (*needs analysis* atau *needs assessment*).

Kini di era desentralisasi penilaian konteks ini menjadi penting untuk semua level pengambil keputusan, apakah pengelola pemerintahan pendidikan di tingkat propinsi dan Kabupaten/ Kota atau pengelola satuan pendidikan, baik pada jalur pendidikan formal, in formal maupun non formal. Sebagaimana diketahui bahwa otonomi dan desentralisasi pendidikan bermakna: 1) desentralisasi pemerintahan pendidikan kepada Pemerintah Propinsi maupun Kabupaten/ Kota, serta 2) otonomi pengelolaan satuan pendidikan. Pemberian kewenangan kepada pemerintahan Propinsi maupun Kabupaten/ Kota berimplikasi pada keharusan adanya perencanaan pembangunan pendidikan yang khas sesuai dengan kebutuhan masing-masing daerah otonom. Memperhatikan kewenangan seperti itu, kegiatan evaluasi konteks menjadi amat penting agar perencanaan yang disusun benar-benar berdasarkan pada kebutuhan masyarakat.

Sebagaimana diketahui bahwa peran Departemen Pendidikan Nasional sebagai pemerintah pusat dalam pembangunan pendidikan di era desentralisasi lebih banyak dalam menyusun standar-standar, memberi bimbingan teknis dan penjaminan mutu pendidikan. Anda tahu bukan, dalam hal kurikulum KBK pemerintah pusat hanya menetapkan standar kompetensi dan acuan. Bagaimana kurikulum tersebut diimplementasikan di satu

sekolah, di sekolah-sekolah Kabupaten/ Kota tertentu, di sekolah-sekolah Proponasi tertentu, akan bergantung pada masing-masing kebutuhan. Untuk merancang maupun mengevaluasi rancangan program pada masing-masing level keputusan inilah evaluasi konteks sangat penting artinya bagi keberlangsungan pembangunan pendidikan yang lebih baik.

- 2) *Input Evaluation*, melayani keputusan-keputusan pada kegiatan pengorganisasian. Menentukan sumberdaya yang tersedia, strategi alternative yang perlu dipergunakan dalam program, serta perencanaan yang terbaik bagi pemenuhan kebutuhan, merupakan focus utama penilaian pada level ini.

Jika anda berpengalaman dalam pengelolaan proyek atau paling tidak mengenal dari informasi proyek, unsur input yang paling sering dievaluasi adalah biaya. Apakah biaya yang disediakan untuk suatu proyek mencukupi? Apakah dana yang tersedia itu dialokasikan sesuai komponen-komponen kegiatan proyek yang direncanakan? Pertanyaan-pertanyaan tersebut merupakan contoh fokus evaluasi input. Tentu input program pendidikan bukan satu-satunya biaya, banyak input lain seperti: guru, buku pelajaran, atau lainnya.

- 3) *Process Evaluation*, melayani keputusan-keputusan yang berkaitan dengan implementasi program. Fokus utama pada level ini adalah: a) bagaimana rencana yang sudah ditetapkan dapat dilaksanakan secara baik? b) hambatan-hambatan apa yang dihadapi dan menghambat kesuksesan? c) perbaikan-perbaikan apa yang diperlukan? Untuk bisa menjawab hal-hal tersebut perlu dilakukan evaluasi.

Evaluasi terhadap proses disebut monitoring. Jadi monitoring merupakan upaya melakukan evaluasi terhadap proses pelaksanaan suatu program. Jika di lingkungan kerja anda masih kelihatan kegiatan monitoring sebagai kegiatan 'jalan-jalan'

memotret keadaan lapangan saja, sebenarnya itu belum merupakan monitoring yang sesungguhnya. Dalam tataran pengelolaan proyek ada satu istilah yang seolah-olah tidak bisa dirubah atau diperbaiki saking sudah melembaga, yaitu ME (singkatan dari Monitoring dan Evaluasi). Istilah ini tentu tidak salah, tetapi dalam pemaknaan yang diberikan proyek terhadap kegiatan ME seringkali menjadi kurang bermakna. Salah satu contoh pengertian yang banyak di lingkungan proyek adalah: monitoring merupakan kegiatan mengumpulkan data sedangkan analisis dan interpretasi terhadap data dilakukan dalam evaluasi. Pengertian ini memberi makna seolah-olah M dan E adalah sebuah kegiatan yang berurutan. Padahal kalau anda buka kembali modul dan referensi sebelumnya dinyatakan bahwa setiap kegiatan evaluasi paling sedikit selalu mengandung 3 (tiga) komponen utama, yaitu: data, kriteria, dan *judgment*. Jadi, kalau monitoring merupakan evaluasi terhadap proses maka dalam kegiatan monitoring juga harus terjadi proses analisis dan interpretasi terhadap data dibandingkan dengan kriteria, apakah sesuai, kurang atau melebihi.

Jika ingin proporsional, sebaiknya M-E di diposisikan sebagai dua kegiatan penilaian yang diperlukan dalam rangka pengelolaan proyek. Monitoring dilakukan sebagai penilaian terhadap pelaksanaan kegiatan (*progress*), sedangkan evaluasi merupakan penilaian terhadap hasil proyek.

- 4) *Product evaluation*, melayani keputusan-keputusan yang berkaitan dengan pencapaian program dan kemungkinan perencanaan ulang. Oleh karena itu, focus penilaian ini meliputi:
  - a) hasil-hasil apa yang diperoleh?
  - b) sejauhmana kebutuhan sudah dapat terpenuhi atau berkurang?
  - c) Apa yang harus dilakukan setelah program berjalan seperti itu?



