

## **PENDEKATAN BERORIENTASIKAN KONSUMEN SCHRIVEN TERHADAP EVALUASI**

Peranan yang tepat dari seorang penilai, menurut Michael Scriven, adalah peranan “menjadi wakil konsumen yang mengetahui hal yang sebenarnya”. Dilengkapi dengan keahlian-keahlian untuk memperoleh informasi yang akurat dan ada hubungannya serta dengan pandangan yang dipertimbangkan secara mendalam terhadap etika dan kebaikan umum, penilai harus membantu para profesional di dalam menghasilkan produk dan jasa yang berkualitas tinggi dan sangat berguna bagi konsumen. Yang lebih penting, penilai harus membantu konsumen untuk mengenal dan menilai barang-barang dan jasa-jasa alternatif. Scriven menunjukkan sikap yang sangat kritis terhadap ideologi-ideologi evaluasi yang memfokuskan pada pencapaian tujuan-tujuan developer (pengembang) sebagai yang terbukti bertentangan dengan tujuan memenuhi kebutuhan-kebutuhan konsumen; ia telah mengemukakan sejumlah konsep dan metode yang dimaksudkan untuk mengalihkan evaluasi dari orientasi tujuannya ke orientasi berdasarkan atas kebutuhan-kebutuhan.

Evaluasi sumatif dapat membantu di dalam memungkinkan para penyelenggara untuk memutuskan apakah kurikulum yang telah diselesaikan seluruhnya dan disaring melalui digunakannya proses evaluasi menurut peranan pertamanya (formatif)

menunjukkan kemajuan yang cukup berarti terhadap alternatif-alternatif yang ada untuk mengesahkan pengeluaran adopsi oleh sebuah sistem sekolah.

Michael Scriven adalah seorang pakar filsafat ilmu pengetahuan yang telah banyak sekali menyumbang kepada perkembangan profesi evaluasi. Ia mengkritik secara tajam konsepsualisasi evaluasi klasik dan modern. Ia mendasarkan pandangan konsumerisnya tentang evaluasi pada suatu posisi penting filsafat dan telah menghasilkan konsep-konsep dan metode-metode untuk membantu mengartikulasi dan menerapkan pendekatannya. Ia juga telah menjadi salah seorang pemimpin terkemuka di dalam upaya untuk memprofesionalisasikan pekerjaan evaluasi.

Ia dilahirkan dan dibesarkan di Victoria, Australia. Ia memperoleh gelar sarjana dalam bidang matematika terapan dan logika simbolik di University of Melbourne dan menyelesaikan gelar Ph.D.-nya dalam bidang filsafat ilmu pengetahuan di Oxford University. Ia mengabdikan diri di dalam peranan profesionalnya di University of Minnesota dari tahun 1952 sampai 1956, di Swarthmore College dari tahun 1956 sampai 1960, Indiana University dari tahun 1960 sampai 1965, University of California di Berkeley dari tahun 1965 sampai 1975, dan University of San Francisco dari tahun 1975 sampai 1982. Dengan karya tulisnya, ia

menjadi Guru Besar Universitas Terkemuka di University of Western Australia di Perth.

Ia memilih untuk menamakan pendekatan Dr. Scriven sebagai “evaluasi yang berorientasikan pada konsumen” untuk menggolongkan pendekatan filsafat utamanya. Ia mendefinisikan evaluasi sebagai penilaian sistematis tentang hal-hal yang berharga atau bermanfaat dan menekankan bahwa para penilai harus mampu sampai pada keputusan-keputusan nilai yang dapat dipertahankan daripada sekedar mengukur sesuatu atau menentukan apakah tujuan telah dicapai atau belum. Daripada mengakui tujuan-tujuan seorang developer sebagai yang sudah ditentukan, menurut Scriven seorang penilai harus menilai apakah tujuan-tujuan yang dicapai akan menyumbang kepada kesejahteraan konsumen atau tidak. Terlepas dari tujuan-tujuan tersebut, penilai harus mengetahui hasil-hasil yang sesungguhnya dan memperkirakan nilainya dari perspektif kebutuhan-kebutuhan konsumen. Bagi Scriven, peranan profesional penilaian adalah menjadi “wakil konsumen yang mengetahui hal yang sebenarnya”, membantu sebagai yang mewakili suara hati masyarakat yang diberitahu adalah apa yang ia pandang sebagai “batu fondasi etika profesional di dalam pekerjaan evaluasi”. Pendekatannya yang praktis terhadap evaluasi pada umumnya menuntut upaya mengidentifikasi dan menggolongkan program-

program dan produk-produk pilihan yang tersedia bagi konsumen, berdasarkan atas harga dan pengaruhnya yang relatif serta dengan memperhatikan kebutuhan-kebutuhan konsumen yang dinilai. Ia sering memperkenalkan majalah *Consumer Reports* sebagai contoh yang baik dari apa yang dapat disumbangkan oleh evaluasi profesional.

Di dalam unit ini kami meneliti dan menelaah kontribusi-kontribusi teoritis, metodologis, dan profesional dari Dr. Scriven. Penilaiannya yang kritis tentang pendekatan-pendekatan alternatif diperkenalkan dan dibahas. Kami membahas konsepsualisasi formatif-sumatif dan memperkenalkan konsep-konsep pekerjaan utamanya. Kritiknya pada tahun 1983 terhadap ideologi-ideologi evaluasi yang kuat dan beberapa saran metodologisnya juga dibahas. Kami menutup bab ini dengan mencatat prestasi-prfestasinya di dalam membentuk Jaringan Evaluasi dan *Evaluation News*.

