

PENERAPAN RISET EVALUASI DALAM BIDANG PENDIDIKAN: SEBUAH PEDOMAN PRAKTIS

Oleh: Muhammad Akhyar

ABSTRACT

Evaluation Research is an applied social science that developing so fast nowadays. It develops due to increasing society needs. The results of this research are used for making public decisions. The model of evaluation that altered by researchers depend on the objective of evaluation that they design. CIPPO model - one of several existing evaluation models-can be altered, because it can provide some informations more comprehensive and accurate. Therefore this model suits on evaluating education program. The application of this research in education is not only based on the comprehensive model but also based on the complete methods.

Key words: *Riset evaluasi, model CIPPO, bidang pendidikan.*

I. Pendahuluan

Di negara-negara yang sedang berkembang termasuk Indonesia kualitas pelayanan sosial seperti pelayanan di bidang pendidikan, kesehatan, dan pelayanan publik lainnya masih belum memadai. Akibatnya di mana-mana bermunculanlah organisasi non pemerintah atau Lembaga Swadaya Masyarakat (LSM) yang mencoba mengambil fungsi sebagai penggerak dalam upaya memperoleh kualitas pelayanan tersebut.

Untuk meningkatkan kualitas pelayanan terhadap masyarakat peranan riset evaluasi menjadi penting. Belum banyak riset evaluasi yang berkualitas muncul kepermukaan. Riset evaluasi yang ada kebanyakan berupa survei awal yang tidak menyentuh substansi permasalahan. Bahkan banyak riset evaluasi dilakukan hanya untuk memenuhi pesanan pemerintah atau sponsor. Padahal temuan-temuan riset evaluasi dapat dijadikan masukan berharga bagi pemerintah atau bagi pemilik program untuk mengambil keputusan penting terutama yang berkaitan dengan perbaikan kualitas pelayanan pendidikan. Manakala kebijakan yang dihasilkan tidak tepat bahkan salah maka risiko sosial dan politisnya sangat tinggi, karena ia menyentuh kepentingan masyarakat banyak.

Riset evaluasi sebagai kegiatan ilmiah memerlukan perencanaan yang matang, desain yang tepat, instrumen pengumpulan data yang akurat, dan teknik analisis yang benar sebagaimana penelitian pada umumnya. Hanya saja kalau riset pada umumnya (non evaluasi) kegiatannya hanya sampai kepada menguji hipotesis atau menjawab pertanyaan-pertanyaan penelitian, sementara riset evaluasi dirancang untuk mengambil sebuah keputusan. Keputusan yang diambil berdasarkan kriteria-kriteria tertentu.

Perspektif periset tentang makna atau definisi evaluasi sangat mewarnai desain riset yang akan digunakan. Berbagai macam definisi evaluasi dikemukakan oleh para pakar evaluasi seperti Scriven, Tyler, Kaufman, Stufflebeam dan lain-lain. Tetapi dari sekian banyak definisi evaluasi yang ditawarkan oleh para pakar tersebut semuanya bermuara kepada empat definisi utama. Keempat definisi utama itu adalah pertama, proses menentukan seberapa jauh sasaran sebuah program telah tercapai; kedua, proses menyajikan informasi guna mengambil keputusan; ketiga, proses penilaian terhadap nilai atau manfaat sebuah program; dan keempat, proses membandingkan pelaksanaan program dengan standar ukuran yang telah ditentukan.

Berdasarkan konteks persoalan di atas betapa penting melakukan riset evaluasi secara tepat dan akurat agar temuan-temuannya dapat dijadikan dasar perbaikan program pendidikan. Artikel ini mencoba memaparkan penerapan riset evaluasi dalam konteks dunia pendidikan.

II. Prosedur Riset Evaluasi

Ketika periset menerima tawaran pekerjaan untuk mengevaluasi program pendidikan, periset tidak serta merta menerima tawaran pekerjaan tersebut. Ia terlebih dahulu menetapkan evaluabilitas program yang hendak dievaluasi. Bila telah terjadi kesepakatan antara periset dan pemesan evaluasi, maka barulah ia menerima tawaran tersebut. Kesepakatan tersebut mencakup beberapa hal yakni mengapa program terkait perlu dievaluasi, apakah ada pesan pihak-pihak tertentu yang bersifat politis sehingga mengakibatkan biasanya hasil evaluasi dan apakah pengambilan data akan mengakibatkan risiko karena responden merasa terusik masa depannya.

