

Evaluasi Program Pendidikan

Dr. Heri Retnawati



PENDAHULUAN

Pendidikan diyakini oleh masyarakat dan pemerintah untuk meningkatkan kualitas sumber daya manusia. Berbagai program dilakukan pemerintah untuk meningkatkan kualitas pendidikan, baik melalui pemberian bantuan dana, sarana, dan prasarana, peningkatan kualitas proses pendidikan, maupun peningkatan kualitas pendidik dan tenaga kependidikan, maupun peningkatan kualitas peserta didik. Untuk mengetahui keberhasilan program pendidikan yang dilaksanakan, diperlukan suatu evaluasi, yang disebut dengan evaluasi program. Karena khusus mengevaluasi program pendidikan, maka sering disebut dengan evaluasi program pendidikan.

Pada modul ini akan dikaji evaluasi program pendidikan, mulai dari definisi, ruang lingkup evaluasi termasuk kaitan antara tujuan evaluasi dan tujuan program, serta perbedaan evaluasi program pendidikan dengan penelitian pendidikan.

Setelah mempelajari modul ini, Anda secara umum dapat menjelaskan evaluasi program pendidikan, sedangkan secara khusus Anda dapat menjelaskan kembali pengertian evaluasi program pendidikan, menjelaskan ruang lingkup evaluasi pendidikan, serta dapat membedakan evaluasi pendidikan dengan penelitian pendidikan.

Untuk mempelajari modul ini dengan mudah, Anda dapat mengingat atau membaca kembali materi-materi tentang penelitian, karena ada beberapa kesamaan antara evaluasi program dengan penelitian pendidikan. Modul ini bermanfaat bagi mahasiswa yang mempelajari evaluasi program, dengan terlebih dahulu memahami pengertian program pendidikan dan evaluasi program pendidikan.

Modul ini terdiri dari dua kegiatan pembelajaran, yang dapat dipelajari secara berurutan. Pada Kegiatan Belajar 1 disajikan pengertian dan program evaluasi, dan pada Kegiatan Belajar 2 disajikan sejarah evaluasi dan perbedaan antara evaluasi dengan penelitian. Selanjutnya Anda dapat

memperoleh inti dari rangkuman, dan mencoba mengerjakan latihan, dan membandingkan jawaban Anda dengan petunjuk pengerjaan. Hasilnya dapat didiskusikan dengan mahasiswa lain ataupun dosen tutor.

KEGIATAN BELAJAR 1**Pengertian Program dan Evaluasi****A. PROGRAM PENDIDIKAN**

Acap kali kita mendengar tentang istilah program. Apakah sebenarnya program itu? Program merupakan serangkaian kegiatan yang mempunyai tujuan-tujuan tertentu. Karena serangkaian, berarti kegiatannya lebih dari satu. Kegiatan yang satu dengan kegiatan lain merupakan suatu rangkaian yang saling berhubungan, yang masing-masing mempunyai tujuan tertentu. Tujuan-tujuan tiap kegiatan ini merujuk pada satu tujuan pokok dari program yang selanjutnya dilaksanakan.

Demikian pula halnya dengan program pendidikan. Program pendidikan ini merupakan serangkaian kegiatan pendidikan yang mempunyai tujuan-tujuan tertentu. Sebagai contoh misalnya program pelatihan guru. Program pelatihan ini tentunya terdiri dari rangkaian kegiatan yang saling terkait. Kegiatan-kegiatan tersebut di antaranya persiapan pelatihan, proses pelatihan, dan evaluasi hasil pelatihan. Jika dirinci lebih jauh, persiapan pelatihan meliputi penentuan kurikulum pelatihan, pemilihan peserta pelatihan, penentuan tempat dan waktu, persiapan sarana dan prasarana.

Contoh-contoh program pendidikan lain, di antaranya program:

1. Rintisan Sekolah Berstandar Internasional;
2. Sertifikasi Guru;
3. Seleksi Masuk Sekolah dan Perguruan Tinggi;
4. Peluncuran Dana Bantuan Operasional Sekolah (BOS);
5. Beasiswa Unggulan untuk Pegawai;
6. Beasiswa Siswa/Mahasiswa Berprestasi, dan lain-lain;
7. Pendidikan Bahasa Inggris;
8. Pelaksanaan Kurikulum KTSP;
9. pelaksanaan Kurikulum 2013;
10. Pelaksanaan Sekolah Gratis;
11. Kuliah Kerja Lapangan;
12. Pengentasan Buta Aksara;
13. Pendidikan Anak Usia Dini;
14. Pendidikan Dasar 9 Tahun.

Masing-masing program pendidikan tersebut dilaksanakan berdasarkan aturan tertentu, dasar hukum tertentu, dan terdiri atas serangkaian program tertentu yang masing-masing mempunyai tujuan. Contoh detail dari program pendidikan disajikan sebagai berikut.

Contoh 1:

Program Rintisan Sekolah Berstandar Internasional

Dalam rangka memenuhi kebutuhan nasional akan sumber daya manusia yang unggul dan dapat bersaing secara internasional, pemerintah telah mencanangkan program Rintisan Sekolah Bertaraf Internasional atau RSBI. RSBI ini merupakan calon dari Sekolah Bertaraf Internasional (SBI). RSBI (Rintisan Sekolah Berstandar Internasional), atau yang sering kita kenal dengan SBI adalah suatu program pendidikan yang ditetapkan oleh Menteri Pendidikan Nasional berdasarkan Undang-Undang No. 20 Tahun 2003 pasal 50 ayat 3, yang menyatakan bahwa Pemerintah dan/atau Pemerintah Daerah menyelenggarakan sekurang-kurangnya satu pendidikan pada semua jenjang pendidikan untuk dikembangkan menjadi satuan pendidikan yang bertaraf internasional.

RSBI didefinisikan sebagai Sekolah Bertaraf Internasional adalah satuan pendidikan yang diselenggarakan dengan menggunakan Standar Nasional Pendidikan (SNP) dan diperkaya dengan standar salah satu negara anggota *Organization for Economic Co-operation and Development* (OECD) dan/atau negara maju lainnya (Panduan Penyelenggaraan program Rintisan SMA bertaraf Internasional, 2009:9). Menurut Pedoman Penjamin Mutu Sekolah/ Madrasah Bertaraf Internasional pada Jenjang Pendidikan Dasar dan Menengah, Sekolah/Madrasah Bertaraf Internasional merupakan "Sekolah/Madrasah yang sudah memenuhi seluruh Standar Nasional Pendidikan dan diperkaya dengan mengacu pada standar pendidikan salah satu negara anggota *Organization for Economic Co-operation and Development* dan/atau negara maju lainnya yang mempunyai keunggulan tertentu dalam bidang pendidikan sehingga memiliki daya saing di forum internasional". Pada prinsipnya, Sekolah/Madrasah Bertaraf Internasional harus bisa memberikan jaminan mutu pendidikan dengan standar yang lebih tinggi dari Standar Nasional Pendidikan (Pedoman Penjamin Mutu Sekolah/ Madrasah Bertaraf Internasional pada Jenjang Pendidikan Dasar dan Menengah, 2007:12).

Rintisan Sekolah Bertaraf Internasional adalah realisasi Undang-Undang Nomor 20 tahun 2003 tentang Sistem Pendidikan Nasional (UUSPN 20/2003) Pasal 50 Ayat 3 yang menyebutkan bahwa Pemerintah dan/atau Pemerintah Daerah menyelenggarakan sekurang-kurangnya satu satuan pendidikan pada semua jenjang pendidikan untuk dikembangkan menjadi satuan pendidikan bertaraf internasional. Adapun tujuan umum pengembangan program Rintisan Sekolah Bertaraf Internasional selain untuk meningkatkan mutu pendidikan nasional juga sebagai sarana yang memberikan peluang baik kepada siswa maupun sekolah yang berpotensi untuk mencapai kualitas dan prestasi bertaraf nasional dan internasional.

Mengingat **Rintisan Sekolah Bertaraf Internasional (RSBI)** nanti diharapkan menjadi **Sekolah Bertaraf Internasional (SBI)**, yang merupakan sebuah jenjang sekolah nasional di Indonesia dengan standar mutu internasional. Standar Kompetensi Lulusan, Kurikulum, Proses Belajar Mengajar, Sumber Daya Manusia, Fasilitas, Manajemen, Pembiayaan, dan Penilaian Standar Internasional. Demikian pula standar penggunaan bahasa pengantar dalam pendidikan, yakni bahasa Indonesia dan bahasa Inggris.

Standar internasional yang dimaksudkan meliputi proses pembelajaran yang diharapkan mengacu pada proses pembelajaran kelas internasional. Pada proses belajar mengajar di sekolah ini menekankan pengembangan daya kreasi, inovasi, dan eksperimentasi untuk memacu ide-ide baru yang belum pernah ada. Pembelajaran yang digunakan menggunakan bahasa pengantar bahasa asing, yakni bahasa Arab dan bahasa Inggris misalnya. Sistem penilaian yang digunakan misalnya mengacu pada penilaian internasional, minimal mengacu pada penilaian dari sekolah mitra di luar negeri. Hal ini mengakibatkan, lulusan diharapkan dapat dengan mudah melanjutkan ke perguruan tinggi lain di luar negeri.

Disarikan dari berbagai sumber.

Contoh 2:

Program Bantuan Operasional Sekolah (BOS)

Kebijakan pembangunan pendidikan dalam kurun waktu 2004-2009 berorientasi pada peningkatan akses rakyat terhadap pendidikan yang lebih berkualitas. Implementasinya melalui pelaksanaan program Wajib Belajar Pendidikan Dasar 9 Tahun dan pemberian akses yang lebih besar kepada kelompok masyarakat yang selama ini kurang terjangkau layanan pendidikan. Kenaikan harga BBM yang cukup tinggi, menurunkan kemampuan daya beli

penduduk miskin, selanjutnya dapat menghambat upaya penuntasan Program Wajib Belajar Dasar 9 Tahun karena penduduk miskin semakin sulit memenuhi kebutuhan biaya pendidikan. Oleh sebab itu, Program PKPS - BBM (Program Kompensasi Pengurangan Subsidi Bahan Bakar Minyak) bidang pendidikan perlu dilanjutkan. Hal tersebut mendasari program pemerintah tentang pemberian Bantuan Operasional Sekolah (BOS). Untuk pengelolaan BOS dilakukan Dinas Pendidikan, sedangkan Bantuan Operasional Madrasah adalah BOS dikelola Departemen Agama.

Program bantuan operasional sekolah bertujuan untuk membebaskan biaya pendidikan bagi siswa tidak mampu dan meringankan bagi siswa yang lain, agar mereka memperoleh layanan pendidikan dasar yang lebih bermutu sampai tamat dalam rangka penuntasan Wajib Belajar 9 Tahun

Sasaran program BOS adalah sekolah setingkat SD dan SMP, baik negeri maupun swasta di seluruh Propinsi di Indonesia.

Besar dana BOS yang diterima oleh sekolah/madrasah dihitung berdasarkan jumlah siswa dengan ketentuan:

1. SD/MI/SDLB/Salafiyah/Sekolah agama non-Islam setara SD sebesar **Rp254.000/siswa/tahun;**
2. SMP/MTs/SMPLB/Salafiyah/Sekolah agama non-Islam setara SMP sebesar **Rp354.000/siswa/tahun.**

Besar dana BOS Buku untuk jenjang SD dan SMP sebesar Rp22.000 hanya satu kali dalam satu tahun.