Tujuan-tujuan instruksional unit ini dicatat di bawah ini. Perhatikan tentang tujuan-tujuan ini, bacalah tulisan berikutnya, kemudian selesaikan ujian pengetahuan dan pertanyaan-pertanyaan aplikasi yang muncul pada bagian akhir unit ini. Menilai jawaban-jawaban anda terhadap pertanyaan-pertanyaan penting, dan jika ada jawaban anda yang tidak benar, perhatikan bagian-bagian yang ada hubungannya dari unit ini. Anda mungkin

ingin melakukan penelitian independen lebih lanjut dengan mencoba “Pertanyaan-pertanyaan Tanpa Jawaban” (atau membahasnya bersama teman-teman mahasiswa anda) dan menanyakan sumber-sumber informasi tentang karya Dr. Scriven yang dicatat di bagian belakang unit ini.

### **Tujuan**

Unit ini dimaksudkan untuk membantu anda mencapai tujuan-tujuan berikut :

1. Mengetahui dan menerangkan definisi evaluasi dari Scriven.
2. Mengetahui dan menilai kritikan-kritikannya terhadap pandangan-pandangan lain tentang evaluasi oleh Tyler dan Cronbach.
3. Menerangkan posisinya tentang peranan keputusan di dalam evaluasi.
4. Menerangkan peranan dari penilaian kebutuhan-kebutuhan di dalam pendekatannya.
5. Menerangkan perbedaan-perbedaannya antara evaluasi dan estimasi pencapaian tujuan, tujuan dan peranan evaluasi, evaluasi formatif dan sumatif, evaluasi intrinsik dan payoff, evaluasi amatir dan profesional, evaluasi bebas tujuan dan berdasarkan tujuan, dan evaluasi serta metaevaluasi.

6. Membicarakan Daftar Pemeriksaan Evaluasi Penting dengan mengenal dasar pemikiran yang mengarahkannya, mencatat dan menetapkan langkah-langkah, dan menerangkan bagaimana konsepsi formatif-sumatif berkaitan dengan Daftar Pemeriksaan Evaluasi Penting.
7. Menerangkan penilaiannya tentang ideologi-ideologi evaluasi yang diberi nama separatisme, positivisme, manajemen, dan relativisme.
8. Mengenal peranan dan kontribusi-kontribusi utamanya di dalam memprofesionalisasikan evaluasi.

Tujuan-tujuan dapat dicapai hanya dengan membaca materi yang ada di dalam unit ini. Untuk menambah pengetahuan anda tentang gagasan-gagasan Scriven, anda harus mempelajari tulisan-tulisannya yang orisinal, terutama tulisan-tulisan yang dicatat pada akhir unit ini. Sebagian dari materi yang dikemukakan di sini tidak mudah didapat pada literatur yang diumumkan dan menggambarkan satu hubungan kerja penulis yang erat dengan Dr. Scriven selama periode lebih dari 10 tahun.

### **Evaluasi yang Didefinisikan**

Selama bertahun-tahun, Dr. Scriven telah mengembangkan definisi yang diusulkannya tentang evaluasi, tetapi pesan intinya tetap sama. Di dalam sebuah tulisan klasik (1967) ia

mendefinisikan evaluasi sebagai suatu aktivitas metodologis yang “meliputi hanya mengumpulkan dan menggabungkan data prestasi dengan sekumpulan skala tujuan berbobot untuk menghasilkan penilaian-penilaian numeris atau komparatif, dan di dalam membenarkan (1) instrumen-instrumen pengumpulan data, (2) pembobotan, dan (3) pemilihan tujuan-tujuan”. Di dalam membahas tujuan dari makna evaluasi yang dimaksudkannya, kami sering mendengarnya mengatakan bahwa evaluasi pada intinya adalah penentuan sistematis dan obyektif tentang nilai atau manfaat dari suatu benda, dan bahwa evaluasi dapat dilakukan dengan sangat baik dengan cara mendorong penilaian yang independen untuk memberikan suatu keputusan tentang suatu obyek atau benda berdasarkan atas bukti yang dikumpulkan tentang bagaimana membandingkannya dengan obyek-obyek yang bersaing dalam rangka memenuhi kebutuhan-kebutuhan konsumen. Menurut pandangan ini, evaluasi lebih disukai yang bersifat komparatif; melalui implikasinya ia memandang kerugian-kerugian dan juga keuntungan-keuntungan komparatif; hal ini berhubungan dengan bagaimana sebaiknya dapat memenuhi kebutuhan-kebutuhan konsumen; secara optimal, evaluasi merupakan suatu aktivitas profesional yang memerlukan prosedur-prosedur sistematis; evaluasi harus dilakukan seobyektif mungkin; dan evaluasi harus mencapai

puncaknya pada keputusan-keputusan dan rekomendasi-rekomendasi.

### **Kritikan terhadap Persuasi-persuasi Lain**

Dengan memperhatikan definisi sebelumnya, Scriven mengkritik secara tajam pandangan-pandangan lain tentang evaluasi dan menggunakan analisis kritisnya untuk memperluas posisinya sendiri. Ia menuduh bahwa tradisi Tyler – yang memandang evaluasi sebagai yang menentukan apakah tujuan-tujuan telah tercapai atau belum – pada dasarnya adalah cacat, karena tradisi ini pada intinya bebas nilai. Ia memandang evaluasi berdasarkan atas pendekatan ini sebagai yang sangat tidak sah, karena tujuan-tujuan pengembangnya dapat menjadi tidak bermoral, tidak realistis, tidak mewakili kebutuhan-kebutuhan konsumen yang dinilai, atau terlalu sempit untuk membayangkan efek-efek samping yang mungkin sangat penting. Daripada menggunakan tujuan-tujuan untuk mempedomani dan menilai pengaruh-pengaruh yang mungkin ada, Scriven berpendapat bahwa para penilai harus menilai tujuan-tujuan dan tidak dipaksakan kepada mereka di dalam mencari hasil-hasil. Apakah program tersebut diarahkan oleh tujuan-tujuan yang bermanfaat, ia percaya, para penilai harus mencari semua hasil dari sebuah program, menilai kebutuhan-kebutuhan konsumen, dan menggunakan kedua

perangkat penilaian tersebut untuk sampai pada kesimpulan tentang manfaat dan nilai dari program-program yang ada.