Temuan-temuan riset evaluasi tak seperti temuan riset pada umumnya. Temuan riset evaluasi memerlukan konfirmasi atau penguatan dari responden. Hal ini penting dilakukan untuk mendapatkan temuan yang akurat sehingga tak ada pihak-pihak yang dirugikan. Besar kemungkinan temuan riset evaluasi meskipun telah dianalisis atas dasar data yang diberikan responden mendapat

penolakan dari responden itu sendiri. Dalam konteks persoalan tersebut periset perlu melakukan langkah-langkah riset yang tepat dan akurat. Posavac dan Carey (1980) mengusulkan langkah-langkah perencanaan riset evaluasi berikut ini.

Langkah pertama adalah mengkaji literatur. Langkah ini sama halnya dengan penelitian pada umumnya. Mengkaji literatur sebelum mendesain riset atau sebelum mengembangkan instrumen adalah penting. Melalui artikel atau jurnal penelitian misalnya periset dapat mempelajari keberhasilan atau kegagalan periset lain, di antaranya mengenai metode riset yang tepat atau mempelajari kesulitan apa yang dihadapi periset ketika berada di lapangan. Selain itu periset dapat mengetahui apakah riset terhadap program sejenis telah pernah dilakukan, desain apa yang digunakannya, apakah teknik pengukuran baru telah dikembangkan, dan jenis analisis apa yang telah digunakan. Informasi ini menjadi sangat berharga bagi periset untuk melaksanakan sebuah riset evaluasi.

Langkah kedua adalah menentukan metode riset. Setelah mempelajari literatur periset siap untuk mengambil keputusan metodologis berkenaan dengan strategi dan desain, teknik pengumpulan data dan analisis data.

Langkah ketiga adalah menyajikan proposal tertulis untuk dinegosiasikan kepada pemilik program. Setelah masing-masing pihak yakni periset dan pemilik program atau pemesan evaluasi menyetujui proposal tersebut barulah periset mulai melakukan riset. Ketiga langkah riset yang ditawarkan Posavac dan Carey bukan saja harus dilakukan oleh periset evaluasi tapi juga oleh periset pada umumnya. Hanya saja dalam praktiknya, riset evaluasi memerlukan negosiasi dan kesepakatan tentang proposal secara lebih informal dibandingkan dengan riset pada umumnya (non evaluasi). Secara lebih spesifik *Improving the Efficiency of Educational Systems* (1986) menawarkan sebelas langkah pelaksanaan riset evaluasi.

Pertama adalah merumuskan titik pandang (*formulate a point of view*). Dalam tahap ini periset mencoba mempertimbangkan beberapa macam model evaluasi. Model-model evaluasi ini diwarnai oleh perbedaan pemahaman peneliti tentang makna mengevaluasi. Perbedaan umumnya terletak pada perspektif filosofi dan ideologi tentang evaluasi. Tidak ada titik pandang yang terbaik. Periset dapat menggunakan model evaluasi yang berbeda pada waktu yang berbeda atau dapat menggunakan kombinasi dari model yang ada. Bahkan periset dapat menggunakan model evaluasi yang ia kembangkan sendiri. Tentu saja pengembangan tersebut didasari oleh pemikiran-pemikiran yang mempunyai alasan ilmiah yang cukup kuat. Titik pandang periset tentang makna evaluasi akan menentukan desain evaluasinya.

Kedua adalah mengidentifikasi tujuan evaluasi (*identify purpose*) yang diminta pemesan evaluasi. Berbagai macam alasan mengapa evaluasi dilakukan. Terkait dengan alasan tersebut, pada umumnya ada tiga jenis evaluasi yakni analisis kebutuhan (*needs assessment*), evaluasi formatif (*formative evaluation*) dan evaluasi sumatif (*sumative evaluation*). Analisis kebutuhan dilakukan terhadap program yang belum berjalan. Evaluasi formatif dilakukan terhadap program yang sudah berjalan dengan tujuan untuk melakukan perbaikan-perbaikan program. Evaluasi sumatif dilakukan untuk menjustifikasi apakah sebuah program perlu diteruskan atau dihentikan. Yang perlu diperhatikan bahwa jenis evaluasi apapun yang dilakukan, periset hanya bertugas memberikan informasi belaka dan tidak berhak melakukan pemaksaan untuk harus melaksanakan hasil-hasil riset evaluasi. Kebijakan yang dikeluarkan sebagai aplikasi dari temuan riset evaluasi hanya wewenang pemilik, pengelola program atau pemesan evaluasi.