Evaluasi hasil (*product evaluation*) merupakan kegiatan evaluasi yang selama ini banyak dilakukan. Mungkin keadaan ini berakar pada tradisi memahami langkah-langkah pengelolaan (*management*) yang selalu menempatkan kegiatan evaluasi, pengendalian, dan istilah lain yang sejalan, pada akhir kegiatan pengelolaan. Sehingga evaluasi yang banyak dikenal adalah evaluasi yang dilakukan jika suatu program sudah berakhir. Ingat di sekolah kita juga masih seperti itu, misalnya ujian akhir semester (untuk mengevaluasi akhir semester), ujian akhir nasional (untuk mengevaluasi akhir program tiap jenjang pendidikan secara nasional), dll. Evaluasi yang hanya memfokuskan pada hasil biasanya identik dengan pendekatan evaluasi yang berorientasi tujuan sebagaimana dikemukakan pada pembahasan sebelumnya.

Dalam evaluasi hasil, biasanya berkembang beberapa focus, yakni: 1) evaluasi yang memfokuskan pada hasil nyata sesuai tujuan program (*out-put evaluation*), evaluasi terhadap manfaat yang diperoleh dari suatu program (*benefit evaluation*), dan 3) evaluasi terhadap dampak yang ditimbulkan dari program (*impact/ outcome evaluation*). Masing-masing bertujuan mengevaluasi hasil program, hanya hasil yang diperoleh antara lain dipengaruhi oleh jangka waktu yang berbeda. Penilaian dampak biasanya dilakukan beberapa lama setelah program berakhir, lain halnya penilaian manfaat mungkin lebih cepat pelaksanaannya daripada penilaian dampak. Sedangkan penilaian out-put biasanya dilakukan langsung setelah program berakhir. Oleh karena itu, penilaian manfaat dan dampak program lebih rumit dan memerlukan pemikiran yang lebih luas dari sekedar scenario program yang sudah dirancang.

Untuk dapat merancang kegiatan evaluasi pada masing-masing level itu, Stufflebeam mengajukan beberapa langkah yang perlu ditempuh sebagai berikut:

1) Membuat Fokus Evaluasi

- a) Identifikasi, level keputusan apa yang utama akan dilayani dengan evaluasi tersebut, apakah tingkat local, regional, atau nasional?
- b) Untuk tiap level pengambilan keputusan, perkirakan situasi keputusan yang akan dilayani dan gambarkan masing-masing focus, tingkat kekritisannya, waktu, dan alternative komposisinya
- c) Definisikan criteria untuk masing-masing situasi keputusan dengan membuat: (1) spesifikasi variable untuk kepentingan pengukuran dan (2) standar-standar yang akan digunakan dalam memberi pertimbangan alternative.
- d) Definisikan pada area kebijakan mana evaluator harus melakukan seluruh kegiatan penilaian

2) Mengumpulkan Informasi

- a) Spesifikasikan sumber-sumber informasi yang akan digunakan
- b) Spesifikasikan instrument-instrumen dan metode pengumpulan data yang diperlukan
- c) Spesifikasikan prosedur penentuan sample yang akan dilakukan
- d) Spesifikasikan kondisi dan jadwal pelaksanaan pengumpulan data.

3) Mengorganisasikan Informasi

- a) Sediakan format-forat untuk merekap informasi yang akan dikumpulkan
- b) .....**analisis** terhadap informasi yang diperoleh

- 4) Menganalisis Informasi
  - a) Pilih prosedur analisis yang akan digunakan
  - b) Tentukan teknik-teknik yang dipergunakan dalam analisis
- 5) Melaporkan Informasi
  - a) Definisikan siapa yang menjadi sasaran laporan evaluasi ini
  - b) Tentukan teknik penyajian informasi yang akan dipakai
  - c) Tetapkan format laporan
  - d) Buat jadwal pelaporan
- 6) Administrasi Evaluasi
  - a) Rangkum jadwal evaluasi
  - b) Tetapkan keperluan staf dan sumberdaya serta jadwal pengadaanya
  - c) Tentukan teknik-teknik untuk mengadakan keperluan-keperluan pelaksanaan evaluasi
  - d) Menilai rancangan evaluasi yang potensial untuk menyediakan informasi yang valid, reliable, kredibel, tepat waktu, dan pervasive.
  - e) Menentukan dan mejadwalkan updating rancangan evaluasi secara periodic
  - f) Menyediakan anggaran biaya untuk kegiatan evaluasi program secara total.

Alkin (1969) sebagai Direktur Pusat Kajian Evaluasi di UCLA mengembangkan kerangka kerja evaluasi yang hamper parallel dengan pemikiran yang dikemukakan dalam model CIPP. Kerangka kerja evaluasi ini lebih populer disebut sebagai *UCLA Evaluation Model*. Secara garis besar, model evaluasi ini terdiri dari lima tipe, yaitu:

- 1) *System Assesment*, meyediakan informasi yang berkaitan dengan latar belakang suatu system (identik dengan *context evaluation* dalam CIPP Model);

- 2) *Program Planning*, membantub memilih program-program yang lebih khusus dan mungkin efektif dalam memenuhi kebutuhan pendidikan tertentu (hamper sama dengan *input evaluation*);
- 3) *Program Implementation*, menyediakan informasi tentang: apakah program dilaksanakan untuk kelompok sasaran yang tepat sesuai yang diharapkan;
- 4) *Program Improvement*, menyediakan informasi tentang: bagaimana suatu program berfungsi, apakah tujuan-tujuan yang ditetapkan berusaha untuk terus dicapai, apakah ada hal-hal yang tidak diharapkan muncul (hampir sama dengan *process evaluation*);
- 5) *Program Certification*, menyediakan informasi tentang nilai dari suatu program dan kemungkinannya untuk dipergunakan lebih jauh (sama dengan *product evaluation*)

Dari gambaran di atas kedua model menampakkan langkah-langkah yang linier dan berurutan, walalupun menurut para pengembangnya hal itu bisa saja tidak terjadi. Proses evaluasi program bisa dilakukan tanpa adanya evaluasi konteks atau input, misalnya.