Dana bos diutamakan digunakan untuk:

- a. pembiayaan seluruh kegiatan dalam rangka penerimaan siswa baru: biaya pendaftaran, penggandaan formulir, administrasi pendaftaran, dan pendaftaran ulang, serta kegiatan lain yang berkaitan langsung dengan kegiatan tersebut;
- b. pembelian buku teks pelajaran (di luar buku yang telah dibeli dari dana BOS Buku) dan buku referensi untuk dikoleksi di perpustakaan;
- c. pembiayaan kegiatan pembelajaran remedial, pembelajaran pengayaan, olahraga, kesenian, karya ilmiah remaja, pramuka, palang merah remaja, dan sejenisnya;
- d. pembiayaan ulangan harian, ulangan umum, ujian sekolah, dan laporan hasil belajar siswa;

- e. pembelian bahan-bahan habis pakai: buku tulis, kapur tulis, pensil, bahan praktikum, buku induk siswa, buku inventaris, langganan koran, kopi, teh, dan gula untuk kebutuhan sehari-hari di sekolah;
- f. pembiayaan langganan daya dan jasa: listrik, air, telepon, termasuk untuk pemasangan baru jika sudah ada jaringan di sekitar sekolah;
- g. pembiayaan perawatan sekolah: pengecatan, perbaikan atap bocor, perbaikan pintu dan jendela, perbaikan mebel air, perbaikan sanitasi sekolah, serta perawatan fasilitas sekolah lainnya;
- h. pembayaran honorarium guru honorer dan tenaga kependidikan honorer;
- i. pengembangan profesi guru: pelatihan, KKG/MGMP dan KKKS/MKKS;
- j. pemberian biaya transportasi bagi siswa miskin yang menghadapi masalah biaya transportasi dari dan ke sekolah;
- k. pembiayaan pengelolaan BOS: alat tulis kantor (ATK), penggandaan, surat menyurat dan penyusunan laporan;
- l. khusus untuk pesantren dan sekolah keagamaan non-Islam, dana BOS dapat digunakan untuk biaya asrama/pondokan dan membeli peralatan ibadah;
- m. bila seluruh komponen tersebut di atas telah terpenuhi pendanaannya dari BOS dan masih terdapat sisa dana maka sisa dana BOS tersebut dapat digunakan untuk membeli alat peraga, media pembelajaran, dan mebelair sekolah.

Penggunaan dan BOS untuk transportasi dan uang lelah bagi guru PNS diperbolehkan hanya dalam rangka penyelenggaraan suatu kegiatan sekolah selain kewajiban jam mengajar, besarnya mengikuti batas kewajaran yang ditetapkan pemerintah daerah.

Dana bos tidak boleh digunakan untuk:

- a. disimpan dalam jangka waktu lama dengan maksud dibungakan;
- b. dipinjamkan kepada pihak lain;
- c. membiayai kegiatan yang tidak menjadi prioritas sekolah dan memerlukan biaya besar, misal studi banding, studi *tour*, dan sejenisnya;
- d. membayar bonus, transportasi, atau pakaian yang tidak berkaitan dengan kepentingan siswa;
- e. digunakan untuk rehabilitasi sedang dan berat;
- f. membangun gedung/ruang baru;

- g. membeli bahan/peralatan yang tidak mendukung proses pembelajaran;
- h. menanam saham;
- i. membiayai segala jenis kegiatan yang telah dibiayai dari sumber dana pemerintah pusat atau daerah.

1. Ketentuan yang Harus Diikuti Sekolah Penerima Bos

Sekolah yang menyatakan menerima BOS dibagi menjadi 2 (dua) kelompok, dengan hak dan kewajiban berikut ini.

- a. Apabila di sekolah tersebut terdapat siswa miskin, maka sekolah diwajibkan membebaskan segala jenis pungutan/sumbangan/iuran seluruh siswa miskin. Sisa dana BOS (bila masih ada) digunakan untuk menyubsidi siswa lainnya. Dengan demikian, sekolah tersebut menyelenggarakan pendidikan gratis terbatas. Bila seluruh siswa tergolong miskin dan atau bila dana BOS cukup untuk membiayai seluruh kebutuhan sekolah maka otomatis sekolah tersebut menyelenggarakan pendidikan gratis.
- b. Bagi sekolah yang tidak mempunyai siswa miskin, maka dana BOS digunakan untuk menyubsidi seluruh siswa sehingga dapat mengurangi pungutan/sumbangan/iuran yang dibebankan kepada orang tua siswa, minimum senilai dana BOS yang diterima sekolah.

2. Pengawasan Program Bos

Pengawasan terhadap pelaksanaan program BOS dilakukan oleh lembaga pengawas yang berkompeten antara lain oleh BPK, BPKP, Inspektorat Jenderal serta BAWASDA Propinsi dan kota.

Lembaga tersebut bertanggung jawab untuk melakukan audit terhadap kinerja dan administrasi keuangan

Selain itu dalam rangka transparansi pengelolaan dana BOS, program ini juga dapat dimonitor oleh:

- a. unsur masyarakat;
- b. unit-unit pengaduan masyarakat yang terdapat di sekolah, kota, propinsi dan pusat.

Lembaga tersebut dapat melakukan monitoring dalam rangka pelaksanaan program BOS di sekolah, tetapi tidak melakukan audit. Apabila terdapat indikasi penyimpangan dalam pengelolaan BOS, diharap agar segera dilaporkan kepada lembaga yang berwenang.

Sumber: <http://mediainfokota.jogjakota.go.id/>

Pelaksanaan program pendidikan dapat ditinjau dari berbagai sudut pandang. Sudut pandang tersebut di antaranya dari pemerintah selaku pembuat kebijakan, dari masyarakat sebagai pengguna, dari pendidik, misalnya ditinjau dari sisi efektivitas program, kebermanfaatan program, hasil dan dampak program, dan lain-lain. Namun, dari berbagai sudut pandang tersebut, satu hal yang menjadi kata kunci yakni harapan akan perubahan dan perkembangan ke arah yang lebih baik. Agar memenuhi harapan tersebut, kegiatan pemantauan dan evaluasi program perlu dilakukan secara objektif, *reliabel*, dan menghasilkan laporan yang bermanfaat dalam rangka perbaikan dan membuat keputusan yang lebih baik.

B. DEFINISI EVALUASI

Sebelum sampai pada definisi evaluasi, perlu didefinisikan dulu pengukuran dan penilaian, yang merupakan hal-hal yang erat kaitannya dengan evaluasi. Dalam pelaksanaan program, ada beberapa kegiatan yang saling terkait. Kegiatan tersebut adalah pengukuran, penilaian, dan evaluasi. Menurut Griffin dan Nix (dalam Djemari Mardapi, 2008), pengukuran, penilaian, dan evaluasi merupakan suatu hierarki. Pengukuran membandingkan hasil pengamatan dengan suatu kriteria yang dianggap lebih baku. Penilaian merupakan upaya menjelaskan hasil pengukuran, sedangkan evaluasi merupakan penetapan nilai-nilai atau implikasi suatu perilaku. Sifat hierarki ini menunjukkan bahwa kegiatan evaluasi melibatkan pengukuran dan penilaian.

C. PENGUKURAN

Pengukuran (*measurement*), Penilaian (*Assesment*) dan evaluasi (*Evaluation*) merupakan 3 hal yang saling berkaitan. Pengukuran dipelajari tersendiri dalam teori pengukuran. Allen dan Yen (1979) mengemukakan bahwa *measurement theory is a brach of applied statistics that attempt to describe, categorisizes, and evaluate the the quality of measurement, improve the usefulness, accuracy, and meaningfulness of measurements, and propose method for developing new and better measurement instruments*". Berdasarkan pendapat ini, diperoleh bahwa teori pengukuran merupakan cabang dari statistik terapan yang berusaha mendeskripsikan, membuat kategorisasi, dan mengevaluasi kualitas instrumen, meningkatkan kegunaan,

ketepatan, dan keberartian pengukuran serta mengajukan instrumen pengukuran baru dan lebih baik.

Dengan pengukuran, suatu gejala dapat di kuantifikasi. Sebagai contoh tinggi badan. Mendeskripsikan tinggi badan biasanya dilakukan dengan mengukur tinggi terlebih dahulu, dengan membandingkan tinggi badan dengan ukuran panjang yang telah terstandar. Demikian pula objek pengukuran lain yang tidak nampak. Sebagai contoh misalnya kemampuan seseorang sebagai hasil belajar. Hasil belajar sendiri tidak dapat di kuantifikasi secara langsung. Dengan menggunakan suatu instrumen, misalnya instrumen yang mengukur kemampuan, dapat diperoleh skor yang dapat berupa angka. Pada suatu tes matematika misalnya. Dengan perangkat tes matematika, seseorang dapat dilihat kemampuan matematikanya. Kemudian kemampuan matematika ini diskor, baik dengan skala 0-10 atau 0-100. Berdasarkan contoh ini, dapat dikatakan bahwa pengukuran identik dengan kuantifikasi sementara dari sesuatu, dengan membandingkannya dengan ukuran yang lebih standar.

Ahli lain menyatakan bahwa upaya mendeskripsikan karakteristik atau sifat sesuatu dalam bentuk angka, di antaranya adalah Ebel dan Frisbie. Menurut Ebel dan Frisbie (1986) pengukuran merupakan proses mendapatkan angka atau skor individu dengan instrumen tes atau nontes tertentu. Penentuan angka ini dapat dilakukan secara sistematis. Penentuan angka dalam suatu pengukuran merupakan upaya untuk menggambarkan suatu objek, baik objek yang nampak maupun yang tidak nampak.

Pengukuran dapat didefinisikan sebagai *the process by which information about the attributes or characteristics of thing are determined and differentiated* (Oriondo & Antonio: 1998). Guilford mendefinisi pengukuran dengan *assigning numbers to, or quantifying, things according to a set of rules* (Griffin & Nix: 1991). Pengukuran dinyatakan sebagai proses penetapan angka terhadap individu atau karakteristiknya menurut aturan tertentu (Ebel & Frisbie: 1986). Allen & Yen (2002) mendefinisikan pengukuran sebagai penetapan angka dengan cara yang sistematis untuk menyatakan keadaan individu.

Berdasarkan pendapat-pendapat di atas, dapat diperoleh kesimpulan bahwa pengukuran merupakan upaya mendeskripsikan sesuatu dalam bentuk angka atau merupakan upaya menguantifikasi sesuatu, baik objek yang tampak maupun yang tidak tampak dengan membandingkan dengan ukuran lain yang lebih standar.

Tes merupakan salah satu alat untuk melakukan pengukuran, yaitu alat untuk mengumpulkan informasi karakteristik suatu objek. Objek ini bisa berupa kemampuan peserta didik, sikap, minat, maupun motivasi.

Berdasarkan bentuknya, tes prestasi belajar dapat dikelompokkan menjadi tiga jenis, yaitu: 1) objektif, yang sederhana terdiri dari bentuk jawaban singkat, benar-salah atau dua pilihan, dan menjodohkan, serta objektif pilihan ganda dengan alternatif jawaban lebih dari dua; 2) uraian. Respons peserta tes terhadap sejumlah pertanyaan menggambarkan kemampuan dalam bidang tertentu. Tes merupakan bagian tersempit dari evaluasi.

Dengan demikian, esensi dari pengukuran adalah kuantifikasi atau penetapan angka tentang karakteristik atau keadaan individu menurut aturan-aturan tertentu. Keadaan individu ini bisa berupa kemampuan kognitif, afektif dan psikomotor atau keterampilan. Pengukuran memiliki konsep yang lebih luas dari pada tes. Kita dapat mengukur karakteristik suatu objek tanpa menggunakan tes, misalnya dengan pengamatan, skala *rating* atau cara lain untuk memperoleh informasi dalam bentuk kuantitatif.