Ketiga adalah mengenal klien (*identify client*). Pertanyaan-pertanyaan seperti siapa yang meminta evaluasi dilakukan, siapa yang mendanai evaluasi, dan kepada siapa hasil evaluasi dilaporkan adalah pertanyaan-pertanyaan penting dalam riset evaluasi. Di awal sebuah kegiatan evaluasi periset perlu mengenal secara jelas sponsor, klien, atau orang-orang yang berkepentingan terhadap program yang dievaluasi (*stakeholders*). Sponsor adalah kelompok yang mendanai riset. Klien adalah orang yang memesan riset. Dalam beberapa hal sponsor dan klien adalah sama. Audien kunci atau *stakeholders* adalah kelompok yang memiliki perhatian khusus dalam evaluasi tersebut. Mereka adalah orang atau kelompok yang berkepentingan terhadap program yang di evaluasi dan merasakan dampak evaluasi.

Keempat adalah mengenal audien dan sponsor (*identify audiences and sponsor*). Pertanyaan-pertanyaan yang penting dari langkah ini adalah siapa yang membutuhkan hasil evaluasi dan siapa penyandang dananya. Periset dalam hal ini harus mewaspadai bahwa tidak menutup kemungkinan periset mendapat tekanan-tekanan untuk mengarahkan temuan evaluasinya menurut kemauan pemesan. Biasanya hal ini terkait dengan kebijakan politis.

Kelima adalah mengenal sumber-sumber dana dan kendala (*identify resources and constrains*). Bahan apa, siapa, waktu, kendala, dan data yang tersedia sebelumnya adalah informasi penting bagi periset. Untuk mengidentifikasi kendala yang ada Kaufman (1980) menawarkan beberapa langkah yang harus ditempuh yakni dengan melakukan analisis kebutuhan (*needs assessment*). Langkah yang pertama adalah menemukan kesenjangan antara *apa yang sudah ada* dengan *apa yang harus ada*; kedua, menempatkan kesenjangan itu secara berurutan sesuai dengan tingkat kepentingannya; ketiga, kesenjangan yang paling krusial dikelompokkan

RIWAYAT HIDUP PENULIS

Nur Saidah, lahir di Kediri, 11 Februari 1975 adalah dosen tetap Fakultas Tarbiyah UIN Sunan Kalijaga dpk. pada Institut Seni Indonesia Yogyakarta dengan matakuliah pokok Pendidikan Agama Islam. Menyelesaikan S-1 bidang Bahasa dan Sastra Arab Fakultas Adab lulus tahun 1998 dan S-2 di bidang Pemikiran Pendidikan Islam Program Studi Pendidikan Islam IAIN Sunan Kalijaga Yogyakarta tahun 2003. Pelatihan yang pernah diikuti antara lain Pelatihan Pengembangan Program Pengajaran UPT MPK ISI Yogyakarta, September 2005, Pelatihan Penjaminan Mutu Perguruan Tinggi UPT MPK ISI Yogyakarta Januari 2006, TOT *English Teaching* di Pusat Bahasa UIN Sunan Kalijaga, Februari-Juni 2006, Workshop Metodologi Penelitian Interdisipliner II di Lembaga Penelitian UIN Sunan Kalijaga 2006, Pelatihan Pengembangan Calon Dosen UIN Sunan Kalijaga (Kompetensi Kepribadian, November 2006 dan Kompetensi Pedagogik, Juni 2007) dan Workshop Penelitian Kualitatif Lembaga Penelitian UIN Sunan Kalijaga, Juli 2007.

Maragustam Siregar, lahir di Tapanuli Selatan, 1 Oktober 1959 adalah dosen tetap Fakultas Tarbiyah UIN Sunan Kalijaga. Pendidikan jenjang S-1, S-2, dan S-3 diselesaikan di UIN. Saat ini menjabat sebagai Pembantu Rektor III Bidang Kemahasiswaan. Aktif mengikuti berbagai kegiatan keilmuan, workshop, seminar, dan pelatihan-pelatihan baik tingkat local maupun nasional.