Dalam perkembangan sekarang, CIPP model telah menghasilkan panduan-panduan khusus untuk melakukan masing-masing tipe evaluasi. Pada tahun 1977, Stufflebeam mengembangkan prosedur pelaksanaan evaluasi konteks yang dirancang khusus untuk melakukan penilaian kebutuhan program pendidikan. Pada tahun 1972, Reinhard mengembangkan panduan khusus untuk melakukan evaluasi input, yang disebut sebagai *advocate team technique*. Teknik ini digunakan ketika alternative-alternatif yang bisa diterima untuk merancang program baru tidak tersedia atau obvius. Teknik ini menciptakan rancangan alternatif baru, kemudian dievaluasi, dipilih, diadaptasi, atau dikombinasikan hingga tercipta rancangan alternative yang lebih viable untuk program baru. Teknik ini telah sukses digunakan oleh pemerintah federal (Reinhart, 1972) dan school district (Sanders, 1982)

untuk menciptakan pilihan dan panduan rancangan final program-program pendidikan. Sedangkan dalam hal evaluasi proses, telah dikembangkan prosedur evaluasi oleh Cronbach tahun 1963.

Disamping dua model di atas, model discrepancy sebagaimana diuraikan pada Kegiatan Belajar 1, juga bisa dikategorikan sebagai evaluasi berorientasi manajemen. Dalam beberapa aspek, nampak bahwa dalam evaluasi model ini diarahkan juga untuk penyediaan informasi bagi kepentingan pengambilan keputusan para manajer.

Pendekatan evaluasi yang berfokus pada utilitas (*The utilization-focused evaluation approach*) yang dikembangkan Patton (1978) dalam satu aspek juga dapat dipandang sebagai pendekatan pembuatan keputusan. Dia menekankan bahwa proses indentifikasi dan pengorganisasian relevansi antara pengambil keputusan dan pengguna informasi merupakan langkah pertama dalam evaluasi. Dalam pandangan dia, penggunaan hasil temuan evaluasi perlu memperhatikan pemikiran para pengambil keputusan, informasi apa yang dibutuhkan oleh kelompok sasaran, bagaimana informasi tersebut dikumpulkan, dan bagaimana disajikan kepada mereka. Beberapa ahli juga meyakini bahwa pendekatan analisis system termasuk sebagai pendekatan evaluasi, walaupun rentang analisisnya lebih mengarah pada penelitian/riset.

Model CIPP telah banyak digunakan di berbagai belahan Negara Amerika Serikat, baik oleh pemerintah maupun agen-agen swasta. Penggunaan pendekatan evaluasi ini banyak digunakan dalam rangka menjamin akuntabilitas public dari suatu program pendidikan. Stuffelbeam dan Shinkield (1985) menggambarkan pemanfaatan CIPP model dalam dua kepentingan, yakni pembuatan keputusan (orientasi formatif) dan akuntabilitas (orientasi sumatif), sebagai berikut.

	<b>Orientasi Formatif</b>	<b>Orientasi Sumatif</b>
<b>Konteks</b>	Pedoman untuk memilih tujuan dan menentukan prioritas	Mencatat sejauhmana tujuan yang dipilih berdasar pada kebutuhan, kesempatan, dan masalah
<b>Input</b>	Panduan dan masukan untuk memilih strategi program maupun rancangan procedural	Mencatat strategi dan rancangan yang dipilih, serta alasan-alasannya
<b>Proses</b>	Panduan Implementasi	Mencatat proses yang aktual
<b>Produk</b>	Pedoman untuk menghentikan, melanjutkan, memodifikasi atau instalasi program	Merekam ketercapaian prestasi dan perumusan kembali keputusan-keputusan

Model evaluasi CIPP yang dikemukakan oleh Stufflebeam & Shinkfield (1985) adalah sebuah pendekatan evaluasi yang berorientasi pada pengambil keputusan (*a decision oriented evaluation approach structured*) untuk memberikan bantuan kepada administrator atau leader pengambil keputusan. Stufflebeam mengemukakan bahwa hasil evaluasi akan memberikan alternatif pemecahan masalah bagi para pengambil keputusan. Model evaluasi CIPP ini terdiri dari 4 huruf yang diuraikan sebagai berikut:

- a. *Context evaluation to serve planning decision.* Seorang evaluator harus cermat dan tajam memahami konteks evaluasi yang berkaitan dengan merencanakan keputusan, mengidentifikasi kebutuhan, dan merumuskan tujuan program.
- b. *Input Evaluation structuring decision.* Segala sesuatu yang berpengaruh terhadap proses pelaksanaan evaluasi harus disiapkan dengan benar. Input evaluasi ini akan memberikan bantuan agar dapat menata keputusan, menentukan sumber-sumber yang dibutuhkan, mencari berbagai alternatif yang akan dilakukan, menentukan rencana yang matang, membuat strategi yang akan dilakukan dan memperhatikan prosedur kerja dalam mencapainya.