Alat pengukuran lain adalah bentuk nontes. Bentuk ini dapat berupa angket dengan berbagai skala, daftar cek, dan bentuk lainnya. Jawaban responden terhadap instrumen ini selanjutnya di kuantifikasi atau disajikan dalam bentuk angka, yang selanjutnya disebut sebagai data.

D. PENILAIAN

Penilaian memiliki makna yang berbeda dengan evaluasi. Popham (1995) mendefinisikan *asesmen* dalam konteks pendidikan sebagai sebuah usaha secara formal untuk menentukan status siswa berkenaan dengan berbagai kepentingan pendidikan. Boyer & Ewel mendefinisikan *asesmen* sebagai proses yang menyediakan informasi tentang individu siswa, kurikulum atau program, institusi atau segala sesuatu yang berkaitan dengan sistem institusi, yaitu *processes that provide information about individual students, about curricula or programs, about institutions, or about entire systems of institutions* (Stark & Thomas: 1994).

Penilaian mencakup semua cara yang digunakan untuk menilai individu. Penilaian difokuskan pada individu dikaitkan dengan bukti-bukti yang telah dicapai oleh individu tersebut. Popham (1995) mendefinisikan *asesmen* dalam konteks pendidikan sebagai sebuah usaha secara formal untuk menentukan status siswa berkenaan dengan berbagai kepentingan pendidikan.

Boyer & Ewel mendefinisikan *asesmen* sebagai proses yang menyediakan informasi tentang individu siswa, kurikulum atau program, institusi atau segala sesuatu yang berkaitan dengan sistem institusi, yaitu *processes that provide information about individual students, about curricula or programs, about institutions, or about entire systems of institutions* (Stark & Thomas: 1994). Berdasarkan berbagai uraian di atas dapat disimpulkan bahwa *assessment* atau penilaian dapat diartikan sebagai kegiatan menafsirkan data hasil pengukuran melalui pemberian status seseorang.

E. EVALUASI

Beberapa ahli evaluasi mencoba mendefinisikan arti evaluasi. Ralph Tyler (1950) menyatakan bahwa *evaluation is the process of determining to what extent the educational objectives are actually being realized*. Definisi ini memiliki makna bahwa evaluasi merupakan suatu kegiatan untuk menentukan seberapa jauh suatu tujuan pendidikan tercapai. Evaluasi dilaksanakan dengan tujuan untuk mengetahui tingkat keberhasilan program. Tingkat keberhasilan program tersebut dapat diketahui melalui kegiatan penilaian.

1. Malcolm Provus (1951) mendefinisikan “*evaluation is the comparison of performance to some standard to determine whether discrepancies existed*”. Evaluasi merupakan cara untuk melihat kesesuaian program kegiatan yang telah dilakukan dengan penampilan standar yang telah ditentukan sebelumnya, atau membandingkan hasil yang telah dicapai dengan kriteria yang telah ditetapkan.
2. William Dunn menyatakan “*evaluation is the policy analytical methods used to produce information about the value or worth of past and/or future course of action*”. Evaluasi merupakan metode untuk menganalisis kebijakan, menghasilkan informasi tentang program yang telah dilaksanakan, menilai kegunaan, dan manfaat yang telah dicapai untuk mengungkapkan, menyelidiki, menginterpretasi sebagai dasar pengambilan kebijakan tentang program untuk masa mendatang.
3. Scriven (1967) menyatakan bahwa evaluasi sebagai metode atau cara untuk menilai kegunaan dan kemanfaatan suatu program yang telah dilaksanakan.
4. Stufflebeam (1971) memandang evaluasi sebagai suatu proses untuk mengidentifikasi dan mengungkap permasalahan yang terjadi terkait

- dengan program dalam suatu organisasi, mencari dan menganalisis data, dan menyajikan informasi untuk pembuatan keputusan.
5. Cronbach mengatakan bahwa evaluasi sebagai cara untuk meningkatkan kualitas pendidikan. Evaluasi akan menghasilkan informasi yang digunakan sebagai umpan balik dan pembenahan terhadap program pendidikan dan pembelajaran.
 6. Worthen & Sanders (1973). Evaluasi merupakan aktivitas untuk menentukan kelayakan dari sesuatu, misalnya program, produk, prosedur, dan tujuan atau kegunaan potensial dan alternatif pendekatan untuk mencapai tujuan khusus.

Pendapat lain dinyatakan oleh Caffarella (Douglas, 1998), yakni *“evaluation is process used to determine whether the design and delivery of program were effective and whether the proposed outcomes were met”*. Berdasarkan pendapat ini, diperoleh bahwa evaluasi merupakan proses yang digunakan untuk menentukan apakah rancangan dan pelaksanaan program sudah efektif, dampak peningkatan sudah tercapai. Dengan diketahuinya informasi ini, tingkat keberhasilan program dapat diketahui.

Evaluasi program merupakan suatu metode untuk mengetahui kinerja suatu program dengan membandingkan dengan kriteria yang telah ditentukan atau tujuan yang ingin dicapai dengan hasil yang dicapai. Hasil yang dicapai dalam bentuk informasi digunakan sebagai bahan pertimbangan untuk pembuatan keputusan dan penentuan kebijakan. Evaluasi ini dilakukan secara sistematis dengan melalui proses pengumpulan dan analisis data yang dapat dipertanggungjawabkan kebenarannya, untuk mengetahui tingkat keberhasilan suatu program.

F. RUANG LINGKUP EVALUASI PROGRAM PENDIDIKAN

Stufflebeam dan Shrinkfield (1985) menyatakan bahwa *“...the most important purpose of evaluation is not to prove, but to improve...”*. Kalimat ini mengandung pengertian bahwa tujuan evaluasi untuk meningkatkan, bukan membuktikan. Peningkatan mengandung makna bahwa penilaian dilakukan berkaitan dengan manfaat atau nilai, dengan kata lain kegiatan evaluasi berhubungan secara khusus dengan pernyataan ‘seberapa efektif atau seberapa tidak efektif’, ‘seberapa memadai atau seberapa tidak memadai’, ‘seberapa buruk atau seberapa tidak buruk’, ‘seberapa bernilai atau seberapa

tidak bernilai', seberapa cocok atau seberapa tidak cocok', dan seterusnya dari sebuah tindakan, proses, atau produk dari suatu program.

Pendapat lain mengenai tujuan evaluasi program dikemukakan oleh Kauffman dan Thomas (1980) yang menyebutkan bahwa tujuan evaluasi adalah "... *to collect data (result), convert the data into information, and use the information to make decisions. If decision are not made, the evaluation might just as well have been skipped*". Pendapat tersebut memuat makna bahwa evaluasi dimaksudkan untuk mengumpulkan data dan mengolah data tersebut menjadi informasi yang dapat membantu dalam pengambilan keputusan yang bermanfaat. Pengambilan keputusan dilakukan agar program yang dilakukan memperoleh masukan perubahan, perbaikan atau tindak lanjut berupa perpanjangan atau keterlanjutan dari suatu program. Hasil dari evaluasi berupa kebijakan atau keputusan.

Namara (1977) berpendapat bahwa "*program evaluating is carefully collecting information about a program or some aspect of program in order to make necessary decision about the program*". Terkait dengan hal ini, evaluasi program bertujuan untuk mengumpulkan informasi terkait dengan program atau aspek-aspek program untuk membuat keputusan yang terkait dengan program. Namara selanjutnya mengidentifikasi beberapa pertanyaan terkait dengan program sebagai berikut.

1. *How were the program goals (and objectives is applicable) established? Was the process effective?*
2. *What is the status of the program's progress toward achieving the goals?*
3. *Will the goals be achieved according to the timelines specified in the program implementation or operations plan? If not, then why?*
4. *Do personal have adequate resource (money, equipment, facilities, training, etc) to achieve the goals?*
5. *How should priorities be changed to put more focus on achieving the goals?*
6. *How should timelines be changed?*
7. *How should goals be changed?*
8. *How should goals be established in the future?*

Pertanyaan-pertanyaan berikut maknanya kurang lebih sebagai berikut.

1. Apakah tujuan yang ditetapkan sudah tepat dan dapat diwujudkan, seberapa efektif proses yang dilaksanakan.

2. Bagaimanakah kondisi kemajuan program dalam mencapai tujuan.
3. Akankah tujuan yang ditetapkan dapat tercapai sesuai dengan waktu yang ditentukan, jika tidak tercapai, apa yang menyebabkannya?
4. Apakah sumber daya yang ada mampu untuk mencapai tujuan?
5. Hal apa sajakah yang mendesak untuk dibenahi agar lebih fokus pada pencapaian tujuan?
6. Bagaimanakah waktu yang diperlukan untuk pelaksanaan program disesuaikan?
7. Bagaimanakah seharusnya tujuan ditetapkan di masa mendatang?

Terkait dengan hal-hal tersebut, informasi dari evaluasi digunakan untuk membuat keputusan dalam memperbaiki program terkait dengan tujuan, proses, waktu, sumber daya, prioritas perubahan, dan penetapan tujuan di masa mendatang.

Puet (2008) memiliki rumusan yang agak berbeda mengenai tujuan evaluasi program, yakni

1. *demonstrate program effectiveness to funders;*
2. *improve the implementation and effectiveness of program;*
3. *better manage limited resource;*
4. *document program accomplishment;*
5. *justify current program funding;*
6. *support the need for increased level of funding;*
7. *satisfy ethical responsibility to client to demonstrate positive and negative effects of program participations;*
8. *document program development and activities to help ensure successful replication.*

Menurut Puet tersebut, tujuan evaluasi program, yaitu

1. menunjukkan efektivitas program kepada penyumbang dana;
2. meningkatkan implementasi dan efektivitas program;
3. mengelola sumber daya yang terbatas dengan lebih baik;
4. membuktikan hasil yang diperoleh suatu program;
5. menunjukkan kondisi pendanaan program terkini;
6. memberikan dukungan yang diperlukan untuk menambah pembiayaan;
7. memberikan kepuasan yang layak dan bertanggung jawab kepada pelanggan untuk menunjukkan dampak positif dan negatif kepada peserta program;

8. menunjukkan kemajuan program dan kegiatan untuk membantu menjamin keberhasilan program yang sejenis.

Sudhardsono (1994) menyampaikan tujuan evaluasi yang disampaikan oleh Brooks. Menurut Brooks, evaluasi bertujuan untuk mengetahui:

1. ketercapaian suatu program dengan membandingkan tujuan program;
2. pengaruh relatif variabel-variabel kunci dari suatu program yang dievaluasi;
3. peran program bila dibandingkan dengan variabel eksternal.

Setiap program kegiatan yang direncanakan perlu diakhiri dengan evaluasi. Evaluasi ini dimaksudkan untuk melihat kembali apakah program tersebut dilaksanakan sesuai dengan perencanaan dan sesuai dengan tujuannya. Berdasarkan informasi hasil evaluasi, dapat dibandingkan apakah suatu program sesuai dengan kriteria yang telah ditetapkan sebelumnya. Berdasarkan hasil evaluasi, selanjutnya diambil keputusan apakah program akan diteruskan, direvisi, dihentikan karena menimbulkan banyak masalah, atau dirumuskan kembali disesuaikan dengan tujuan, sasaran, dan alternatif baru yang berbeda dengan sebelumnya.

Contoh kasus untuk membedakan pengukuran, penilaian, dan evaluasi sebagai berikut.