Ichsan, lahir di Boyolali pada 26 Februari 1963, adalah dosen Fakultas Tarbiyah UIN Sunan Kalijaga Yogyakarta. Ia alumni Sarjana Muda (BA) Jurusan Tadris IPA Fak. Tarbiyah IAIN Sunan Kalijaga tahun 1985 kemudian menyelesaikan Sarjana (S1) jurusan PAI Pada Fakultas yang sama tahun 1989 dan S2 lulus tahun 2004 dari Universitas Negeri Yogyakarta (UNY) dengan konsentrasi Pendidikan Nilai. Ia aktif dalam berbagai kegiatan pelatihan, seminar dan lokakarya dan pada tahun 2007 mengikuti *short Corse* Pendidikan Dasar di Pascasarjana UNY dalam rangka pembukaan Prodi PGMI Fak. Tarbiyah UIN Sunan Kalijaga Yogyakarta.

Karwadi, lahir di Jepara, 15 Maret 1971 dosen tetap Fakultas Tarbiyah UIN Sunan Kalijaga Yogyakarta. Pendidikan jenjang Sarjana (S-1) di selesaikan di IAIN Jambi, S-2 IAIN Ar-Raniry Banda Aceh, dan sekarang sedang menyelesaikan penulisan disertasi untuk mnyelesaikan studi S-3 di UINSunan Kalijaga. Di samping mengajar juga aktif mengikuti kegiatan seminar, workshop dan peltaihan-pelatihan.

Muhammad Akhyar, lahir di Pangkalpinang tanggal 29 Juli 1961. Bekerja sebagai dosen di Program Studi Pendidikan Teknik Mesin FKIP UNS sejak tahun 1991. Pendidikan sarjana (S1) Pendidikan Teknik Mesin Jurusan Pendidikan Teknik dan Kejuruan (1987). Pendidikan Pascasarjana Penelitian dan Evaluasi Pendidikan diselesaikan tahun 1996. Di samping mengajar, aktif juga dalam berbagai kegiatan ilmiah, baik lokal, nasional maupun internasional. Beberapa karya tulis, baik buku, makalah dan hasil penelitian juga telah dihasilkan.

Supa'at, adalah dosen tetap pada Jurusan Tarbiyah STAIN Kudus, saat ini sedang menyelesaikan studi program doktor pada Universitas Negeri Yogyakarta, Prodi Penelitian dan Evaluasi Pendidikan

Atikah Syamsi, adalah alumni Jurusan PAI, lulus tahun 2007. Semasa kuliah aktif dalam organisasi intra kampus dan berbagai kegiatan lain di luar kampus yang berhubungan dengan keilmuan.

sebagai prioritas kendala; keempat, mengatasi kendala tersebut. Kendala yang paling penting dan mendesak harus didahulukan penanganannya.

Keenam adalah menentukan pertanyaan evaluasi (*specify the evaluation question*). Banyak isu penting yang dihadapi oleh audien. Isu tersebut bervariasi dan kadangkala semuanya penting. Periset tidak harus mengambil semua isu untuk tujuan evaluasi. Ia harus memilih mana yang paling penting dan mendesak. Identifikasi isu amat penting. Banyak periset mengalami kegagalan karena riset evaluasinya tidak mengarah kepada isu penting tersebut. Ada beberapa alasan mengapa sebuah riset evaluasi gagal mengarah kepada isu penting, yakni pertama, karena tidak ada pertanyaan evaluasi yang dirumuskan, atau gagal merumuskan pertanyaan evaluasi secara jelas sehingga kegiatan riset mereka tidak didasari oleh pertanyaan penelitian; kedua, pertanyaan evaluasi pada awal riset masih relevan tetapi kemudian perlahan-lahan kepedulian periset terhadap kliennya berubah. Pada umumnya hal tersebut terjadi pada periset muda. Oleh sebab itu untuk menghasilkan temuan yang akurat dibutuhkan pengalaman dan komitmen yang tinggi.