- c. *Process evaluation to serve implementing decision.* Pada evaluasi proses ini berkaitan dengan implementasi suatu program. Ada sejumlah pertanyaan yang harus dijawab dalam proses pelaksanaan evaluasi ini. Misalnya, apakah rencana yang telah dibuat sesuai dengan pelaksanaan di lapangan? Dalam proses pelaksanaan program adakah yang harus diperbaiki? Dengan demikian proses pelaksanaan program dapat dimonitor, diawasi, atau bahkan diperbaiki.
- d. *Product evaluation to serve recycling decision.* Evaluasi hasil digunakan untuk menentukan keputusan apa yang akan dikerjakan berikutnya. Apa manfaat yang dirasakan oleh masyarakat berkaitan dengan program yang digulirkan? Apakah memiliki pengaruh dan dampak dengan adanya program tersebut? Evaluasi hasil berkaitan dengan manfaat dan dampak suatu program setelah dilakukan evaluasi secara seksama. Manfaat model ini untuk pengambilan keputusan (*decision making*) dan bukti pertanggung jawaban (*accountability*) suatu program kepada masyarakat. Tahapan evaluasi dalam model ini yakni penggambaran (*delineating*), perolehan atau temuan (*obtaining*), dan penyediaan (*providing*) bagi para pembuat keputusan.

### **Kelebihan dan Kelemahan Pendekatan Berorientasi Manajemen**

Keunggulan atau kelebihan yang paling menonjol dari pendekatan ini adalah penekanannya pada informasi-informasi yang penting dan dibutuhkan pada setiap level pengambilan keputusan. Dengan demikian informasi yang dikumpulkan sangat terarah pada isu-isu pokok yang memerlukan jawaban pada setiap level dan aspek program, sehingga evaluasi lebih terfokus. House (1982) mendeskripsikan keunggulan pendekatan ini dalam sebuah kalimat sebagai berikut:

“...pendekatan ini telah memberikan masukan yang sangat bernilai terhadap evaluasi, dengan penekanannya pada informasi yang bermanfaat dan penting. Informasi evaluasi telah digunakannya secara berarti. Menghubungkan kegiatan evaluasi dengan pengambilan keputusan telah melandasi bermanfaatnya evaluasi. Secara praktis hal ini telah **shape** evaluasi secara konseptual dengan pertimbangan-

pertimbangan pengambilan keputusan. Walaupun seseorang tidak dapat mendefinisikan secara tepat alternative-alternatif keputusan, tetapi seseorang dapat mengeliminasi sejumlah hal-hal yang tidak relevan”

Pendekatan penilaian berorientasi manajemen telah mengantarkan seorang evaluator atau pendidik untuk tidak menunggu waktu berjalannya program sebelum melakukan evaluasi. Faktanya, seorang pendidik dapat melakukan evaluasi sebelum program dilakukan, misalnya ketika ide program muncul. Pelaksanaan evaluasi yang dilakukan setelah program berjalan saja akan mengakibatkan kerugian investasi sumberdaya dan kesempatan yang besar. Hal ini merupakan keunggulan atau kelebihan lain dari pendekatan penilaian ini.

Hal lain yang dianggap kelebihan pendekatan ini adalah memungkinkannya dilakukan perangkat heuristic secara sederhana melalui pengembangan pertanyaan-pertanyaan dalam setiap level dan aspek program. Dengan demikian pendekatan ini terbuka untuk dapat mengali hal-hal lain yang lebih mendalam dan mudah dijelaskan kepada kelompok sasaran pengguna informasi evaluasi.

Kelemahan yang potensial dari pendekatan ini adalah ketidakmampuan evaluator untuk merespon pertanyaan-pertanyaan atau isu-isu yang bermakna karena takut terjadi pertentangan atau setidaknya tidak pas dengan perhatian para pengambil keputusan yang *nota bene* mengendalikan evaluasi. Berkaitan dengan ini House (1980) mengingatkan:

“Mengapa harus pengambil keputusan, yang seringkali diidentifikasi sebagai administrator program, yang diberi banyak peran? Apakah hal ini tidak menempatkan evaluator sebagai pelayan manajemen puncak dan membuat evaluator sebagai *the hired gun* dari penetapan program? Apakah juga tidak membuat evaluasi secara potensial tidak adil dan mungkin tidak demokratis? Ini adalah kelemahan potensial dari pendekatan penilaian yang berorientasi pengambilan keputusan.”



Untuk memperhatikan peringatan yang disampaikan House, Cronbach dkk. (1980) mengembangkan apa yang disebutnya sebagai *Policy-Shaping Community*, terdiri dari:

- 1) pelayan public seperti kantor-kantor yang bertanggungjawab pada level program dan kebijakan, serta personil yang mengoperasikan program secara actual;
- 2) public, terdiri dari bukan hanya perwakilan rakyat, tetapi juga orang-orang yang berpengaruh, seperti komentator, akademisi/ilmuwan, filosof, gadflies, dan sastrawan.

Dari fakta yang ada, sedikit sekali studi-studi kebijakan yang telah menemukan adanya pengaruh langsung dari policy shaping community, tetapi evaluasi dapat dan terus berjalan mempengaruhi kelompok sasaran ini dari waktu ke waktu. Kebijakan sebagai refleksi nilai-nilai public mungkin nampak sebagai aspek pendidikan yang tidak pernah berakhir dan terus berlanjut dengan berbagai revisi, reformasi, dll. Berkaitan dengan ini, Cronbach mengingatkan bahwa peran pertama yang penting dari seorang evaluator adalah mengilluminate, bukan mendikte keputusan. Peran evaluasi adalah membantu klien untuk memahami kompleksitas isu-isu, bukan memberi jawaban sederhana atas sejumlah pertanyaan.