Guru ingin mengetahui kemampuan siswa. Untuk itu, guru menyelenggarakan ujian bagi siswanya. Mengetahui kemampuan siswa dengan menyelenggarakan ujian, kemudian hasil ujian diberi skor disebut dengan pengukuran. Hasil pengukuran kemudian diberi makna, misalnya kalau yang sama dengan atau di atas kriteria ketuntasan minimal (KKM) dikatakan lulus, yang kurang dari KKM dikatakan tidak lulus. Jika hampir semua siswa dalam suatu kelas tersebut lulus, program pembelajaran bisa dilanjutkan. Namun, jika 75% lulus dan sisanya tidak lulus, perlu dilakukan tindakan terkait dengan program pembelajaran di kelas tersebut, misalnya dengan memberikan pengayaan bagi yang lulus dan remedial bagi yang tidak lulus. Pada kasus ini, guru melakukan pengukuran dengan melakukan tes kepada siswa. Hasil tes selanjutnya diberi skor, yang merupakan proses kuantifikasi. Membandingkan hasil penskoran tersebut dengan KKM merupakan penilaian. Hasil penilaian dimanfaatkan untuk evaluasi, jika 65% siswa mencapai KKM maka pembelajaran dapat dilanjutkan. Jika belum,

perlu dilakukan perbaikan terlebih dahulu, atau metode pembelajarannya diganti sehingga hasil pembelajaran diharapkan menjadi lebih baik. Dalam hal ini hasil evaluasi akan memberikan saran apakah program akan berjalan terus atau dihentikan.

G. EVALUASI FORMAL DAN INFORMAL

Proses evaluasi melibatkan dua aktivitas utama yaitu mengumpulkan informasi yang mendukung untuk menyusun suatu keputusan atau kebijakan dan menerapkan suatu kriteria dari informasi yang telah tersedia untuk menentukan suatu kebijakan. Kedua langkah utama ini dilaksanakan secara sistematis dengan langkah-langkah yang jelas dan dapat dipertanggungjawabkan. Perolehan data disimpan dan dilaporkan, yang dapat ditelusuri ketika diperlukan, dan hasilnya dapat dikomunikasikan secara jelas dan tepat. Hal ini dimaksudkan agar temuan evaluasi dapat dimanfaatkan dan tersedia bagi masyarakat yang akan menggunakannya di masa mendatang.

Tanpa kita sadari, setiap hari kita melakukan evaluasi. Memilih baju yang akan kita kenakan sesuai dengan ketersediaan, suasana dan tuntutan lingkungan merupakan evaluasi. Bagi guru, menentukan apakah suatu metode pembelajaran efektif atau tidak, sehingga perlu mengganti dengan metode lain merupakan suatu evaluasi. Kepala sekolah yang akan memberikan sanksi kepada pelanggar kedisiplinan sekolah baik siswa ataupun guru merupakan evaluasi. Macam-macam kegiatan tersebut dilakukan setiap hari, baik di rumah, di tempat kerja, ataupun tempat yang lainnya. Namun, evaluasi tersebut dilakukan tidak menggunakan suatu prosedur yang sistematis dan tentunya tidak dapat dipertanggungjawabkan. Hal tersebut merupakan evaluasi informal.

Evaluasi informal bisa jadi menghasilkan suatu kebijakan yang keliru. Hal ini disebabkan karena kekurangcukupan prosedur yang sistematis dan pengumpulan datanya hanya bersifat ketidaksengajaan dan bukan sesuatu yang formal. Hasilnya juga hanya bersifat kebetulan, belum tentu bisa digunakan untuk semua kasus. Misalnya keputusan penggunaan metode pembelajaran tertentu. Metode pembelajaran tertentu dapat digunakan untuk suatu kasus baik misalnya untuk materi tertentu, namun metode tersebut belum tentu cocok untuk siswa dengan karakteristik yang berbeda. Dengan kata lain, hasil evaluasi informal tidak dapat digeneralisasikan atau diberlakukan secara umum.

Beberapa ahli menyampaikan pendapat lain mengenai tujuan evaluasi program. Weiss dan Henry dalam Fitzpatrick, Sanders, dan Worthen, 2004 menyatakan bahwa tujuan evaluasi adalah untuk membawa perbaikan sosial. Perbaikan ini terkait dengan pengurangan masalah sosial dan memenuhi kebutuhan manusia. Mark, Henry, dan Joullnes menekankan empat tujuan evaluasi, yakni menilai kepastasan dan kebernilaian, pemahaman dan pemenuhan, peningkatan organisasi dan program dan penambahan pengetahuan. Hal penting yang menjadi penekanan biasanya pada pemahaman dan pemenuhan yang ditunjukkan dengan pencapaian kepastasan dan kebernilaian tetapi kadang secara umum aktivitas terfokus pada layanan kepada pengguna yang sesuai.

H. PERAN DAN AKTIFITAS EVALUATOR PROFESSIONAL

Dalam aktivitas evaluasi orang yang melakukan evaluasi disebut dengan evaluator. Evaluator sebagai praktisi memerankan beberapa peranan penting dalam melaksanakan evaluasi. Peranan utama evaluator di lapangan adalah peningkatan organisasi dan program tidak hanya terfokus pada kepastasan dan kebernilaian. Dalam perannya, evaluator merupakan guru yang membantu praktisi mengembangkan penilaian yang kritis. Menurut Patton beberapa peran evaluasi yang berbeda meliputi fasilitator kolaborator, guru, konsultan manajemen, ahli pengembangan organisasi dan agen perubahan sosial. Lebih jauh lagi beberapa ahli menyatakan bahwa salah satu peran penting evaluator, yakni sebagai pengguna (*stake holder*). Evaluator tidak hanya melaporkan informasi atau menyediakannya untuk pengguna yang memerlukan informasi ini namun juga mendorong terbentuknya keputusan secara demokratis. Ahli lain, seperti Bickman dan Chen menekankan pentingnya evaluator dalam perencanaan program. Evaluator memberikan saran tentang teori-teori program atau model logis dari suatu program. Dalam peran penting tersebut evaluator merupakan teman yang mengkritisi desain evaluasi. Evaluator bisa berperan sebagai fasilitator dan kolaborator dalam suatu grup yang ditunjuk untuk mencari pemecahan suatu masalah.

Dalam memerankan berbagai peranya evaluator melaksanakan beberapa kegiatan. Kegiatan-kegiatan tersebut yakni melakukan negosiasi dengan grup pengguna untuk mendefinisikan tujuan evaluasi, mengembangkan kontrak, mengontrol staf, mengatur anggaran, mengumpulkan, menganalisis, menginterpretasikan data kuantitatif dan kualitatif, mengomunikasikan

dengan macam-macam *stakeholder* untuk mendapatkan masukan tentang evaluasi, menulis laporan, menyebarkan informasi, menyampaikan laporan dan perkembangan hasil dan mencari pihak lain untuk mengevaluasi hasil evaluasi. Saat ini, evaluator sudah merupakan pekerjaan yang menjanjikan atau profesional yang dilakukan oleh orang-orang yang telah terlatih dan terdidik sebagai evaluator dan telah ada organisasi profesinya yang menyelenggarakan pertemuan tahunan. Di Indonesia organisasi tersebut namanya HEPI (Himpunan Evaluasi Pendidikan Indonesia).



LATIHAN

Untuk memperdalam pemahaman Anda mengenai materi di atas, kerjakanlah latihan berikut!

- 1) Apakah yang dimaksud dengan program pendidikan?
- 2) Berikanlah contoh program pendidikan!
- 3) Apakah perbedaan antara pengukuran, penilaian, dan evaluasi? Jelaskan!
- 4) Apakah tujuan melaksanakan evaluasi?
- 5) Bagaimanakah pemanfaatan evaluasi?

Petunjuk Jawaban Latihan

- 1) Program pendidikan ini merupakan serangkaian kegiatan pendidikan yang mempunyai tujuan-tujuan tertentu.
- 2) Contoh program pendidikan yaitu Program Pengentasan Buta Aksara, Program pelatihan peningkatan kompetensi guru, program sertifikasi guru, dan lain-lain.
- 3) Evaluasi sangat terkait dengan kegiatan pengukuran dan penilaian, juga kebijakan. Pengukuran merupakan kegiatan menguantifikasi suatu fenomena, dan hasilnya disajikan berupa bilangan dengan cara yang sistematis. *Assessment* atau penilaian dapat diartikan sebagai kegiatan menafsirkan data hasil pengukuran melalui pemberian status seseorang. Evaluasi merupakan proses yang digunakan untuk menentukan apakah rancangan dan pelaksanaan program sudah efektif, dampak peningkatan sudah tercapai.

- 4) Evaluasi ini dimaksudkan untuk melihat kembali apakah program tersebut dilaksanakan sesuai dengan perencanaan dan sesuai dengan tujuannya.
- 5) Berdasarkan informasi hasil evaluasi, dapat dibandingkan apakah suatu program sesuai dengan kriteria yang telah ditetapkan sebelumnya. selanjutnya diambil keputusan apakah program akan diteruskan, direvisi, dihentikan.



RANGKUMAN

1. Program pendidikan ini merupakan serangkaian kegiatan pendidikan yang mempunyai tujuan-tujuan tertentu, contohnya program peningkatan kompetensi guru.
2. Evaluasi sangat terkait dengan kegiatan pengukuran dan penilaian, juga kebijakan.
3. Pengukuran merupakan kegiatan menguantifikasi suatu fenomena, dan hasilnya disajikan berupa bilangan dengan cara yang sistematis.
4. *Assessment* atau penilaian dapat diartikan sebagai kegiatan menafsirkan data hasil pengukuran melalui pemberian status seseorang.
5. Evaluasi merupakan proses yang digunakan untuk menentukan apakah rancangan dan pelaksanaan program sudah efektif, dampak peningkatan sudah tercapai.
6. Evaluasi ini dimaksudkan untuk melihat kembali apakah program tersebut dilaksanakan sesuai dengan perencanaan dan sesuai dengan tujuannya.
7. Berdasarkan informasi hasil evaluasi, dapat dibandingkan apakah suatu program sesuai dengan kriteria yang telah ditetapkan sebelumnya. Selanjutnya diambil keputusan apakah program akan diteruskan, direvisi, dihentikan.



TES FORMATIF 1

Jawablah pertanyaan berikut dengan tepat!

- 1) Program apa sajakah yang dicanangkan dan dilaksanakan oleh pemerintah atau oleh suatu lembaga? Deskripsikanlah program-program tersebut!

- 2) Apakah yang melatarbelakangi program tersebut? Bagaimana dan apa dasar hukumnya? Kegiatan-kegiatan apa saja yang terangkai dalam program-program tersebut?
- 3) Berikanlah contoh yang membedakan antara pengukuran, penilaian, dan evaluasi!
- 4) Deskripsikan kembali tentang sejarah perkembangan evaluasi program pendidikan.

Tindak Lanjut

Untuk mengetahui tingkat penguasaan materi kegiatan belajar ini, yang perlu Anda lakukan adalah:

1. Buatlah rangkuman materi dengan mengambil konsep-konsep utama dan kata-kata kunci pada petunjuk pengerjaan tes formatif.
2. Kerjakanlah latihan yang diberikan dengan mengikuti petunjuk jawaban latihan
3. Cocokkan hasil latihan dengan petunjuk jawaban, dan berilah skor jawaban Anda. Jika sesuai dengan petunjuk jawaban, berilah skor masing-masing butir maksimal 5. Skor dapat disesuaikan dengan tingkat kebenaran jawaban Anda.
4. Hitunglah tingkat penguasaan Anda. Tingkat penguasaan Anda dikatakan baik jika mencapai minimal 80.

$$\text{Tingkat Penguasaan} = \frac{\text{Jumlah Skor Semua Butir}}{5 \times \text{Jumlah Soal}} \times 100\%$$

5. Diskusikanlah hasil latihan yang Anda kerjakan bersama teman-teman serta penjelasan tutor.
6. Apabila hasil latihan yang baik (minimal 80) dan adanya penguatan baik dari tutor, silakan melanjutkan ke materi berikutnya.