Ketujuh adalah merumuskan desain evaluasi (*formulate an evaluation design*). Salah satu pertimbangan yang penting untuk menentukan desain evaluasi adalah strategi menentukan responden dan prosedur mengumpulkan data. Periset juga harus selalu mengingat-ingat perbedaan antara riset evaluasi dengan riset pada umumnya ketika melakukan riset evaluasi. Tuntutan hubungan sebab akibat yang kuat seperti pada penelitian korelasional bukan menjadi bagian dari riset evaluasi. Tujuan riset evaluasi antara lain adalah untuk mendukung keputusan tentang nilai program berdasarkan kriteria tertentu –tidak sekedar meneliti hubungan kausalitas. Generalisasi bagi riset evaluasi bukan hal yang penting, karena riset evaluasi didesain untuk mengarah pada keputusan tertentu dalam konteks tertentu yang relevan dengan program. Ini bukan berarti riset evaluasi tidak mampu menemukan teori.

Kesesuaian desain dengan tujuan riset adalah persoalan dasar dalam memilih desain evaluasi. IEES (1986) menyarankan alternatif desain yang dapat digunakan periset tergolong dalam tiga kategori yakni desain eksperimental, korelasional, dan desain kualitatif. Pemilihan desain tersebut melalui beberapa pertimbangan yakni 1) sejauhmana riset mengikuti prosedur yang tepat dan sistematis; 2) sejauhmana validitas internal dan eksternalnya; 3) sejauhmana desain evaluasi sesuai dengan tujuan riset dan pertanyaan evaluasi; 4) sejauhmana desain sesuai dengan konteks sosial; dan 5) sejauhmana feasibilitas riset. Ini berarti bahwa kaedah penelitian ilmiah juga harus diikuti secara ketat oleh periset evaluasi.

Kedelapan adalah memilih prosedur pengumpulan data (*select a data collection procedure*). Riset evaluasi menggunakan banyak pilihan prosedur pengumpulan data seperti, kuesioner, tes, wawancara, observasi partisipan, dan simulasi bahkan periset evaluasi sangat disarankan untuk menggabungkan berbagai metode pengumpulan data agar kesimpulan yang dihasilkan lebih akurat. Periset evaluasi yang paling berhasil adalah yang memiliki bermacam-macam keterampilan dalam pengumpulan data.

Kesebelasan adalah mengumpulkan data (*collect data*). Hal-hal yang penting dipertanyakan adalah bagaimana periset dapat menjamin data dikumpulkan tanpa bias. Seringkali data riset evaluasi yang terkumpul sangat rendah kualitasnya. Hal ini disebabkan oleh dua hal yakni pertama, tidak adanya instrumen yang cocok; dan kedua periset sering mengalami kegagalan dalam memahami kultur responden sehingga gagal mendapatkan data yang tepat. Namun bila periset memegang komitmen yang tinggi tentang prinsip-prinsip riset yang ketat dan peka terhadap realitas responden maka hal tersebut tak akan terjadi.

Kesepluluh adalah menganalisis data (*analyze data*). Pertanyaan yang paling penting diajukan dalam langkah ini adalah prosedur apa yang paling sederhana, paling jelas, dan paling cocok untuk menganalisis data. Seringkali audien tidak menyukai analisis statistik yang canggih karena umumnya mereka sudah terbiasa dengan metode kualitatif. Akibatnya mereka tidak menghargai hasil riset evaluasi yang telah dilakukan dengan susah payah oleh periset.

Kesebelas adalah menginterpretasikan dan melaporkan hasil riset (*interpret and report results*). Pertanyaan yang penting diajukan adalah bagaimana melaporkan data yang terkumpul. Bagaimana hasil riset dilaporkan sehingga pembaca dapat memahaminya dengan cara yang paling mudah. Periset kadangkala tak dapat menghindari dari analisis yang sarat dengan angka-angka statistik. Namun bila periset mampu melaporkan hasil temuannya dengan ramuan bahasa yang baik dan mudah dipahami, maka hal ini akan memberi peluang diterimanya temuan tersebut secara baik oleh audien.

Sebelas langkah riset di atas pada dasarnya sama dengan langkah riset pada umumnya (riset non evaluasi) hanya saja riset evaluasi tidak menekankan pada kausalitas dan generalisasi tetapi ia lebih mengutamakan menentukan nilai suatu program dan membantu memberikan alternatif keputusan tentang program berdasarkan kriteria-kriteria tertentu.