Kelemahan yang paling akan banyak dihadapi dalam pendekatan ini adalah mahalnya biaya dan kompleksitas ruang lingkup evaluasi. Jika prioritas isu dan pertanyaan-pertanyaan evaluasi tidak ditentukan, penilaian dengan mempergunakan pendekatan ini cenderung akan menjadi rumit dan memerlukan biaya besar. Oleh karena itu, dalam merencanakan evaluasi dengan pendekatan ini seorang evaluator harus mempertimbangkan waktu dan sumberdaya yang tersedia. Jika sumberdaya dan waktu terbatas, sebaiknya evaluasi dilakukan dengan mempertimbangkan pendekatan lain.

Hal lain yang perlu diperhatikan adalah bahwa sejumlah asumsi yang melandasi pendekatan penilaian berorientasi manajemen tidak

selamanya terjamin, misalnya asumsi bahwa: 1) keputusan-keputusan penting dapat diidentifikasi secara mudah, 2) alternative keputusan yang jelas dapat dirumuskan dengan baik. Hal tersebut seringkali justru sulit dilakukan atau disederhanakan dalam pelaksanaannya.

### **C. NATURALISTIC AND PARTICIPANT APPROACH**

Pendekatan naturalistic atau partisipatif dalam penilaian merupakan suatu pendekatan yang secara paradigmatic sangat berbeda dengan pendekatan-pendekatan lain yang diuraikan sebelumnya. Sesuai dengan istilah yang dipergunakan, pada pendekatan ini kegiatan evaluasi diharapkan berjalan natural dan ada keterlibatan (partisipasi) evaluator pada lapangan yang menjadi sasaran evaluasi. Jika pada pendekatan-pendekatan lain peran evaluator cenderung 'diluar' sasaran yang akan dievaluasi, baik sejak perencanaan, pengumpulan informasi, analisis maupun pelaporan, maka pada pendekatan naturalistic-partisipatif seorang evaluator dituntut 'masuk ke dalam' situasi-situasi yang menjadi sasaran evaluasi. Pendekatan ini cocok terutama dalam rangka penilaian proses atau implementasi program.

#### **Konsep dasar dan Sejarah Perkembangan**

Pada awal tahun 1967 beberapa ahli evaluasi mulai melakukan reaksi terhadap apa yang mereka sebut sebagai dominasi pendekatan evaluasi mekanistik dan insensitive pada bidang pendidikan. Pendekatan yang mendominasi ini oleh para ahli diekspresikan sebagai pendekatan yang terlalu menekankan pada : 1) seorang evaluator harus memulai dengan merumuskan dan mengklasifikasi tujuan, 2) merancang system evaluasi secara terurai, 3) mengembangkan instrument-instrumen obyektif yang secara teknis sangat defensible, 4) mempersiapkan laporan-laporan teknis yang panjang. Keempat hal tersebut pada akhirnya akan menempatkan seorang evaluator jauh dari apa yang terjadi secara riil di lapangan pendidikan.

Kritik yang paling diarahkan pada pendekatan tradisional ini adalah bahwa banyak kegiatan evaluasi pendidikan dengan skala besar dilakukan tanpa evaluator menginjakkan kaki berpartisipasi secara langsung di dalam kelas. Banyak praktisi pendidikan mulai mempublikasikan berbagai pertanyaan bahkan mempertanyakan: apakah para evaluator memahami fenomena yang sesungguhnya terjadi di lapangan, yang terjadi di balik angka-angka dalam table dan gambar-gambar yang dilaporkan sebagai hasil evaluasi? Segmen masyarakat pendidikan yang mempertanyakan unsur manusia dalam kompleksitas program pendidikan sebagai factor yang sering diabaikan dalam kegiatan evaluasi makin meningkat. Oleh karena itu, orientasi baru evaluasi telah dilahirkan, dengan penekanan pada pentingnya pengalaman tangan pertama dalam aktifitas dan setting program pendidikan. Pendekatan ini berkembang sangat cepat pada era tahun 1970-1980-an.

Stake (1967) merupakan ahli evaluasi pertama yang melakukan reaksi terhadap pendekatan-pendekatan penilaian pendidikan yang selama ini mendominasi. Paper dia yang berjudul *The Countenance of Educational Evaluation* telah merubah secara dramatis pola pikir evaluasi pada decade-dekade berikutnya. Dalam tulisan-tulisan selanjutnya (Stake, 1975a, 1975b, 1977, 1980), ia menyajikan konsepsi dan prinsip-prinsip yang menjadi arah perubahan pendekatan ini.

Menurut model 'Countenance', penilaian harus mengandung langkah-langkah berikut; menerangkan program; melaporkan keterangan tersebut kepada pihak yang berkepentingan; mendapatkan dan menganalisis 'judgment; melaporkan kembali hasil analisis kepada pelanggan. Seterusnya, model responsif mencadangkan perhatian yang terus menerus oleh penilai dan semua pihak yang terlibat dengan penilaian. Stake (1975) telah menentukan 12 langkah interaksi antara penilai dan pelanggan dalam proses penilaian.