KEGIATAN BELAJAR 2

Sejarah Evaluasi Program Pendidikan dan Perbedaan antara Evaluasi dengan Penelitian

A. SEJARAH EVALUASI PROGRAM PENDIDIKAN

Evaluasi program pendidikan telah berkembang sekitar 40 tahun terakhir. Pada tahun 1965, *Elementary and Secondary Education Act* (ESEA) memandatkan evaluasi, untuk membandingkan sekolah nasional dan sekolah lokal. Evaluasi diharapkan untuk (1) memastikan bahwa dana yang digunakan dapat dimanfaatkan oleh siswa-siswa yang kurang beruntung, (2) menyediakan informasi untuk memberdayakan orang tua dan masyarakat untuk melaksanakan pendidikan yang lebih baik. Pendapat lain menyatakan bahwa informasi dari program-program dan keefektifannya dapat digunakan untuk meningkatkan performa sekolah. Departemen kesehatan, pendidikan, dan kesejahteraan Amerika Serikat menyambut hangat adanya informasi ini, baik program-program yang telah diselenggarakan, pelayanan masyarakat, dan strategi pembelajaran yang digunakan. Setelah itu, sekretariat departemen tersebut mempromosikan pentingnya evaluasi dalam arti menemukan pekerjaan apa yang merupakan langkah awal untuk menyebarluaskan pengalaman-pengalaman yang efektif.

Evaluasi pendidikan bukan merupakan kegiatan yang baru di dunia pendidikan. Pada awal 2000 SM, pejabat China telah melakukan ujian untuk mengukur kemampuan dari pelamar yang akan menempati posisi pemerintah. Dalam pendidikan, Socrates menggunakan evaluasi verbal sebagai bagian dari proses pembelajaran. Namun, setelah itu evaluasi secara formal dilengkapi dengan keyakinan agama dan politik terkait dengan keluasan dalam rangka membuat keputusan di bidang sosial dan pendidikan.

Kegiatan lain terkait evaluasi dilakukan oleh kapten kapal Inggris di sekitar tahun 1700. Kapten melakukan perlakuan (*treatment*) atau semacam perlakuan kepada awak kapal untuk mengonsumsi sejenis sari buah yang disebut dengan limes dengan kelompok kontrol pelaut yang mengonsumsi makanan biasa. Walaupun hasilnya tidak dijelaskan secara pasti, namun setelah masa tersebut, awak kapal dipaksa selalu mengonsumsi buah-buahan yang sampai saat ini tetap diberlakukan di Inggris.

1. Tahun 1800-1940

Selama abad 19, ketidakpuasan dengan berbagai program pendidikan dan sosial di Inggris mendorong dilaksanakannya evaluasi institusi terkait. Demikian pula di Amerika Serikat, beberapa sekolah telah melaporkan hasil *asesmen* untuk mengetahui kualitas sistem sekolah dan menilai efektivitas sekolah. Di akhir 1800-an, pembaharu Joseph Rice melaksanakan studi komparatif terhadap berbagai metode pembelajaran. Hasil dokumentasi kegiatan ini menunjukkan kurang efisiennya waktu di sekolah. Akreditasi sekolah juga sudah mulai dilaksanakan di akhir 1500an khususnya di sekolah menengah dan di universitas-universitas di Amerika.

Evaluasi juga sudah dilakukan di awal 1900-an, seiring dengan perkembangan teori dan teknologi pengukuran. Tokoh yang terkenal di antaranya Thorndike, yang mencoba mengembangkan berbagai level kemampuan. Pada pertengahan 1930an, telah dikembangkan tes terstandar dan pelaksanaan tes normatif termasuk tes kemampuan dan tes personal. Selanjutnya 1950an, Tyler mengembangkan pelaksanaan tes berdasarkan kriteria sebagai alternatif pelaksanaan tes berdasarkan norma.

2. Tahun 1940-1964

Selama perang dunia II, peneliti-peneliti di bidang sosial terapan membantu tentara dan veteran perang untuk meningkatkan kualitas moralnya, meningkatkan kemampuannya, dan menempatkan tentara dan veteran tersebut dalam suatu pekerjaan. Program-program tersebut meliputi *training* pekerjaan, perencanaan keluarga dan perumahan, dan perkembangan masyarakatnya. Bidang sosial terapan ini berkembang dalam ruang lingkup dan skala yang semakin luas.

Penelitian evaluasi merupakan istilah untuk penelitian sosial yang diterapkan untuk meningkatkan kualitas program. Pada akhir 1950-an, evaluasi dilaksanakan untuk mengetahui keberhasilan beberapa program, proyek rehabilitasi, terapi-terapi, program perumahan publik, dan aktivitas organisasi. Model ini menyebar di berbagai Negara dan benua. Amerika dan Afrika misalnya, yang mengevaluasi kesehatan dan pemberian makanan yang bergizi. Kedua program ini diperuntukkan untuk pengembangan masyarakat yang padat penduduknya.

Pada masa antara 1940-1965, pola perkembangan yang agak berbeda terjadi pada evaluasi program pendidikan. Masa ini merupakan masa konsolidasi perkembangan evaluasi awal. Beberapa hal yang dikembangkan

pada masa ini yakni mengembangkan tes yang terstandar, mendesain *quasi* eksperimen, akreditasi, dan survei sekolah. Termasuk pula pada awal 1960-an, telah ada upaya mengevaluasi pembelajaran untuk pendidikan sesuai dengan domain kognitif dari Bloom, bahwa kedalaman kognitif peserta didik dapat diketahui pada level pengetahuan, pemahaman, penerapan, analisis, sintesis, dan evaluasi.

Setelah Sputnik berhasil diluncurkan oleh Rusia 1957, Amerika Serikat menerapkan pembelajaran yang efektif untuk siswa Amerika, khususnya mata pelajaran matematika dan sains. Berbagai program pengembangan kurikulum baru kemudian dirancang. Meskipun tidak semua proyek didanai, namun hal-hal tersebut mengarahkan pembuat kebijakan mengembangkan dan mendanai evaluasi.

Pada masa ini, beberapa penelitian telah dihasilkan, namun teori evaluasi belum ada. Metode yang di adaptasi dari penelitian sosial dan behavioral tak mencukupi bahkan gagal untuk diterapkan, yang mendorong perkembangan evaluasi pada tahun-tahun selanjutnya.

3. Tahun 1964-1972

Di era ini, pemerintah Amerika Serikat melakukan berbagai program di semua sektor baik pendidikan, kesehatan, perumahan, pengangguran, dan lain-lainnya. Program-program ini kemudian dievaluasi efektivitasnya, yang memberikan sumbangan pada evaluasi program selanjutnya dan perkembangan teori evaluasi pada umumnya.

Konsep evaluasi lain dikembangkan oleh perusahaan motor dan mobil FORD, yang menerapkan sistem *Planning, Programming, dan Budgeting System* (PPBS). Sistem ini diterapkan di bidang lain, untuk meningkatkan sistem efisiensi, efektivitas dan keputusan tentang alokasi *budget*. Selanjutnya, program *the Elementary and Secondary Education Act* (ESEA) 1965 yang merupakan program pemberian bantuan dana di sekolah lokal, lembaga-lembaga pendidikan, dan universitas. Presiden Robert F. Kennedy menekankan setiap penerima bantuan melaporkan evaluasi tentang hasil yang diperoleh dari pemanfaatan dana tersebut. Berdasarkan penekanan ini, selain tuntutan akuntabilitas, digunakan pula tes yang terstandar untuk menggambarkan hasil belajar siswa dihubungkan dengan tujuan pembelajaran. Keberadaan ESEA ini merupakan tonggak awal kelahiran evaluasi program modern. Selanjutnya evaluasi ini berkembang di berbagai wilayah, mulai dari pendidikan, kesehatan, sosial, dan lain-lain yang mendorong berbagai pakar, baik politik, ekonomi, dan sosiologi untuk

bersama mengatur dan mengevaluasi berbagai program di instansi-instansi yang terkait dengan bidang kepakarannya.

Pada saat itu, teori dan metodologi yang mendukung belum ada. Kondisi ini mendorong berbagai pakar untuk memberikan sumbangan teori pada evaluasi program berdasarkan metodologi yang telah berkembang sebelumnya misalnya menggunakan desain eksperimen, ilmu pengukuran (psikometri), penelitian sosial, survei dan etnografi. Beberapa buku dan artikel dipublikasikan. Suchman (1967) memublikasikan teks tentang *review* berbagai metode evaluasi dan Chambell (1969) menyampaikan perlunya menggunakan eksperimen sosial untuk menguji efektivitas program. Selanjutnya, Stake (1967), Scriven (1967), Stufflebeam (1968) mulai menulis evaluasi, baik praktik maupun teorinya. Selanjutnya, Wholey (1970) mengenalkan aspek politis evaluasi yang dilaksanakan. Carol Weiss (1972) memublikasikan buku tentang evaluasi. Selanjutnya, artikel evaluasi sering dipublikasikan di berbagai jurnal profesional. Selanjutnya, di bawah payung penelitian, pendidikan, dan evaluasi, menjadi program spesialisasi di pascasarjana di berbagai universitas luar negeri dan dalam negeri.

4. Tahun 1973-1984

Perkembangan selanjutnya yakni perkembangan evaluasi yang meluas di berbagai sektor dan memerankan peranan yang penting dalam pemerintahan. Pada saat ini, beberapa tulisan tentang evaluasi program yang baru diajukan. Nilai-nilai dan standar-standar juga ditekankan, termasuk pentingnya kepastan dan kebernilaian program yang dilaksanakan oleh suatu institusi atau suatu lembaga. Scriven (1972) mencoba menggeser pemahaman evaluasi, tidak hanya berbasis tujuan namun juga *goal-free* (tanpa tujuan), untuk menguji proses dan konteks program dalam menemukan hasil yang tidak teridentifikasi. Stufflebeam (1971) menekankan perlunya evaluasi untuk membuat keputusan *informative* dengan mengembangkan model *Context, Input, Process, and Product* atau dikenal dengan CIPP. Stake (1975) mengajukan evaluasi responsif agar *evaluator* tidak menggunakan eksperimen saja. Guba & Lincoln (1981) berdasarkan hasil kualitatif dari Stake, mengajukan evaluasi naturalistik, yang menggabungkan metode kuantitatif dan kualitatif. Dengan berbagai pandangan penulis yang berbeda-beda tentang penelitian, evaluasi pada dekade ini merupakan alat teknis dan politis yang menggabungkan konsep baru dan pemahaman baru dari metodologi penelitian yang telah ada.

5. Tahun 1985-sekarang

Dibentuk oleh Asosiasi Evaluasi Amerika (AEA) (1985) yang merupakan gabungan organisasi profesi evaluator, yakni Evaluation Research Society dan Evaluation Network mengadakan pertemuan tahunan. Dampak ini meluas sampai ke Kanada dan kemudian merambah ke seluruh dunia.