III. Kriteria Riset Evaluasi

Riset evaluasi adalah sebuah kegiatan mengukur dan menilai. Sebuah program tak dapat dinilai bila tidak ada kriteria sebagai tolok ukur pembandingan. Periset akan kesulitan ketika memberikan pertimbangan nilai tanpa adanya kriteria tertentu. Tanpa kriteria, pertimbangan yang diberikan adalah tanpa dasar. Dengan kata lain tak ada evaluasi tanpa kriteria. Hasan (1988) mengatakan bahwa ada empat macam pendekatan dalam menemukan kriteria evaluasi. Keempat pendekatan tersebut yakni, pendekatan *preordinate*, pendekatan *fidelity*, pendekatan *process*, dan pendekatan *mutual adaptive*.

Kriteria yang dikembangkan dengan pendekatan *preordinate* ditentukan sebelum evaluasi dilaksanakan. Kriteria ini bersifat mengikat dan tak akan diubah sampai riset itu selesai. Kriteria ini dikembangkan berdasarkan pendapat teoritis atau kesepakatan para pakar. Dengan demikian kriteria ini bersifat umum. Kriteria evaluasi dengan pendekatan *fidelity* ditentukan sebelum evaluasi dilaksanakan seperti halnya pendekatan *preordinate*. Tetapi kriteria ini dikembangkan berdasarkan persepsi para pengembang program yang akan dievaluasi. Dengan demikian kriteria ini bersifat lokal. Kriteria evaluasi dengan pendekatan *process* dikembangkan selama proses evaluasi berlangsung. Kriteria ini diperoleh antara lain melalui wawancara, observasi, atau studi dokumentasi. Dengan demikian kriteria yang ditetapkan sangat terikat dengan masalah yang dihadapi oleh pelaksana program. Kriteria dengan pendekatan *mutual adaptive* dikembangkan dari gabungan ketiga pendekatan sebelumnya yakni *preordinate*, *fidelity*, dan *process*. Kriteria ini dikembangkan berdasarkan karakteristik program yang dievaluasi misalnya dari pandangan teoritis para pelaksana dan pemakai program.

Pemilihan pendekatan kriteria yang akan digunakan periset bergantung kepada tujuan riset itu sendiri. Makin universal pendekatan yang digunakan maka makin tinggi daya aplikatif dan transferebilitasnya. Dengan kata lain makin universal pendekatan kriteria digunakan, maka makin luas manfaat hasil riset. Dengan demikian pendekatan kriteria yang dipilih sangat bergantung kepada tujuan evaluasi yang diinginkan dan ketersediaan dana.

IV. Penerapan dalam Bidang Pendidikan

Penerapan riset evaluasi dalam pendidikan harus dilakukan secara cermat. Kebijakan pendidikan yang salah sebagai akibat pelaksanaan riset evaluasi yang salah akan memberikan dampak yang sangat luas bagi masyarakat. Oleh sebab itu data atau informasi yang dijaring oleh riset evaluasi harus dilakukan secermat mungkin sehingga dampak akibat kesalahan pengukuran dapat ditekan.

Sebuah pendekatan evaluasi yang paling komprehensif terhadap suatu program adalah pendekatan yang ditawarkan oleh Stufflebeam. Model tersebut mencakup empat komponen yakni *Context*, *Input*, *Process* dan *Product*. Komponen *Context* berkaitan kelemahan dan kekuatan program. Komponen *Input* berkaitan dengan masukan yang akan diproses. Komponen *Process* berkaitan dengan pengolahan *Input* untuk mencapai tujuan. Komponen *Product* berkaitan dengan hasil dari proses.

Untuk mengevaluasi program pendidikan, pendekatan yang ditawarkan Stufflebeam merupakan sebuah pilihan. Pilihan ini didasari oleh tuntutan untuk memperoleh kecermatan data. Makin lengkap model evaluasi makin lengkap pula data atau informasi yang diperoleh. Oleh sebab itu model evaluasi tidak hanya berkaitan dengan keempat komponen model di atas tetapi juga ditinjau dari komponen *Outcomes*. *Outcomes* menunjuk pada manfaat program menurut penilaian masyarakat luas atau pengguna tenaga lulusan. Dengan demikian, untuk mengevaluasi pelaksanaan pendidikan meliputi lima komponen, yakni *Context*, *Input*, *Process*, *Product*, dan *Outcomes* (model CIPPO).