Scriven juga mengambil isu dengan saran yang diberikan oleh Cronbach (1963). Cronbach telah mengkritik praktek umum mengevaluasi program-program dengan menggunakan percobaan-percobaan yang mengacu pada norma untuk membandingkan hasil dari kelompok-kelompok kontrol dan percobaan dan menganjurkan digunakannya pendekatan yang berorientasikan lebih banyak pada perkembangan. Cronbach menganjurkan terhadap penggunaan eksklusif model-model percobaan komparatif dan menyarankan agar berbagai macam ukuran sebaiknya digunakan untuk mempelajari suatu program khusus secara mendalam sambil dikembangkan dan agar hasil-hasilnya dapat digunakan untuk membantu mengarahkan perkembangan. Scriven mengatakan bahwa saran oleh Cronbach ini menutupi perbedaan antara tujuan dan peranan evaluasi dan sebenarnya menyamakan evaluasi dengan hanya salah satu peranannya, atau yang disebut oleh Scriven sebagai "evaluasi formatif". Berdasarkan atas kritikan ini, Scriven memperluas pandangannya tentang evaluasi di dalam tulisan klasiknya, "Metodologi Evaluasi" (1967).

## **Evaluasi Formatif dan Sumatif**

Di dalam tulisan ini, Scriven berpendapat bahwa tanggung jawab utama penilai adalah membuat keputusan-keputusan yang diberitahukan. Ia menekankan bahwa tujuan dari evaluasi adalah selalu sama – untuk mempertimbangkan nilai. Tetapi peranan dari evaluasi sangat berbeda-beda. Mereka dapat “menjadi bagian dari aktivitas pelatihan guru, proses perkembangan kurikulum, percobaan lapangan yang berhubungan dengan perbaikan teori pembelajaran, penyelidikan pendahuluan terhadap suatu keputusan tentang pembelian atau penolakan materi-materi . . .,” dan sebagainya. Ia menerangkan bahwa tidak mempunya untuk membedakan antara tujuan evaluasi (untuk mempertimbangkan nilai dari sesuatu) dan peranannya (penggunaan konstruktif data evaluatif) telah menyebabkan dilusi terhadap apa yang disebut evaluasi sehingga tidak lagi dapat mencapai tujuannya dari memperkirakan nilai. Dengan kata lain, ia mengatakan bahwa para penilai terlalu sering, di dalam mencoba untuk membantu para pendidik memperbaiki program-program mereka, menjadi suka memilih-milih dan tidak mampu menilai program-program tersebut. Bagi Scriven, suatu penilaian obyektif tentang nilai adalah sine qua non evaluasi.

Dengan semakin pentingnya tujuan evaluasi ditetapkan secara tegas, Scriven mulai menganalisa peranan-peranan

evaluasi. Ia menyimpulkan bahwa ada dua faktor utama: formatif, untuk membantu di dalam mengembangkan program-program dan obyek-obyek lain; dan sumatif, untuk memperkirakan nilai dari obyek apabila obyek ini telah dikembangkan dan dipasarkan.

Evaluasi formatif merupakan bagian integral dari proses perkembangan. Evaluasi ini memberikan umpan-balik terus-menerus untuk membantu di dalam merencanakan dan kemudian menghasilkan suatu obyek. Di dalam perkembangan kurikulum, evaluasi menghadapi pertanyaan-pertanyaan tentang validitas isi kurikulum, tingkat perbendaharaan kata, sifat kegunaan, ketepatan dari media, daya tahan materi, efisiensi, susunan pegawai, dan persoalan-persoalan lainnya. Pada umumnya, evaluasi formatif dilakukan untuk membantu staf memperbaiki apapun yang mereka kerjakan atau dikembangkan.

Di dalam peranan sumatif, evaluasi “dapat membantu untuk memungkinkan para penyelenggara memutuskan apakah kurikulum telah diselesaikan seluruhnya atau belum, disaring dengan menggunakan proses evaluasi di dalam peranan pertamanya (formatif), menunjukkan suatu kemajuan yang cukup berarti terhadap alternatif-alternatif yang ada untuk membenarkan biaya adopsi oleh sebuah sistem sekolah”. Evaluasi sumatif biasanya harus dilakukan oleh seorang penilai eksternal untuk meningkatkan obyektivitas, dan temuan-temuannya harus

diumumkan. Evaluasi semacam ini mencari semua pengaruh dari obyek dan mengujinya terhadap kebutuhan-kebutuhan konsumen relevan yang dinilai. Evaluasi ini membandingkan biaya dan pengaruh dari obyek tersebut dengan biaya dan pengaruh dari apa yang disebut oleh Scriven sebagai pesaing-pesaing kritis, terutama alternatif-alternatif yang lebih murah. Di dalam kasus audiensi yang dapat dipengaruhi untuk menilai hanya hasil-hasil terhadap tujuan-tujuan pengembang, evaluasi sumatif memberikan penilaian-penilaian tentang sejauh mana tujuan-tujuan tersebut menggambarkan dengan benar kebutuhan-kebutuhan yang dinilai. Secara keseluruhan, evaluasi sumatif membantu konsumen dengan memberikan penilaian independen kepada mereka yang membandingkan biaya, manfaat, dan nilai-nilai dari program atau produk yang bersaing.