Model evaluasi **Stake** (1967), merupakan analisis proses evaluasi yang membawa dampak yang cukup besar dalam bidang ini, meletakkan dasar yang sederhana namun merupakan konsep yang cukup kuat untuk perkembangan yang lebih jauh dalam bidang

evaluasi. Stake menekankan pada dua jenis operasi yaitu deskripsi (*descriptions*) dan pertimbangan (*judgments*) serta membedakan tiga fase dalam evaluasi program yaitu :

- ✓ Persiapan atau pendahuluan (*antecedents*)
- ✓ Proses/transaksi (*transaction-processes*)
- ✓ Keluaran atau hasil (*outcomes, output*)

Stake menganggap hanya ada dua aktifitas utama dalam kegiatan evaluasi, yaitu: deskripsi dan pertimbangan (*judgment*), yang dikenal sebagai *Two Countenances of Evaluation*. Untuk membantu evaluator dalam mengorganisasikan pengumpulan dan interpretasi data, Stake menciptakan kerangka kerja sebagaimana digambarkan berikut ini.

R A T I O N A L E	DESCRIPTION MATRIX			JUDGMENT MARIX	
	INTENTS	OBSERVATIONS		STANDARDS	JUDGMENT
	congru	ence	ANTE	DENT	c o n t i
	congru	ence	TRANS	TIONS	n g e n
	congru	ence	OUT	MES	c i e s
			CE	AC	CO

Dengan menggunakan kerangka kerja ini, seorang evaluator akan :

- 1) menyajikan latar belakang, justifikasi dan deskripsi dari rasional program (termasuk kebutuhan);
- 2) membuat daftar antecedent yang diharapkan (input, sumberdaya, dan kondisi yang ada), transaksi yang diharapkan (aktifitas dan proses), serta hasil-hasilnya;
- 3) mencatat antecedent, ransaksi, dan hasil-hasil yang terobservasi (termasuk al-hal yang tidak diharapkan);
- 4) menyatakan secara eksplisit standar-standar (criteria, harapan-harapan, kinerja program yang setara) untuk membuat pertimbangan atas antecedent, ransaksi, dan hasil-hasil program;
- 5) mencatat pertimbangan-pertimbangan yang dibuat tentang kondisi-kondisi antecedent, ransaksi, dan hasil.

Seorang evaluator akan menganalisis informasi dalam matrik deskripsi dengan melihat kongruensi antara yang diharapkan dan hasil observasi, serta ketergantungan atau kontingensi antara hasil yang dicapai dengan transaksi dan anteseden maupun ketergantungan transaksi atas anteseden. Pertimbangan akan dibuat dengan menerapkan standar terhadap data deskriptif.

Parllet dan Hamilton (1976) melakukan kritik terhadap dominasi paradigma riset yang berbasis pada ahli agricultural-botani, yang telah **deficient** pada studi-studi program pendidikan inovatif. Atas dasar kritik tersebut ia mengemukakan pendekatan alternative yang disebut sebagai *illuminative evaluation* sebagai tindak lanjut dari paradigma antropologi.

Pendekatan evaluasi iluminatif melibatkan studi intensif tentang program pendidikan secara keseluruhan, seperti rasional, evolusi, operasi, prestasi, dan kesulitan-kesulitan dalam konteks sekolah atau lingkungan belajar. Manfaat pendekatan ini, khususnya pada program yang berskala kecil, adalah untuk meniluminasi masalah-masalah, isu-isu, feature program yang bermakna. Pendekatan ini didasarkan atas paradigma antropologi social, dan dalam beberapa hal juga pada riset

partisipatif dalam Psikiatri dan Sosiologi, yang dianggap tepat bagi manusia/orang. Evaluasi iluminatif ini lebih menekankan pada deskripsi dan interpretasi, bukan pada pengukuran dan prediksi. Tidak ada upaya melakukan manipulasi atau variable control, tetapi lebih menekankan pada kompleksitas konteks pendidikan yang ada dan berusaha untuk memahaminya.

Proses evaluasi yang diajukan Parlett dan Hamilton ini meliputi tiga tahapan pokok, yaitu:

- 1) *Observation*, menjelajahi dan menjadi familier dengan realitas keseharian setting program yang akan dievaluasi;
- 2) *Further Inquiry*, memfokuskan studi dengan melakukan pencarian berlanjut untuk memilih isu-isu;
- 3) *Explanation*, mencari penjelasan-penjelasan dari pola yang terobservasi dan mencari hubungan sebab-akibatnya.

Dengan menekankan pada proses-proses di kelas, informasi subyektif dan pencarian secara naturalistic, pendekatan evaluasi iluminatif sangat bergantung pada data dari observasi, wawancara, kuesioner dan tes, serta sumber-sumber dokumen. Kombinasi triangulasi dari data-data tersebut diperlukan untuk mendapatkan gambaran yang lebih akurat tentang suatu realitas. Oleh karena itu, evaluasi dengan pendekatan seperti ini memerlukan waktu yang banyak di lapangan.

Ahli lain yang memiliki pandangan relative sejalan dengan pendekatan naturalistic-partisipatif adalah Rippey (1973). Ia meragukan sensitifitas dari pendekatan-pendekatan evaluasi yang ada. Untuk itu, ia mengajukan *transactional evaluation* sebagai alternative yang dianggap tepat. Pendekatan ini lebih menekankan pada perubahan-perubahan yang sedang berjalan daripada hasil-hasil dari sebuah system aktifitas. Evaluasi transaksional ini merupakan strategi tepat untuk mengelola ketidakberfungsian yang terjadi dalam suatu organisasi di tengah-tengah upaya perubahan yang dilakukan (misalnya, ketika suatu

sekolah memperkenalkan program baru atau melakukan revisi kurikulum).