Dalam hubungannya dengan penyusunan instrumen evaluasi, sebaiknya instrumen yang digunakan untuk mengevaluasi program pendidikan menggunakan skala semantik diferensial. Sofian Effendi (1989) mengatakan bahwa instrumen dengan skala semantik diferensial cukup sesuai digunakan pada penelitian evaluatif, karena skala tersebut menunjukkan unsur kualitas dari objek yang diteliti. Beberapa pakar evaluasi menyarankan skala penilaiannya paling sedikit berjumlah 5 dan paling banyak 9. Namun pada umumnya untuk memudahkan dalam menganalisis periset biasanya menggunakan skala berjumlah 5. Tetapi kalau periset dituntut untuk menghasilkan data yang lebih responsif dan akurat maka ia dapat menggunakan jumlah skala 9. Berikut ini adalah penerapan model evaluasi yang merupakan pengembangan model yang ditawarkan Stufflebeam.

Tabel 1. Kisi-kisi Model CIPPO dalam Konteks Pendidikan

JENIS EVALUASI	KOMPONEN	INDIKATOR
INPUT	1. Persiapan kultur	1.1. Persiapan kedisiplinan 1.2. Persiapan etos kerja 1.3. Persiapan budaya membaca 1.4. Persiapan budaya komunikasi 1.5. Persiapan iklim kerjasama 1.6. Persiapan iklim saling percaya
	2. Persiapan fasilitas pendukung belajar	2.1. Persiapan ruangan belajar 2.2. Persiapan perpustakaan 2.3. Persiapan laboratorium 2.4. Persiapan gedung sekolah 2.5. Persiapan administrasi/staf
	3. Persiapan sumber daya manusia	3.1. Kesempatan studi lanjut 3.2. Kesempatan dalam training 3.3. Akses ke lab komputer 3.4. Kesempatan berkreasi
PROCESS	1. Pelaksanaan Kurikulum	1.1. Model pembelajaran 1.2. Variasi metode mengajar 1.3. Pemanfaatan fasilitas 1.4. Iklim pembelajaran
	2. Kendala pelaksanaan kurikulum	2.1. Kendala kultur sekolah 2.2. Kendala fasilitas belajar 2.3. Kendala sumber daya manusia
	3. Keterlibatan guru	3.1. Mengembangkan gaya belajar 3.2. Mengembangkan variasi metode mengajar 3.3. Memberdayakan fasilitas belajar 3.4. Membangun kreativitas siswa
	4. Keterlibatan kepala sekolah	4.1. Mengunjungi kelas 4.2. Memberikan pengarahan kepada staf 4.3. Kepedulian terhadap PBM
	5. Keterlibatan siswa	5.1. Komitmen terhadap belajar 5.2. Penggunaan fasilitas belajar

Tabel 1. Kisi-kisi Model CIPPO dalam Konteks Pendidikan (Lanjutan)

JENIS EVALUASI	KOMPONEN	INDIKATOR
PROCESS	6. Keterlibatan masyarakat	6.1. Kepedulian ortu terhadap sekolah 6.2. Kepedulian ortu terhadap kemajuan anaknya 6.3. Kepedulian pengguna tenaga kerja terhadap PBM di sekolah
PRODUCT	1. Penguasaan pengetahuan	1.1. Pengetahuan faktual 1.2. Pengetahuan konseptual 1.3. Pengetahuan prosedural 1.4. Pengetahuan metacognitif
	2. Kemampuan psikomotor	2.1. Keterampilan lulusan 2.2. Prestasi kerja lulusan
	3. Kemampuan afektif	3.1. Sikap terhadap bidangnya 3.2. Minat terhadap bidangnya 3.3. Nilai kerja 3.4. Sikap kerja
OUTCOMES	1. Manfaat bagi masyarakat pengguna lulusan	1.1. Kesiapan kerja lulusan 1.2. Kesiapan berkarier lulusan 1.3. Kemampuan menopang hidup