### **Evaluasi Amatir versus Profesional**

Pada tahap-tahap awal perkembangan kurikulum, Scriven merujuk “evaluasi amatir” (evaluasi diri sendiri) kepada “evaluasi profesional”. Para pengembang, apabila mereka bertindak sebagai penilaian untuk mereka sendiri, dapat menjadi sedikit tidak sistematis dan subyektif, tetapi mereka juga dapat menjadi suportif, tidak mengancam, mengabdikan diri kepada menghasilkan suatu keberhasilan, dan toleran terhadap prosedur-

prosedur perkembangan penyelidikan dan tujuan-tujuan yang tidak jelas. Karena itu, mereka tidak mungkin akan melumpuhkan kreativitas pada tahap-tahap awal perkembangan. Para penilai profesional, jika terlibat terlalu awal, dapat “mengurangi semangat kreatif dari suatu kelompok yang produktif”, memperlambat proses perkembangan dengan mendesak tujuan-tujuan tersebut segera diklarifikasikan, atau kehilangan perspektifnya yang obyektif dengan menjadi terlalu dekat sejajar dengan usaha produksi, di antara pertimbangan-pertimbangan lainnya. Namun para penilai profesional dibutuhkan untuk melaksanakan evaluasi formatif dan sumatif selama tahap-tahap terakhir perkembangan. Kedua jenis evaluasi tersebut memerlukan keahlian-keahlian teknis yang tinggi dan obyektivitas yang jarang dimiliki oleh orang-orang pada staf pengembangan yang tidak pernah dilatih secara khusus dalam bidang teori dan metodologi evaluasi. Scriven menganjurkan agar seorang penilai profesional dimasukkan pada staf pengembangan untuk melakukan evaluasi formatif, dan ia sering menyarankan agar para penilai profesional eksternal ditugaskan untuk melakukan dan melaporkan evaluasi-evaluasi sumatif.

## **Evaluasi Intrinsik dan Payoff**

Scriven juga membedakan antara evaluasi intrinsik dan payoff. Evaluasi intrinsik menilai kualitas-kualitas dari suatu instrumentalitas, tanpa memperhatikan pengaruh-pengaruhnya terhadap klien, dengan menilai segi-segi seperti tujuan-tujuan, struktur, metodologi, kualifikasi dan sikap-sikap staf, fasilitas, kredibilitas umum, dan rekor masa lalu. Evaluasi payoff tidak berurusan dengan program, buku pelajaran, teori, atau obyek lainnya, tetapi sebaliknya dengan pengaruh-pengaruhnya terhadap klien. Pengaruh-pengaruh tersebut dapat meliputi skor-skor ujian, prestasi pekerjaan, atau status kesehatan. Scriven mengakui pentingnya evaluasi intrinsik, tetapi menekankan bahwa seseorang juga harus menentukan dan menilai hasil-hasilnya, karena jarang ada hubungan antara variabel-variabel hasil dengan proses jika pernah diketahui secara pasti. Ia menerangkan bahwa kedua jenis evaluasi tersebut dapat menyumbangkan baik kepada peranan formatif maupun peranan sumatif. Ia telah menjadi sangat kritis terhadap mengakreditasi dewan-dewan pengurus karena kesibukan mereka dengan kriteria intrinsik, seperti jumlah buku di dalam sebuah perpustakaan lembaga, dan kegagalan mereka untuk menilai prestasi para lulusan.

## **Evaluasi Bebas Tujuan**

Di dalam langkah lainnya terhadap kesibukan yang meluas dengan evaluasi berdasarkan tujuan, Scriven memperkenalkan sebuah proposal tandingan di bawah label “evaluasi bebas tujuan”. Menurut pendekatan ini penilai dengan sengaja terus-menerus mengabaikan tujuan-tujuan yang dicetak oleh sebuah program dan mencari semua pengaruh dari sebuah program terlepas dari tujuan-tujuan pengembangnya. Tidak ada efek samping yang dapat diteliti, karena data tentang semua pengaruh, apapun maksud dari program tersebut, adalah sama-sama dapat diterima. Jika sebuah program melaksanakan apa yang telah diusulkan untuk dilakukan, maka evaluasinya harus menegaskan hal ini, tetapi penilainya juga akan lebih memungkinkan untuk mengungkapkan akibat-akibat yang tak diharapkan yang akan luput dari perhatian para penilai berdasarkan tujuan karena kesibukan mereka dengan tujuan-tujuan yang telah ditetapkan. Di dalam setiap kasus, Scriven mengatakan bahwa evaluasi bebas tujuan adalah dapat menjadi terbalik dan komplementer: seseorang dapat memulai tujuan yang bebas untuk mencari semua akibat, kemudian beralih ke pendekatan berdasarkan tujuan untuk menjamin evaluasi akan menentukan apakah tujuan-tujuan telah dicapai atau belum; atau kedua jenis evaluasi tersebut dapat dilakukan secara serentak oleh penilai-penilai yang

berbeda. Keuntungan dari evaluasi bebas tujuan, menurut Scriven, adalah bahwa evaluasi ini kurang intrusif daripada evaluasi berdasarkan tujuan; lebih mampu beradaptasi dengan tujuan midstream yang berubah; lebih baik di dalam menemukan efek-efek samping; tidak cenderung kepada prasangka kognitif, persepsual, dan sosial; lebih menantang secara profesional; dan lebih pantas di dalam memperhitungkan berbagai macam nilai. Pada umumnya, kami memandang evaluasi bebas tujuan sebagai sebuah teknik inovatif – bukan sebuah model – yang sangat membantu di dalam melaksanakan pendekatan berorientasikan konsumen terhadap evaluasi.