B. DASAR-DASAR EVALUASI

Dalam suatu program, ada banyak rangkaian kegiatan yang dilaksanakan, untuk mencapai tujuan dan hasil yang diharapkan. (Tujuan merupakan sesuatu yang dicapai, misalnya kompetensi tertentu yang dicapai oleh peserta, sedangkan hasil berwujud suatu produk atau barang). Dalam melaksanakan evaluasi program, ada beberapa pertanyaan yang terkait dengan evaluasi program.

1. Untuk apa, sejauh mana, jika ada, apakah program mencapai tujuan yang dimaksudkan?
Tujuan program merupakan hal yang penting untuk diketahui dalam evaluasi program. Dengan melihat tujuan program, dapat diketahui dan dirumuskan kriteria keberhasilan program. Kriteria keberhasilan program digunakan untuk membandingkan, apakah tujuan program sudah tercapai.
2. Apakah program efektif (dalam mencapai hasil yang diharapkan)?
Salah satu indikator keberhasilan program adalah hasil. Suatu program dikatakan efektif, jika hasil yang telah dicapai sesuai dengan yang diharapkan. Mengakses efektivitas program merupakan alasan utama melaksanakan evaluasi program.
3. Sampai sejauh mana, hasil yang diamati konsisten dengan hasil yang diharapkan?
Konsistensi juga menjadi tolok ukur keberhasilan program. Suatu program konsisten jika hasil yang teramati konsisten dengan hasil yang diharapkan.
4. Apakah program efisien?
Melihat efisiensi program dapat diketahui dengan 2 cara, yakni efisiensi secara teknis atau efisiensi secara ekonomis. Efisiensi secara teknis terkait dengan rasio *input* khususnya uang dengan program yang telah dilaksanakan dikaitkan dengan kriteria tertentu. Efisiensi secara teknis ini terkait pula dengan produktivitas, unit biaya per satuan yang telah

dikeluarkan. Efisiensi ekonomi difokuskan pada keuntungan suatu program dibandingkan dengan biaya yang telah dikeluarkan.

5. Apakah program efektif dari sisi pembiayaan? Efektivitas pembiayaan diketahui dengan membandingkan biaya program dengan hasilnya. Rasio biaya per unit hasil menjadi tolok ukur evaluasi program. Bisa saja suatu kegiatan memerlukan dana yang banyak, namun tidak efektif karena hasil yang diharapkan tidak sebanding dengan biaya yang dikeluarkan.
6. Apakah program sesuai? Pertanyaan ini terkait dengan struktur program, apakah logis untuk dapat memperoleh hasil seperti yang diharapkan, termasuk pula latar belakang ilmu/pengetahuan, mendesai programnya, dan upaya untuk mencapai hasil (*outcome*) dari program.
7. Apakah rasional atau latar belakang dari program rasional? Jika telah disajikan tujuan program, perlu diketahui apakah program sesuai dengan misi dan tujuan dari suatu instansi? Apakah program sesuai dengan prioritas kebutuhan dan kebijakan? Kesesuaian program ini diketahui berdasarkan suatu *asesmen* kebutuhan (*need assesment*).
8. Apakah program tersebut memadai? Memadai tidaknya suatu program dapat diketahui dengan jangkauan kegiatan, sumber daya program, dan skala pelaksanaan program, apakah program terlalu luas atau terlalu sempit.
9. Seberapa baik program diimplementasikan? Pelaksanaan program juga menjadi indikator keberhasilan. Pelaksanaan program yang baik sesuai dengan yang telah direncanakan.

Evaluasi dapat dilakukan secara sumatif dan formatif. Evaluasi yang dilaksanakan di akhir suatu program dinamai dengan evaluasi sumatif. Namun, jika evaluasi dilakukan ketika program sedang dilaksanakan dinamai dengan evaluasi formatif. Beberapa perbedaan evaluasi sumatif dan formatif disajikan sebagai berikut.

Subjek yang diamati (serig disebut dengan audiens) pada evaluasi sumatif dan formatif berbeda. Pada evaluasi formatif, audiens adalah orang yang melaksanakan program, misalnya orang yang mengembangkan/membuat jadwal baru, melaksanakan training, dll. Pada dasarnya, evaluasi formatif digunakan untuk memperbaiki program sehingga audiens yang paling utama diamati adalah orang yang melaksanakan program kesehariannya. Pada evaluasi sumatif, audiens adalah konsumen yang

potensial, misalnya guru, siswa, manajer, yang menggunakan program. Selain itu, audiens untuk evaluasi sumatif misalnya pembuat kebijakan atau administrator, yang nantinya menentukan apakah program tersebut dilanjutkan atau tidak. Perbandingan antara evaluasi formatif dan sumatif disajikan sebagai berikut.

Aspek	Evaluasi Formatif	Evaluasi Sumatif
Tujuan	menentukan nilai atau kualitas	menentukan nilai atau kualitas
Pemanfaatan	meningkatkan program	membuat keputusan tentang masa depan program
Audiens (subjek yang diamati)	manajer program dan staf	administrator, pembuat keputusan, pengguna potensial, penyandang dana
Dilaksanakan oleh	lebih utama dilakukan oleh evaluator internal didukung oleh evaluator eksternal	pada umumnya evaluator eksternal, didukung oleh evaluator internal
Ciri utama	menyediakan masukan terhadap program agar dapat ditingkatkan	menyediakan informasi yang memungkinkan pembuat keputusan memutuskan keberlanjutan program, atau pihak lain mengadopsinya
Batasan desain	informasi apa yang diperlukan? Kapan?	Kejadian apa yang diperlukan untuk membuat keputusan utama?
Tujuan Pengumpulan Data	diagnostik	Keputusan
Frekuensi Pengumpulan Data	sering	agak sering
Ukuran sampel	kecil	biasanya besar
Pertanyaan yang ditanyakan	Apa saja yang dilaksanakan? Apakah memerlukan peningkatan? Bagaimana meningkatkannya?	bagaimana hasilnya? Oleh siapa? Bagaimana kondisi yang mendukung? Dengan training apa? Bagaimana biayanya?

Evaluasi baik sumatif maupun formatif merupakan hal yang penting. Dengan evaluasi formatif, tahapan-tahapan perkembangan program dapat ditingkatkan, Peningkatan program berujung pada kelayakan program di masa mendatang, Pada evaluasi sumatif, kegiatan difokuskan pada standar pencapaian, yang hanya menekankan hasil. Sebagai contoh misalnya pada program penggunaan kurikulum yang baru. Selama pelaksanaan program,

dengan evaluasi formatif diperoleh masukan terkait dengan perbaikan bertahap dari kurikulum misalnya pelaksanaan proses oleh guru, maupun pembelajaran yang dilaksanakan. Namun, pada evaluasi sumatif, keberhasilan kurikulum lebih ditekankan pada hasil belajar.

Perbedaan antara evaluasi sumatif dan formatif terletak pada jenis keputusan hasil evaluasi. Perbedaan ini menjadi dasar pijakan awal kegiatan evaluasi. Beberapa ahli menggunakan istilah *needs assessment*, *process*, dan *outcome* untuk menjawab beberapa permasalahan terkait dengan apakah kebutuhan atau suatu masalah ada dan kemudian menyusun rekomendasi untuk mengurangi permasalahan. Pada tahapan proses atau monitoring, dideskripsikan bagaimana program dilaksanakan. *Outcome* terfokus pada menggambarkan, mengeksplorasi, menentukan perubahan yang terjadi dalam penerima program, pengguna yang lain atau masyarakat umum terhadap tujuan final dan beberapa hasil sampingan atau dampaknya. Istilah ini dapat dimanfaatkan untuk membedakan antara evaluasi sumatif dan evaluasi formatif. Pada evaluasi formatif, terkait dengan evaluasi proses sedangkan pada evaluasi sumatif, lebih pada hasil.

C. EVALUASI INTERNAL DAN EVALUASI EKSTERNAL

Terkait dengan pelaksanaannya, evaluasi dapat dilakukan eksternal maupun internal. Evaluasi internal menggunakan evaluator dari dalam instansi yang melakukan evaluasi (evaluator internal). Ada beberapa keuntungan menggunakan evaluator internal. Evaluator internal dapat mengamati dan mendeskripsikan berbagai hal yang tidak diketahui oleh pihak eksternal. Namun, kedekatan evaluator dengan organisasi/program menutup/mengurangi munculnya inovasi baru untuk perbaikan program yang sedang atau telah dilaksanakan.

Alternatif lain yang dapat digunakan, yakni menggunakan evaluator eksternal. Evaluator eksternal merupakan orang yang berada di luar sistem instansi yang dievaluasi, sehingga lebih bisa dipercaya. Evaluator eksternal juga lebih dapat melihat secara objektif permasalahan dan perkembangan program yang dievaluasi.

Evaluasi formatif dan sumatif dapat dikombinasikan dengan dimensi internal dan eksternal. Sebagai contoh, adalah evaluasi program pemanfaatan dana BOS dapat dilakukan secara internal dan eksternal. Biasanya, selama masih proses pelaksanaan, dilakukan evaluasi oleh orang “dalam” atau

personil yang juga terlibat dalam pemanfaatan dana BOS, baik dilihat dari ketercapaian kegiatan maupun kesesuaian dengan rencana. Evaluasi ini dikenal dengan evaluasi formatif internal. Pada akhir program atau menjelang akhir program, ada evaluator dari luar (biasanya dari dinas pendidikan atau lembaga pihak ketiga) melakukan evaluasi pemanfaatan dana BOS tersebut, baik kegiatan, kesesuaian kegiatan dengan anggaran, maupun ketercapaiannya. Hal tersebut disebut dengan sumatif eksternal.

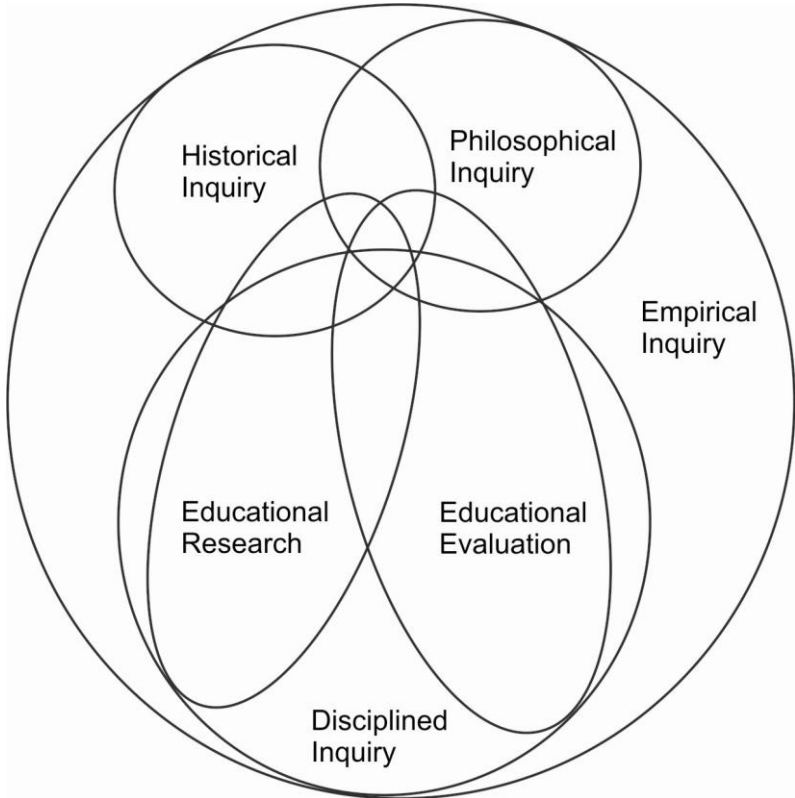
Evaluasi merupakan hal yang penting untuk dilakukan terkait dengan suatu program yang dilaksanakan oleh suatu instansi, baik pemerintah maupun nonpemerintah. Evaluasi mengarahkan suatu tujuan dengan cara yang lebih baik dengan melihat kelayakan dan berharganya suatu program. Adapun keterbatasan suatu evaluasi misalnya seberapa baik evaluator dapat melihat kelemahan dan kekuranglayakan program untuk merumuskan perbaikan pada evaluasi formatif. Evaluasi menghasilkan informasi tentang kekuatan dan kelemahan program, namun perbaikan program tergantung pada pelaksanaan program.