Tabel di atas menunjukkan bahwa setiap jenis evaluasi mencakup beberapa komponen dan setiap komponen dijabarkan menjadi beberapa indikator. Indikator tersebut menjadi dasar penyusunan instrumen evaluasi. Apabila periset evaluasi secara konsisten dan utuh menggunakan model CIPPO di atas, maka akurasi data yang terkumpul akan lebih terjamin. Riset evaluasi yang hanya menggunakan satu metode pengumpulan data misalnya angket betapapun baik kualitasnya, maka temuan penelitian belum terjamin keakuratannya. Sifat riset evaluasi adalah melibatkan banyak metode. Makin banyak metode yang digunakan maka makin lengkap dan akurat data yang diperoleh. Pendekatan model CIPPO dalam mengevaluasi program pendidikan bukan satu-satunya model yang dapat digunakan. Masih banyak lagi model evaluasi lain sebagai pilihan.

V. Penutup

Periset evaluasi adalah ilmuwan sosial yang bekerja untuk meningkatkan kualitas pelayanan kepada sekelompok manusia bukan perorangan. Gerakan konsumen, *keprihatinan profesional*, upaya meningkatkan keefektifan manajerial, pembatasan sumber-sumber, dan mandat pemerintah telah mendorong pemilik program untuk membuat keputusan yang rasional guna meningkatkan kualitas pelayanan kepada masyarakat.

Sebelum merencanakan riset evaluasi periset perlu mengambil keputusan tertentu sebelum pengumpulan data dimulai. Periset perlu mempelajari tujuan dan mekanisme kerja program, siapa yang mensponsori program, siapa yang ada di dalamnya, dan bahkan kelompok mana yang nampaknya cenderung menentang riset tersebut. Setelah mempelajari informasi ini, barulah periset memutuskan apakah evaluasi dapat dilakukan atau tidak.

Ada tiga prinsip yang dapat membimbing periset dalam melakukan prosedur pengumpulan data yakni periset harus transparan terhadap bagaimana cara mengumpulkan data dan keterbatasannya, data sebaiknya dapat dikumpulkan melalui beberapa teknik pengumpulan data, dan peneliti harus mencari cara untuk menggunakan pengukuran ganda.

Sebelum memutuskan untuk melakukan riset evaluasi dan sebelum menulis proposal secara lengkap, periset perlu melakukan pertemuan awal dengan orang-orang relevan untuk memperoleh informasi tentang siapa yang ingin dievaluasi, jenis evaluasi apa yang diinginkan –needs assessment, evaluasi formatif atau sumatif-, mengapa mereka minta dievaluasi, kapan program akan dievaluasi, dan sumber apa yang tersedia. Setelah menyusun pertemuan awal periset kemudian menetapkan evaluabilitas program. Penetapan evaluabilitas program ini dilakukan dengan metode deskriptif dan analitik untuk memperoleh kepastian apakah evaluasi dapat dilakukan atau tidak. Penerapan riset evaluasi dalam pendidikan menggunakan prinsip bahwa riset harus mampu memberikan informasi yang komprehensif dan akurat. Salah satu model yang dapat dipilih adalah model CIPPO. Meskipun model ini bukan satu-satunya alternatif model tetapi model ini mampu memberikan informasi yang komprehensif. Tingkat keakuratan informasi dapat diperoleh dengan penerapan berbagai macam metode pengumpulan data secara terpadu. Dengan menggunakan model CIPPO dan metode pengumpulan data yang komplit, temuan riset akan dapat memberikan sumbangan yang bukan saja menyeluruh tetapi juga akurat. Penerapan riset evaluasi dalam pendidikan sangat dianjurkan menggunakan pendekatan ini.

DAFTAR PUSTAKA

- Improving the Efficiency of Educational Systems (IEES). (1986). *The Evaluation of Efficiency in Educational Development Activities*. Florida: Educational Efficiency Clearinghouse Learning Systems Institute.
- Kaufman, Roger. (1980). *Evaluation Without Fear*. New York: Library of Congress Cataloging.
- Posavac, E.J., and Carey, R.G. (1980). *Program Evaluation: Methodes and Case Studies*. New Jersey: Prentice Hall, Inc.
- Said Hamid. Hasan. (1988). *Evaluasi Kurikulum*. Jakarta: Ditjendikti Depdikbud
- Sofian Effendi (1989). *Metode Penelitian Survei*. Jakarta: LP3ES.