### **Penilaian Kebutuhan**

Salah satu kesulitan di dalam menggunakan evaluasi bebas tujuan menyangkut bagaimana cara memberikan makna nilai kepada temuan-temuan. Jika hasil-hasil diketahui tanpa memandang apa yang sedang diusahakan oleh seseorang untuk dicapai, maka bagaimana seseorang dapat memisahkan konsekuensi yang diinginkan dari konsekuensi yang tidak diinginkan? Jawaban Scriven adalah bahwa seseorang harus membandingkan hasil-hasil yang diamati dengan kebutuhan-kebutuhan konsumen yang dinilai. Tetapi jika suatu kebutuhan merupakan kesenjangan antara sesuatu yang nyata dan sesuai

yang ideal dan jika sesuatu yang ideal adalah suatu tujuan, maka adakah kebutuhan-kebutuhan penilaian yang berdasarkan atas tujuan, dan karena itu tidakkah evaluasi-evaluasi bebas tujuan juga berdasarkan atas tujuan?

Bagi Scriven suatu kebutuhan adalah “sesuatu yang sangat penting untuk suatu cara kehidupan yang memuaskan, sesuatu tanpa cara kehidupan atau tingkat prestasi mana yang akan jatuh di bawah tingkat yang memuaskan”. Contoh-contoh yang ia gunakan secara khusus adalah vitamin C dan melek huruf fungsional. Dalam situasi tidak adanya hal-hal ini, seseorang akan menjadi sakit secara fisik atau lemah secara intelektual dan sosial; karena itu orang tersebut membutuhkan hal-hal tersebut. Bagi Scriven, kebutuhan-kebutuhan penilaian merupakan sebuah proses untuk menemukan fakta-fakta tentang apakah hal-hal yang jika tidak diberikan atau jika ditarik akan mengakibatkan konsekuensi-konsekuensi yang sangat buruk, dengan adanya standar-standar baik dan buruk yang tepat. Dengan adanya hasil-hasil dari kebutuhan penilaian tersebut, seseorang kemudian dapat menilai hasil-hasil yang diamati untuk suatu program tertentu sebagai yang baik, buruk, atau berbeda tergantung pada apakah mereka melakukannya atau akan menyumbang kepada mememnuhi kebutuhan-kebutuhan yang diidentifikasi.

## **Daftar Pemeriksaan Evaluasi Penting**

Daftar Pemeriksaan Evaluasi Penting ini mensintesa gagasan-gagasan Scriven sebelumnya. Ini menggambarkan pandangannya bahwa evaluasi melibatkan banyak dimensi, harus menggunakan banyak perspektif, melibatkan banyak tingkat ukuran, dan harus menggunakan banyak metode. Karena itu, ia kadang-kadang menyebut Daftar Pemeriksaan Evaluasi Penting sebagai multimodel evaluasi. Delapan belas pokok pemeriksaannya adalah sedbagai berikut :

1. *Deskripsi*. Apa yang akan dievaluasi? Yaitu *evaluand*, dijelaskan seobyektif mungkin. Apakah ia memiliki komponen-komponen? Apa saja hubungan-hubungannya?

2. *Klien*. Siapa yang menugaskan evaluasi tersebut? Yaitu *klien* untuk evaluasi tersebut, yang mungkin sebagai atau mungkin bukan sebagai *inisiator* permintaan terhadap evaluasi tersebut dan mungkin sebagai atau mungkin bukan sebagai *instigator* *evaluand*.

3. *Latar belakang dan Konteks* (a) *evaluand* dan (b) evaluasi. Meliputi identifikasi stakeholder (seperti nonklien yang dicatat pada nomor 2, sponsor, wakil masyarakat, dsb.).

4. *Sumber-sumber* (“mendukung sistem” atau “memperkuat penilaian”) (a) tersedia untuk atau bagi penggunaan evaluand, (b) tersedia untuk atau bagi penggunaan penilai.

5. *Fungsi*. Apa yang dilakukan oleh evaluand? Membedakan apa yang *diusulkannya* untuk dilakukan – fungsi atau peranan yang *diharapkan atau diduga* – dari apa hal itu sebenarnya dilakukan – fungsi-fungsi *aktual* untuk klien dan konsumen.

6. *System Pengiriman*. Bagaimana evaluand mencapai pasar? Bagaimana ia dipertahankan (dilayani)? Bagaimana memperbaikinya (memperbaharainya)? Bagaimana para pengguna dilatih?

7. *Konsumen*. Siapa yang menggunakan atau menerima (pengaruh-pengaruh dari) evaluand? Membedakan populasi konsumen yang menjadi target – pasar yang diharapkan – dari populasi konsumen yang terkena dampak langsung secara aktual dan potensial – pasar atau pelanggan atau penerima yang sesungguhnya (atau klien untuk evaluand, sering disebut klien).

8. *Kebutuhan dan Nilai-nilai* dari populasi yang terkena dampak secara potensial. Hal ini akan meliputi keinginan serta kebutuhan, dan juga nilai-nilai seperti standar-standar manfaat dan ideal yang dinilai atau dipercaya; tujuan-tujuan program yang telah ditetapkan di mana evaluasi berdasarkan tujuan dilakukan; dan kebutuhan-kebutuhan.