D. PERBEDAAN ANTARA EVALUASI DAN PENELITIAN

Manusia memiliki rasa ingin tahu (*curiosity*). Keingintahuan ini diselesaikan dengan melalui penerapan metode ilmiah, untuk memperoleh pencerahan. Kegiatan ini sering disebut dengan penyelidikan yang disiplin (*disciplined inquiry*), agar dapat dibedakan antara pendapat dan keyakinan. Penyelidikan dapat dikategorikan menjadi 3 keluarga besar, yakni inkuiri empiris (*empirical inquiry*), inkuiri historis (*historical inquiry*) dan inkuiri filosofis (*philosophical inquiry*). Masing-masing keluarga saling berkaitan atau ada irisan masing-masing. Inkuiri empiris merupakan upaya mendeskripsikan kondisi atau membuktikan kebenaran pernyataan yang ada kaitannya dengan suatu variabel melalui suatu observasi atau eksperimen, dan melalui suatu pengukuran baik terkait dengan kemampuan, perilaku, pendapat, dan lain-lain. Inkuiri historis merupakan studi pengembanan hal-hal yang terkait dengan manusia, kehidupan, perpindahan, dan kejadian-kejadian secara naratif. Inkuiri filosofis terkait dengan analisis secara rasional berdasarkan logika formal dan semantik.

Evaluasi pendidikan dan penelitian pendidikan merupakan dua kegiatan yang sama-sama menggunakan teknik inkuiri yang sistematis. Dua-duanya

juga beririsan dengan inkuiri historis dan inkuiri filosofis. Hasilnya disajikan pada Gambar 1.1.



Sumber: Worthen & Sanders, hal. 15

Gambar 1.1
Hubungan antara Penelitian Pendidikan dan Evaluasi Pendidikan
dalam Inkuiri

Ada 12 karakteristik dari inkuiri yang menjadi dasar untuk membedakan penelitian dengan evaluasi.

1. Motivasi dari *inquirer*. Penelitian dan evaluasi dilaksanakan untuk beberapa alasan yang berbeda. Penelitian dilakukan untuk memenuhi keingintahuan (*curiosity*), sedangkan evaluasi dilakukan untuk

memberikan sumbangan penyelesaian atau solusi dari permasalahan praktis.

Contoh 3:

Penelitian

Seorang peneliti ingin mengetahui dampak dari program RSBI di masyarakat. Penelitian ini dilakukan, semata hanya untuk menjawab pertanyaan peneliti tentang dampak apakah yang terjadi di masyarakat dengan adanya RSBI, misalnya terkait dengan pertanyaan berikut. Bagaimana tanggapan masyarakat? Bagaimana proses pembelajarannya? Bagaimana biayanya? Bagaimana *output* siswanya? dan seterusnya. Informasi yang diperoleh digunakan hanya untuk memperoleh pencerahan tentang dampak adanya RSBI yang terjadi di masyarakat.

Evaluasi

Di masyarakat terjadi keresahan terkait dengan adanya RSBI. Selain pemberian dana dari pemerintah yang cukup besar untuk menyelenggarakan RSBI ini, orang tua yang anaknya diterima di RSBI harus membayar sumbangan pendidikan yang cukup besar. Dengan adanya berbagai fasilitas yang diperoleh di sekolah dengan kategori RSBI, terjadi kecemburuan siswa-siswa di sekolah non RSBI, yang berdampak kurang baik di masyarakat. Terkait dengan hal ini, perlu dievaluasi kembali pelaksanaan RSBI. Dari evaluasi ini diharapkan diperolehnya informasi, apakah RSBI yang ada telah berjalan seperti tujuannya, dilihat dari efisiensi berdampak baik bagi masyarakat, menghasilkan *output* yang cukup setara dengan energi yang dikeluarkan, dan lain-lain. Berdasarkan informasi ini, dapat diperoleh masukan, apakah program tersebut dapat dipertahankan, diperbaiki ataukah dihentikan.

2. Tujuan penelitian dan evaluasi. Penelitian dan evaluasi mempunyai tujuan akhir yang berbeda. Menurut Cronbach dan Suppes (dalam Worthen dan Sanders, perbedaan tersebut dikatakan dengan berorientasi pada keputusan (*decision-oriented*) dan berorientasi pada kesimpulan (*conclusion-oriented*). Ketika tujuan berorientasi pada keputusan, invertigator menyediakan informasi yang diinginkan oleh pembuat keputusan, seperti halnya administrator sekolah, pembuat kebijakan

pendidikan, manajer proyek pengadaan buku teks, dan sejenisnya. Pembuat keputusan menginginkan informasi untuk memandu langkah-langkah yang akan dilakukan berikutnya. Pada kegiatan yang berorientasi pada kesimpulan, peneliti merumuskan sendiri pertanyaan penelitiannya. Kegiatan dilakukan untuk mengonsep dan memahami gejala. Kegiatan difokuskan pada orang-orang atau keadaan yang diharapkan dapat memberikan pencerahan pada masalah yang diteliti. Penelitian berorientasi pada kesimpulan, sedangkan evaluasi berorientasi pada keputusan.

Penelitian

Pada penelitian tentang dampak yang timbul di masyarakat dengan adanya program RSBI, fokus utamanya adalah menjawab pertanyaan penelitian. Informasi yang diperoleh dengan adanya dampak penelitian RSBI, digunakan untuk memperoleh pemahaman. Misalnya, dengan adanya RSBI, siswa lebih berani berkomunikasi dalam bahasa Inggris. Setelah memperoleh pemahaman, tidak ada yang dilakukan lagi, selain menambah khasanah pengetahuan.

Evaluasi

Evaluasi program RSBI berdampak pada masyarakat, dampaknya kemudian diidentifikasi. Berdasarkan identifikasi ini, dirumuskan keputusan. Jika dampaknya lebih banyak, program RSBI perlu dilanjutkan tentunya dengan berbagai revisi dan modifikasi agar lebih baik. Namun, jika dampak buruknya lebih banyak, perlunya keputusan untuk diberhentikan dan perlu diganti dengan program lain untuk meningkatkan daya saing bangsa, sebagaimana tujuan awal RSBI.

3. Aturan-aturan dan deskripsi-deskripsi. Pada penelitian, *nomothetic* (penjelasan tentang sesuatu yang umum) sedangkan evaluasi merupakan kegiatan *idiographic* (penjelasan hal-hal yang khusus). Pada penelitian, dicari deskripsi hal-hal yang berlaku secara umum sebagai khasanah pengetahuan, sedangkan pada evaluasi dicari penjelasan yang khusus untuk memperoleh informasi dan membuat keputusan apakah program dapat diteruskan, direvisi, atau dihentikan.

4. Peran dari *explanator* memberikan penjelasan secara ilmiah, sedangkan kegiatan evaluasi mencari pemahaman fenomena yang dievaluasi. Suatu produk atau program bagus atau tidak, namun perlu penjelasan yang bisa dipercaya.
5. Otonomi dari *inquiry*. Pada penelitian, peneliti sendiri yang mempunyai permasalahan, namun pada evaluasi klienlah yang merasakan permasalahan.
Sebagai contoh, pada penelitian terkait dengan dampak RSBI, peneliti ingin mengetahui dampak penelitian secara pribadi, namun pada evaluasi, lembaga pemberi dana/pemerintah yang merasakan masalah untuk melakukan revisi program, dalam rangka meningkatkan ketercapaian program.
6. Sifat dari fenomena yang diakses. Evaluasi pendidikan berusaha mengakses kebermaknaan (*worth*) dari suatu program, sedangkan penelitian berusaha mengakses kebenaran ilmiah. Kebermaknaan terkait dengan kegunaan sosial, seperti halnya meningkatkan kebahagiaan, kesehatan harapan hidup, dll., dan mengurangi rasa sakit, kegagalan, dll. Kebenaran ilmiah difokuskan pada 2 hal, mengukuhkan bukti empiris dan konsistensi logis. Peneliti tidak perlu mempermasalahkan suatu program berguna atau tidak.
7. *Generalisabilitas* kejadian yang menjadi fokus studi. Evaluasi program terfokus pada fenomena lintas waktu dan lintas wilayah, sedangkan pada penelitian relatif tetap.
8. Teknik investigasi. Pada dasarnya beberapa ahli mengatakan bahwa peneliti dan evaluator bekerja pada paradigma inkuiri yang sama dan termasuk dalam metodologi penelitian pada umumnya. Namun bedanya, untuk mencari hubungan antar variabel, peneliti melakukan berbagai hal, termasuk mengontrol variabel-variabel yang tidak diperlukan. Hal ini tidak perlu dilakukan oleh evaluator, yang penting bagi evaluator adalah kebermaknaan.
9. Keutamaan pernyataan mengenai nilai. Pertanyaan mengenai nilai digunakan untuk memilahkan informasi yang dicari. Pada penelitian, pertanyaan tentang nilai tidak terlalu menjadi perhatian.
10. Kriteria untuk menilai aktivitas. Pada penelitian, juga terkait dengan validitas internal (untuk perluasan hasil penelitian dan tidak ada kaitannya dengan pengaruh yang tidak perlu, atau adanya keberagaman yang disebabkan oleh kesalahan sistematik) dan eksternal *validity*

(generalisasi hasil pada populasi lainnya). Sedangkan untuk nilai kecukupan evaluasi, yakni *isomorphism* (perluasan informasi pada realitas dengan bentuk yang hampir sama) dan kreativitas (perluasan informasi dipandang sebagai dapat dipercayanya oleh klien yang menggunakan informasi tersebut).

11. Dasar disiplin. Agar penelitian terfokus, peneliti hanya menggunakan satu paradigma saja. Namun, evaluator tidak hanya dengan satu paradigma saja.
12. Training. Training bagi evaluator menggunakan pendekatan multidisiplin. Jika perlu dikaitkan dengan bidang lain yang relevan agar evaluator menjadi lebih sensitif. Pada peneliti, training yang dilakukan semakin mengerucut ke bidang keahlian. Semakin pakar, akan semakin menyempit sesuai bidang keahlian.



LATIHAN

Untuk memperdalam pemahaman Anda mengenai materi di atas, kerjakanlah latihan berikut!

- 1) Jelaskan perkembangan evaluasi mulai tahun 1973 sampai sekarang!
- 2) Jelaskan perbedaan evaluasi formatif dan evaluasi sumatif!
- 3) Jelaskan perbedaan antara evaluasi dan penelitian dilihat dari tujuannya!
- 4) Pada suatu studi tentang pelaksanaan program paket C, disajikan berbagai hasil sebagai berikut.
 - a. Warga belajar adalah lulusan SMP atau program paket B yang tidak mampu melanjutkan ke sekolah umum, sebanyak 24 orang.
 - b. Pendidik adalah warga setempat, 2 guru SMP/SMA yang telah menyelesaikan S1 dan 2 tokoh masyarakat, namun lulusan SMA.
 - c. Pembelajaran dilaksanakan secara konvensional, satu pekan duakali dan per tahunnya efektif 2 semester masing-masing 4 bulan.
 - d. Pada tahun 2012/2013, warga belajar lulus 100%.
- 5) Apakah hasil studi ini merupakan hasil penelitian atau hasil evaluasi? Jelaskan!
 - a. Pada suatu studi tentang pelaksanaan program paket C, disajikan berbagai hasil sebagai berikut.
 - b. Warga belajar adalah lulusan SMP atau program paket B yang tidak mampu melanjutkan ke sekolah umum, sebanyak 24 orang, dengan angka partisipasi 100% (kategori sangat baik).