Adapun langkah-langkah pokok yang dilakukan dalam evaluasi transaksional ini, meliputi:

- 1) Initial, mengidentifikasi titik masalah (*trouble spots*) dengan evaluasi yang neutral;
- 2) Instrumentation, mengumpulkan data dalam berbagai pertemuan kelompok-kelompok berkepentingan;
- 3) Program development, redefinisi untuk merefleksikan tujuan dan nilai-nilai yang merupakan kesepakatan kelompok;
- 4) Program monitoring, kesepakatan kelompok untuk mengimplementasikan dan memonitor program baru;
- 5) Recycling, proses daur ulang sebagai emerge konflik baru.

Rippey mengajukan penggunaan evaluasi berkelanjutan terhadap perubahan, baik perubahan proponen maupun oponen, untuk memecahkan konflik yang berkaitan dengan konsekuensi-konsekuensi dari inovasi program pendidikan, baik konsekuensi yang diharapkan maupun yang tidak diharapkan.

McDonald (1974, 1976) melihat pendekatan-pendekatan evaluasi yang telah ada **misuses** informasi evaluasi untuk kepentingan politik yang dipertanyakan. Pilihan yang ditawarkan adalah *democratic evaluation* yang dirancang untuk melindungi hak-hak dan kebutuhan informasi seluruh masyarakat yang terlibat. Ia membedakan adanya tiga tipe evaluasi, jika dilihat dari pemilihan peran, tujuan, kelompok sasaran, teknik, dan isu-isu, yaitu:

- 1) *Bureaucratic Evaluation*, dimana agen birokrasi yang mensponsori evaluasi, mengendalikan informasi evaluasi dan laporan hasil evaluasi, bukan evaluator;

- 2) *Autocratic evaluation*, dimana evaluator sebagai pemilik evaluasi dan melaporkan temuan-temuannya kepada agen sponsor atau jurnal ilmiah;
- 3) *Democratic evaluation*, dimana evaluator menampilkan informasi evaluasi kepada seluruh lapisan masyarakat, baik sponsor maupun partisipan, dengan tanpa adanya klaim khusus terhadap temuan-temuan evaluasi tersebut dari pihak sponsor.

Selama tahun 1970-an, Stake mulai memperluas tulisan awalnya tentang evaluasi lebih mengarah pada realm partisipatif-naturalistik, dengan megemukakan apa yang ia sebut sebagai *responsive evaluation* (penilaian responsif). Fokus dari penilaian responsif dialamatkan pada perhatian dan isu-isu stakeholders.

Guba (1969) adalah ahli evaluasi yang mendiskusikan kegagalan evaluasi pendidikan dan selanjutnya mencari alternatif pendekatan evaluasi yang lebih rasionalistik. Pada tahun 1981, Guba dan Lincoln melakukan kajian terhadap pendekatan-pendekatan utama penilaian pendidikan yang sering digunakan. Mereka menolak semuanya kecuali pemikiran Stake tentang *responsive evaluation* (penilaian responsif) yang dianggap sejalan dengan pendekatan naturalistic sebagai payung dari semua alternatif pendekatan penilaian pendidikan.

Dalam tulisan-tulisannya tentang pengalaman evaluasi di lapangan tahun 1975-1980, Patton menambah literature tentang evaluasi yang berorientasi partisipan. Beberapa ahli lain yang mengembangkan pemikiran sejalan dengan pendekatan naturalistic-partisipatif antara lain: Kelly (1975), McDonald and Wolker (1977), Kemmis (1977), Hamilton (1976), Stenhouse (1975), Bullock (1982), Fetterman (1984), dan Simon (1984).

Belajar dari berbagai model penilaian yang termasuk pendekatan naturalistic-partisipatif, nampak ada beberapa karakteristik utama dari pendekatan ini, yaitu:



- 1) Berdasar pada alasan-alasan induktif. Pemahaman isu, peristiwa, atau suatu proses dating dari observasi dan penemuan berbasis akar rumput.
- 2) Menggunakan *multiplicity* data. Pemaaman atas suatu persoalan didasarkan pada asimilasi data dari sejumlah sumber. Representasi gejala-gejala yang dievaluasi, baik yang subyektif maupun obyektif, kuantitatif maupun kualitatif digunakan
- 3) Tidak disandarkan pada rencana yang standar. Proses eveluasi berjalan sebagaimana pengalaman yang diperoleh partisipan dalam semua aktifitas program.
- 4) Mencatat realitas yang *multiple* ketimbang *single*. Seseorang melihat sesuatu dan menginterpretasikannya dengan cara yang berbeda-beda. Tidak seorangpun megetahui segala sesuatu yang terjadi di sekolah, dan tidak satu perspektif pun yang diterima sebagai kebenaran. Karena hanya orang tersebutlah yang paling tahu benar apa yang dia alami, semua perspektif diterima sebagai sesuatu yang benar, dan tugas utama evaluator adalah menangkap realitas ini semua dan potertnya tanpa menyederhanakn kompleksitas dunia pendidikan.

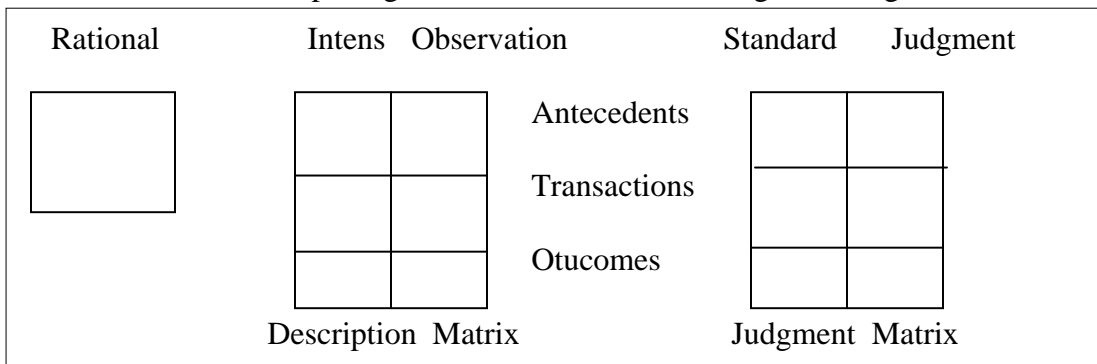
### **3. Model Evaluasi Brinkerhoff**

*Brinkerhoff & Cs.* (1983) mengemukakan tiga golongan evaluasi yang disusun berdasarkan penggabungan elemen-elemen yang sama, seperti evaluator-evaluator lain, namun dalam komposisi dan versi mereka sendiri sebagai berikut :