9. *Standar*. Adakah standar-standar manfaat atau nilai yang diterapkan berlaku secara obyektif? Dapatkah standar-standar tersebut disimpulkan dari klien plus konsumen, fungsi dan kebutuhan-kebutuhan/nilai-nilai?

10. *Proses*. Batas-batas/kerugian/keuntungan apa yang berlaku untuk operasi normal evaluand (bukan untuk pengaruh-pengaruhnya atau hasil-hasil dari nomor 11)? Satu pertimbangan proses ilmiah akan menggunakan indikator-indikator proses hasil akhir yang sah secara ilmiah; pertimbangan proses lainnya akan menggunakan materi yang baik secara ilmiah (secara historis, dsb.) di dalam buku pelajaran atau kursus.

11. *Hasil-hasil*. Pengaruh-pengaruh apa yang dihasilkan evaluand (diharapkan atau tidak diharapkan)? Sebuah matriks pengaruh-pengaruh sangat berguna; populasi yang terpengaruh x jenis pengaruh (kognitif/afektif/psikomotor/kesehatan/sosial/lingkungan) x ukuran masing-masing waktu permulaan z (segera/akhir “perlakuan”/kemudian) x lamanya waktu x masing-masing komponen atau dimensi (jika evaluasi analitis diperlukan).

12. *Generalisabilitas* untuk orang/tempat/waktu/versi lain. (“Orang” berarti staf dan juga penerima). Ini dapat disebut deliverabilitas dan salabilitas/eksportabilitas/durabilitas/modifiabilitas.

13. *Biaya*. Dolar vs. psikologis vs. personil; pertama vs. berulang (termasuk persiapan-pemeliharaan-perbaikan); langsung/tak langsung vs. cepat/terlambat/tertunda; oleh komponen-komponen jika tepat.

14. *Perbandingan* dengan pilihan-pilihan alternatif – meliputi pilihan-pilihan yang diakui dan tidak diakui, pilihan-pilihan yang ada saat ini dan pilihan-pilihan yang dapat dibuat. Pesaing-pesaing terkemuka dalam bidang ini adalah “pesaing-pesaing kritis” dan dikenal atas dasar keefektifan biaya plus.

15. *Arti penting*. Suatu sintesis tentang semua hal yang telah dicatat di atas. Validasi dari prosedur mensintesis sering menjadi salah satu tugas paling sulit di dalam evaluasi. Hal ini biasanya tidak dapat diserahkan kepada klien, yang biasanya kurang berpengalaman atau tidak dilengkapi dengan obyektivitas untuk melakukan hal itu; dan pendekatan-pendekatan formula – misalnya, perhitungan keuntungan dan kerugian.

16. *Rekomendasi*. Rekomendasi dapat atau tidak dapat diminta, dan dapat atau tidak dapat diikuti dari evaluasi; meskipun diminta rekomendasi ini mungkin tidak dapat dilaksanakan untuk memberikan sesuatu karena hanya rekomendasi-rekomendasi yang akan menjadi tepat yang tidak diminta sehingga setiap bukti ilmiah untuk rekomendasi-

rekomendasi khusus dapat diperoleh di dalam bidang penelitian yang relevan.

*17. Laporan.* Perbendaharaan kata, panjangnya, format, alat, waktu, lokasi, dan personil untuk penyajiannya memerlukan penelitian yang cermat, seperti yang dilakukan pada perlindungan/kebebasan pribadi/publisitas dan penyaringan sebelumnya atau sirkulasi draf-draf pertamadan terakhir.

*18. Metaevaluasi.* Evaluasi harus dinilai, lebih baik sebelum sampai pada (a) pelaksanaannya, (b) penyebaran terakhir laporan. Evaluasi eksternal sangat diinginkan, tetapi sebelumnya penilai utama harus menggunakan Daftar Pemeriksaan Evaluasi Penting untuk evaluasi itu sendiri.

Langkah-langkah ini tidak dimaksudkan untuk dilaksanakan di dalam setiap rangkaian khusus, tetapi semuanya harus dihadapi sebelum Daftar Pemeriksaan Evaluasi Penting dilaksanakan secara tepat. Juga, seorang penilai dapat melakukan putaran melalui daftar pemeriksaan beberapa kali selama evaluasi sebuah program. Putaran-putaran pertama adalah evaluasi formatif; putaran terakhir adalah apa yang disebut oleh Scriven "evaluasi sumatif".

Dasar pemikiran tentang Daftar Pemeriksaan Evaluasi Penting adalah bahwa evaluasi pada dasarnya merupakan sebuah proses pengurangan data dengan jalan mana data dalam jumlah

besar diperoleh dan dinilai dan kemudian disintesa menjadi suatu pertimbangan yang menyeluruh tentang nilai. Di dalam menjelaskan proses pengurangan data ini, Scriven menjelaskan bahwa langkah-langkah awal membantu menggolongkan sebuah program atau produk dan langkah-langkah terakhir membantu menilai validitasnya.

### **Metaevaluasi**

Soal terakhir di dalam Daftar Pemeriksaan Evaluasi Penting memerlukan penilaian tentang evaluasi. Scriven memperkenalkan konsep ini pada tahun 1968, ketika ia mengumumkan sebuah tulisan yang menanggapi pertanyaan-pertanyaan tentang bagaimana menilai instrumen-instrumen evaluasi. Ia menyebutnya sebagai salah satu dari banyak perhatian pada metaveluasi dan menekankan bahwa para penilai mempunyai kewajiban profesional untuk menjamin bahwa evaluasi-evaluasi yang mereka selesaikan atau diusulkan sesuai dengan evaluasi kompeten. Dasar pemikirannya adalah bahwa “evaluasi merupakan suatu subyek acuan sendiri yang khusus karena ... evaluasi berlaku untuk proses dan produk-produk dari semua usaha keras manusia yang serius dan juga untuk evaluasi”. Ia menjelaskan bahwa metaevaluasi dapat menjadi formatif, di dalam membantu penilai untuk merencanakan dan melakukan suatu evaluasi yang

baik, atau menjadi sumatif, di dalam memberikan bukti independen kepada klien tentang kemampuan teknis dari penilai utama dan kekuatan dari evaluasinya atau laporan-laporannya.