- c. Pendidik (tutor) adalah warga setempat, 2 guru SMP/SMA yang telah menyelesaikan S1 dan 2 tokoh masyarakat, namun lulusan SMA, atau baru 50% pendidik (tutor) pada kategori sesuai.
- d. Pembelajaran dilaksanakan secara konvensional, satu pekan dua kali dan per tahunnya efektif 2 semester masing-masing 4 bulan, atau pada kategori baik.
- e. Pada tahun 2012/2013, warga belajar lulus 100%, atau pada kategori sangat baik.
- f. Berdasarkan keempat hasil tersebut di atas, pelaksanaan program paket C dapat dikategorikan berhasil, namun ada yang perlu diperbaiki yaitu pendidik/tutornya.

Petunjuk Jawaban Latihan

- 1) Pada masa ini berkembang model evaluasi tanpa tujuan (*goal free*) oleh Scriven untuk menguji proses dan konteks program dalam menemukan hasil yang tidak teridentifikasi, kemudian Stufflebeam dengan CIPP, Stake dengan evaluasi model responsif, dan Guba dan Lincoln dengan metode naturalistik. Selanjutnya setelah tahun 1980 didirikan asosiasi evaluasi di berbagai belahan dunia.
- 2) Perbedaan dapat dilihat dari pemanfaatan, subjek yang diamati, pelaksana, ciri utama, batasan desain, tujuan pengumpulan data, frekuensi pengumpulan data, ukuran sampel, dan pertanyaan yang ditanyakan pada evaluasi, dengan kata kunci meningkatkan program untuk formatif, dan menentukan masa depan program untuk evaluasi sumatif.
- 3) Perbedaan penelitian dan evaluasi dilihat dari tujuannya yaitu penelitian lebih berorientasi atas jawaban pertanyaan penelitian yang berujung pada kesimpulan, namun pada evaluasi lebih kepada keputusan masa depan program, apakah program dihentikan, direvisi, atau dilanjutkan.
- 4) Hasil studi tersebut merupakan penelitian, karena orientasi lebih kepada deskripsi pelaksanaan program.
- 5) Hasil studi tersebut merupakan evaluasi, karena deskripsi pelaksanaan program memuat kriteria ketercapaian, dan berdasarkan ketercapaian disajikan masa depan program, yakni perlu diperbaikinya program tersebut.



RANGKUMAN

1. Pengukuran, penilaian, dan evaluasi merupakan hal-hal yang saling berkaitan, yang diperlukan pada pelaksanaan evaluasi program pendidikan.
2. Evaluasi program merupakan metode sistematis untuk mengumpulkan, menganalisis, dan menggunakan informasi untuk menjawab pertanyaan tentang proyek, kebijakan, dan program, khususnya tentang efektivitas dan efisiensinya.
3. Program pendidikan ini merupakan serangkaian kegiatan pendidikan yang mempunyai tujuan-tujuan tertentu
4. Evaluasi program dapat dilakukan secara internal maupun eksternal, secara sumatif maupun formatif.
5. Sejarah evaluasi dapat dikategorikan berdasarkan waktu perkembangannya.
6. Perbedaan antara penelitian dan evaluasi dapat diketahui dari motivasi peneliti dan evaluator, tujuan, aturan dan deskripsi, peran peneliti dan evaluator, otonomi penyelidikan (*inquiry*), sifat fenomena yang diakses, keberlakuan umum (*generalisabilitas*) kejadian yang menjadi fokus studi, teknik investigasi, keutamaan mengenai nilai, kriteria menilai aktivitas, dasar disiplin, dan *training*.



TES FORMATIF 2

Jawablah pertanyaan di bawah ini dengan tepat!

- 1) Perhatikanlah kasus A dan kasus B berikut!

Kasus A:

A ingin mengetahui apakah pendidikan anak usia dini (PAUD) mempengaruhi kesiapan belajar siswa di sekolah dasar. Untuk menjawab pertanyaan tersebut, A membandingkan sekelompok siswa yang mengikuti PAUD dan siswa yang tidak mengikuti PAUD, dan dibandingkan kesiapannya untuk belajar di Sekolah Dasar.

Kasus B:

B juga melakukan penelitian, yang juga ingin mengetahui kesiapan siswa belajar SD setelah melalui suatu program pendidikan PAUD. Seperti

halnya A, B juga membandingkan sekelompok siswa yang mengikuti PAUD dan siswa yang tidak mengikuti PAUD, dan dibandingkan kesiapannya untuk belajar di Sekolah Dasar. Setelah pertanyaan terjawab, B memberikan saran kepada pembuat kebijakan di PAUD, apakah program tersebut sudah berhasil atau memang benar-benar menyiapkan siswanya untuk belajar di SD dengan kriteria tertentu. Jika memang berhasil, B memberikan saran pembuat kebijakan untuk melanjutkan pendidikan PAUD, ataukah B menyarankan perbaikan pelaksanaan pendidikan PAUD, atau B menyarankan perlunya dihentikan pelaksanaan pendidikan PAUD karena tidak ada hasilnya. Dengan memperhatikan kasus A dan kasus B tersebut, manakah yang merupakan penelitian dan manakah yang merupakan evaluasi ? Jelaskan alasan Anda!

- 2) Berikan sepasang contoh lain berupa kasus beda antara penelitian dan evaluasi! Jelaskan mengapa termasuk kasus penelitian, sedangkan kasus lainnya evaluasi!

Tindak Lanjut

Untuk mengetahui tingkat penguasaan materi kegiatan belajar ini, yang perlu Anda lakukan adalah:

1. Buatlah rangkuman materi dengan mengambil konsep-konsep utama dan kata-kata kunci pada petunjuk pengerjaan.
2. Kerjakanlah latihan yang diberikan dengan mengikuti petunjuk jawaban latihan.
3. Cocokkan hasil latihan dengan petunjuk jawaban, dan berilah skor jawaban Anda. Jika sesuai dengan petunjuk jawaban, berilah skor masing-masing butir maksimal 5. Skor dapat disesuaikan dengan tingkat kebenaran jawaban Anda.
4. Hitunglah tingkat penguasaan Anda. Tingkat penguasaan Anda dikatakan baik jika mencapai minimal 80.

$$\text{Tingkat Penguasaan} = \frac{\text{Jumlah Skor Semua Butir}}{5 \times \text{Jumlah Soal}} \times 100\%$$

5. Diskusikanlah hasil latihan yang Anda kerjakan bersama teman-teman serta penjelasan tutor.
6. Apabila hasil latihan baik (minimal 80) dan adanya penguatan baik dari tutor, silakan melanjutkan ke materi berikutnya.

Kunci Jawaban Tes Formatif

Tes Formatif 1

Kata Kunci Pengerjaan	Skor Maksimum
1) Fokuskan perhatian pada suatu lembaga masyarakat atau lembaga pemerintah. Kemudian pilihlah beberapa program yang terdapat dalam lembaga tersebut, dengan ciri-ciri memiliki tujuan yang jelas dan terdiri atas serangkaian kegiatan.	5
2) Dari program yang Anda sebutkan pada nomor 1, kajilah latar belakang program lembaga tersebut, dasar hukumnya, dan uraikanlah kegiatan-kegiatan yang tercakup dalam program-program tersebut.	5
3) Contoh yang diberikan dapat bervariasi, tergantung apa yang akan diukur. Pada intinya, pengukuran merupakan kuantifikasi dari sesuatu, penilaian merupakan kegiatan memaknai hasil pengukuran, dan evaluasi membandingkan hasil penilaian dengan suatu kriteria.	5
4) Perkembangan evaluasi dapat dibedakan dari jaman perkembangannya, sebelum perang dunia I (tahun 1800-1940), setelah perang dunia I(1940-1964), antara tahun 1964-1972, antara 1973-1984, dan setelah tahun 1984 sampai sekarang.	5
Total Skor	20

Tes Formatif 2

Kata Kunci Pengerjaan	Skor Maksimum
1) Kasus A merupakan penelitian, karena lebih ingin menjawab permasalahan, sedangkan kasus B merupakan evaluasi, karena selain ingin menjawab permasalahan, juga ingin memberikan saran perbaikan.	5
2) Dapat dipilih kasus penelitian dan evaluasi, di mana penelitian hanya ingin mengetahui atau menjelaskan permasalahan saja, namun kalau evaluasi selain ingin mengetahui permasalahan, namun informasinya dapat digunakan untuk memberikan masukan perbaikan program yang dievaluasi.	5

Glosarium

- Evaluasi : kegiatan membandingkan hasil penilaian dengan kriteria untuk memberikan status seseorang.
- Evaluasi formatif : evaluasi yang dilakukan ketika program sedang berjalan.
- Evaluasi sumatif : evaluasi yang dilakukan ketika program telah selesai dilaksanakan
- Evaluasi program : merupakan metode sistematis untuk mengumpulkan, menganalisis, dan menggunakan informasi untuk menjawab pertanyaan tentang proyek, kebijakan, dan program, khususnya tentang efektivitas dan efisiensinya.
- Program : sekumpulan/serangkaian kegiatan yang merupakan satu kesatuan dan mempunyai tujuan, yang dilaksanakan oleh lembaga, baik pemerintah maupun masyarakat.
- Pengukuran : kegiatan menguantifikasi sesuatu, yang biasanya hasilnya dinyatakan dengan bilangan
- Penilaian : kegiatan membandingkan hasil pengukuran dengan kriteria.

Daftar Pustaka

- Allen, M.J.& Yen, W.M. 1979. *Introduction to Measurement Theory*. Belmont, CA: Wadsworth, Inc.
- Babbie, E. 2004. *The Practice of Social Research*. Belmont, CA: Wadsworth.
- Ebel, R.L. & Frisbie, D.A. 1986. *Essentials of Educational Measurement*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Djemari Mardapi. 1991. *Konsep Dasar Teori Respons Butir: Perkembangan dalam Bidang Pengukuran Pendidikan*. *Cakrawala Pendidikan* 3(X). 1-16.
- Fitzpatrick, J.L., Sanders, J.R., Worthen, B.R. 2004. *Program Evaluation: Alternative Approaches and Practical Guidelines*. Boston: Pearson.
- Frechtling, J.A. 2007. *Logic Modeling Methods in Program Evaluation*. San Francisco, CA: John Wiley & Sons.
- McDavid, J.C. & Hawthorn, L.L.R. *Program Evaluation and Performance Measurement: an Introduction to Practice*. Thousand Oaks, CA: Sage Publications.
- Rossi, P.H., Lipsey, M.W., & Freeman, H.E. 2004. *Evaluation: A Systematic Approach*. Thousand Oaks, CA: Sage Publications.
- Wholey, J.S., Hatry, H.P., Newcomer, K.H. 1994. *Handbook of Practical Program Evaluation*. San Francisco: Jossey-Bass Publishers.
- Worthen, B.R. & Sanders, J.R. 1973. *Educational Evaluation: Theory and Practice*. Worthington, Ohio: Charles A. Jhon.