- a. *Fixed vs Emergent Evaluation Design*. Dapatkah masalah evaluasi dan kriteria akhirnya dipertemukan? Apabila demikian, apakah itu suatu keharusan? Belum lengkap penjelasannya
- b. *Formative vs Summative Evaluation*. Apakah evaluasi akan dipakai untuk perbaikan atau untuk melaporkan kegunaan atau manfaat suatu program? Atau keduanya?
- c. *Experimental and Quasi Experimental Design vs Natural/ Unobtrusive Inquiry*. Apakah evaluasi akan melibatkan intervensi ke dalam kegiatan program/mencoba

memanipulasi kondisi, orang diperlakukan, variabel dipengaruhi dan sebagainya, atau hanya diamati, atau keduanya?

Model stake tersebut dapat digambarkan dalam bentuk diagram sebagai berikut :



*Descriptions matrix* menunjukkan *Intents* (goal=tujuan) dan *observations* (*effect*=akibat) atau yang sebenarnya terjadi. *Judgment* berhubungan dengan standar (tolak ukur = kriteria)/dan *judgment* (pertimbangan). Stake menegaskan bahwa ketika kita menimbang-nimbang di dalam menilai suatu program pendidikan, kita tentu melakukan perbandingan relatif (antara satu program dengan standard).

Model ini menekankan kepada evaluator agar membuat keputusan/penilaian tentang program yang sedang dievaluasi secara benar, akurat dan lengkap. Stake menunjukkan bahwa description disatu pihak berbeda dengan pertimbangan (judgment) atau menilai. Di dalam model ini data tentang Antecedent (input), Transaction (process) dan Outcomes (Product) data tidak hanya dibandingkan untuk menentukan kesenjangan antara yang diperoleh dengan yang diharapkan, tetapi juga dibandingkan dengan standar yang mutlak agar diketahui dengan jelas kemanfaatan kegiatan di dalam suatu program.

### 5. Model Evaluasi Metfessel dan Michael

*Metfessel dan Michael* (1967), dapat digunakan oleh guru dan evaluator program. Dalam strategi model Metfessel dan Michael terdapat delapan langkah yaitu

- a. Keterlibatan masyarakat (*envalvement of the community*) yakni : orangtua, ahli-ahli pendidikan dan peserta didik
- b. Pengembangan tujuan dan memilih tujuan menurut skala prioritas.

- c. Menterjemahkan tujuan menjadi bentuk tingkah laku dan mengembangkan pengajaran.
- d. Mengembangkan metode untuk mengukur dan mengevaluasi pencapaian tujuan.
- e. Menyusun dan mengadministrasi ukuran untuk mengevaluasi pencapaian tujuan
- f. Menganalisis hasil pengukuran
- g. Menginterpretasi dan mengevaluasi data
- h. Menyusun rekomendasi untuk mengembangkan pengajaran

Metode ini dilengkapi dengan instrumen pengumpulan data, lengkap dengan kriteria-kriteria yang dapat digunakan untuk mengevaluasi sebuah proyek/kegiatan program. Seperangkat instrumen tersebut meliputi : tes, angket, check list, dan sebagainya serta cara-cara lain untuk menghimpun data penunjang.

## DAFTAR PUSTAKA

Aiken, LR, *Psychological Testing and Assesment (Ninth Edition)*, Boston: Allyn and Bacon, 1997.

Anita E. Woolfolk dan Lorraine McCune-Nicolich, (1984), *Educational Psychology for Teachers*,(New Jersey: Prentice Hall Inc.)

Anne Anastasi dan Susana Urbina, (1997), *Psychological Testing*, (New Jersey: Prentice-Hall Internatinal,Inc.)

Blaine R. Worthen and James R. Sanders, (1987), *Educational Evaluation: Alternative Approaches and Practical Guidelines*, (New York: Longman Inc.)

Daniel L. Stufflebeam, et.al., (1977), *Educational Evaluation and Decision Making* (Illino-is: F.E Peacock Publishers Inc.)

Mohammad Noer, (1987), *Pengantar Teori Tes*, (Jakarta: Depdiknas).

Norman E. Gronlund dan Robert L. Linn, (1985) , *Measurement and evaluation in Teaching*, (New York: Macmillan Publishing Company)

Raka Joni, T, (1981), *Penilaian Program Pendidikan* (Jakarta: P3G Depdikbud)

Saifudin Azwar, (2000), *Tes Prestasi: Fungsi dan Pengembangan Pengukuran Prestasi Belajar*, (Yogyakarta: Pustaka Pelajar).

Stephen Issac dan William B. Michael, (1982), *Handbook in Research and Evaluation: For Education and the Bahavioral Sciences*, (Sandiego-California: Edits Publish-ers).

Sumadi Suryabrata, (2000), *Pengembangan Alat Ukur Psikologis*, (Yogyakarta: Andi).