### **Ideologi Evaluasi**

Seperti yang telah diketahui di dalam pembahasan sebelumnya, Scriven telah menjadi salah seorang kritikus yang sangat vokal dan bijaksana tentang pandangan-pandangan umum evaluasi. Konsisten dengan pendirian kritis ini, ia telah menegaskan bahwa evaluasi merupakan suatu subyek yang mengacu sendiri secara khusus, yang sesuai dengan dukungannya terhadap metaevaluasi. Di dalam sebuah paper yang baru ia menggolongkan pandangan-pandangan umum tentang evaluasi kedalam empat kelompok dan mengupas masing-masing kelompok secara luas dengan harapan akan meyakinkan para penilai untuk mengakui dan melepaskan prasangka-prasangka tertentu, yang dikatakannya telah melemahkan pekerjaan atau karya evaluasi. Kemudian ia menggunakan analisisnya tentang kekuatan dan kelemahan dari masing-masing pendekatan untuk memperkuat dasar pemikirannya bagi Daftar Pemeriksaan Evaluasi Penting. Ia menjelaskan model daftar pemeriksaan ini sebagai model yang meliputi segi-segi terbaik dari semua proposal serius lainnya tentang bagaimana melakukan evaluasi dan model yang

menghindari cacat-cacat yang telah ia perkenalkan di dalam proposal-proposal lainnya. Kami menganggap hal ini akan membantu anda memperoleh wawasan lebih jauh tentang filsafat evaluasi Scriven pada umumnya, dan Daftar Pemeriksaan Evaluasi Penting pada khususnya, jika anda memperhatikan secara cermat analisisnya tentang ideologi-ideologi alternatif. Oleh karena itu, kami telah mencoba di bawah ini untuk menangkap pokok-pokoknya yang paling penting tentang masing-masing dari empat ideologi tersebut. Ideologi-ideologi ini adalah ideologi separatis, ideologi positivis, ideologi manajerial, dan ideologi relativis.

### **Ideologi Separatis**

Scriven memandang ideologi separatis sebagai yang bersumber pada penyangkalan atau penolakan proposisi bahwa evaluasi merupakan suatu aktivitas self-referent. Ideologi ini digambarkan dengan sangat baik di dalam proposal-proposal evaluasi yang memerlukan ditugaskannya penilai-penilai yang sepenuhnya tidak tergantung pada apa yang akan dievaluasi. Menetapkan dan mempertahankan kebebasan penilai ini dari evaluand sering dianggap penting untuk menjamin bahwa laporan-laporan evaluasi tidak menyimpang. Di samping itu, para penilai yang mempraktekkan ideologinya, menurut Scriven sering tidak mengakui atau tidak mau menghadapi keharusan pekerjaan

mereka sendiri dievaluasi. Yang paling mungkin, banyak di antara mereka memandang metaevaluasi tersebut sebagai suatu urusan orang lain, karena para penilai, menurut pandangan separatis mereka, tidak dapat menjadi obyektif di dalam mengevaluasi pekerjaan mereka sendiri. Karena itu, Scriven menunjuk kepada paradoks seorang penilai yang memperoleh suatu kehidupan dengan mengevaluasi pekerjaan orang lain tetapi tidak mampu mencoba, atau bahkan menolak, evaluasi-evaluasi pelayanannya sendiri.

Sebagai yang bertentangan dengan posisi separatis, Scriven berpendapat bahwa para profesional, termasuk penilai-penilai profesional, harus mengakui dan menghadapi secara langsung sifat evaluasi self-referent. Hal yang sangat menentukan dari seorang profesional adalah membiarkan pekerjaan seseorang untuk evaluasi. Fakta bahwa semua evaluasi cenderung akan menyimpang tidak harus mencegah seseorang dari mengevaluasi pekerjaan orang itu sendiri atau menugaskan orang lain untuk melakukan pekerjaan tersebut. Sebaliknya, seseorang harus menanggapi dengan melakukan evaluasi dengan cara sebisa mungkin tidak menyimpang dan membiarkan evaluasi tersebut kepada penelitian yang cermat sesuai dengan standar-standar evaluasi yang baik yang diakui. Di dalam apa yang disebut evaluasi-evaluasi program seseorang harus melihat secara realistis

kepada staf dan juga aspek-aspek lainnya dari program tersebut karena keberhasilan dan kegagalan tidak dapat dipisahkan dari pekerjaan staf tersebut.

### ***Ideologi Positivis***

Scriven memandang ideologi yang kedua, yaitu positivisme logis, sebagai reaksi berlebihan lainnya kepada fobia nilai. Di dalam usaha mereka untuk menghilangkan prasangka dari karya-karya ilmiah, Scriven memandang positivis yang bereaksi kepada tujuan mencoba untuk memberikan ilmu pengetahuan abad dua puluh pada umumnya dan evaluasi pada khususnya sebagai yang bebas nilai. Walaupun para separatis menolak sifat self-referent dari ilmu pengetahuan atau evaluasi, namun para positivis menolak sifat evaluatif dari ilmu pengetahuan tersebut. Scriven menunjuk kepada sejumlah kasus yang bertentangan, misalnya, para ahli psikologi pendidikan yang menegaskan bahwa tidak ada pertimbangan-pertimbangan evaluatif yang dapat dilakukan dengan obyektivitas, namun dengan mudah menghasilkan pertimbangan-pertimbangan evaluatif tentang prestasi dari siswa-siswa mereka. Tentu saja, reaksi Scriven terhadap cacat-cacat dari positivisme adalah memberikan arti yang sangat penting kepada praktek memberikan makna-makna nilai kepada temuan-temuan yang diperoleh di dalam studi-studi evaluasi.

### ***Ideologi Manajerial***

Bagi Michael Scriven, “evaluasi yang dilaksanakan dengan baik” sering berarti jauh dari sekedar evaluasi yang diarahkan oleh seorang penyelenggara yang kompeten. Bahkan hal ini dapat meliputi “suatu kegemeran meyalani diri sendiri di dalam fobia nilai” oleh manajer dan penilai program. Manajer program dapat mengadakan pengawasan yang ketat terhadap evaluasi yang mereka perintahkan sehingga tidak akan ada sesuatu yang mengejutkan. Mereka mungkin hanya menginginkan program mereka dievaluasi, bukan personil yang menjalankan program tersebut dan terutama bukan penyelenggaranya. Dan mereka dapat memastikan bahwa evaluasi terbatas pada menentukan apakah tujuan-tujuan yang telah mereka tetapkan untuk program tersebut dapat dicapai atau tidak dan bahwa hal itu dibatasi dari menilai pekerjaan mereka berdasarkan atas keinginan orang lain untuk program tersebut.

Dari perspektif manajer ideologi manajerial ini jelas meliputi suatu prasangka terhadap menghasilkan laporan-laporan yang menguntungkan. Menurut Scriven, banyak penilai bersedia memenuhi keinginan manajer untuk laporan-laporan yang menguntungkan dan dapat diprediksi karena sejumlah alasan melayani diri sendiri yang sama. Mereka menginginkan kontrak-kontrak di masa mendatang atau mempertahankan posisi evaluasi

mereka di dalam lembaga tersebut, dan memberikan laporan yang menguntungkan, atau setidaknya-tidaknya laporan yang tidak membuat klien dan sponsor mereka menjadi gelisah, adalah demi kepentingan terbaik memperoleh pekerjaan di masa mendatang.

Bagi Scriven, ketaatan yang dipandang secara luas kepada ideologi manajerial dan hubungan-hubungannya dengan praktek-praktek evaluasi yang buruk lainnya merupakan suatu ejekan untuk masyarakat dan untuk profesi evaluasi. Ia telah menggunakan analisis kritisnya tentang pendirian ini sebagai suatu platform dari mana ia mendukung serangkaian pembaharuan, yang dapat dilihat di dalam Daftar Pemeriksaan Evaluasi Penting-nya, yang meliputi :

1. Melakukan penilaian kebutuhan-kebutuhan sebagai suatu dasar untuk mempertimbangkan apakah sebuah program telah memberikan hasil-hasil yang menguntungkan atau tidak.
2. Mengevaluasi “bebas tujuan” bukan untuk menjaidi disibukkan dengan tujuan-tujuan pengembang dan karena tidak dapat memperoleh haisil-hasil penting tetapi tidak diharapkan, baik dan buruk.
3. Membandingkan apa yang sedang dievaluasi dengan alternatif-alternatif yang nyata.
4. Menggabungkan evaluasi personil dan program.

### ***Ideologi Relativis***

Ideologi lainnya yang telah dipandang oleh Scriven sebagai ideologi cacat dan lemah di dalam pengaruhnya terhadap pekerjaan atau karya evaluasi adalah ideologi relativis. Scriven memandang ideologi ini sebagai suatu reaksi yang berlebihan terhadap masalah-masalah yang berhubungan dengan ideologi positivis. Sedangkan para positivis sering mengemukakan pandangan bahwa ada suatu kenyataan yang obyektif yang dapat diketahui oleh seseorang yang dapat dan akan menggunakan prosedur-prosedur penilaian tidak menyimpang, para relativis telah menegaskan bahwa konstruksi ini terlalu sederhana dan hanya dapat menghasilkan penilaian-penilaian sempit yang memberikan keunggulan eksklusif dan tak semestinya kepada perspektif dari suatu kelompok yang berkuasa, dengan pandangan yang salah bahwa perspektif dan penilaian mereka adalah obyektif.

Scriven memandang banyak gerakan di dalam bidang evaluasi ini sebagai suatu reaksi yang berlebihan yang kadang-kadang menyangkal kemungkinan determinasi-determinasi obyektif mengenai gambaran-gambaran yang berharga atau bahkan benar secara obyektif tentang program-program. Walaupun ia juga menolak keberadaan dari suatu gambaran yang benar, namun ia menyarankan kepada kita supaya tidak meninggalkan gagasan bahwa ada suatu kenyataan yang obyektif.

Malahan ia menyarankan agar kita harus dapat merelativisasi gambaran-gambaran kita kepada kelompok-kelompok orang yang berbeda.

### **Profesionalisasi Evaluasi**

Konsisten dengan perhatiannya untuk mengevaluasi gagasan-gagasan dan karya-karya dari para penilai sesamanya, Scriven telah menjadi suatu kekuatan utama untuk memprofesionalisasikan evaluasi. Ia telah menjadi seorang pendukung terkemuka perkembangan dan penggunaan *Standar-standar untuk Penilaian Evaluasi Pendidikan* (Panitia Bersama, 1981). Ia membantu menetapkan dan mengembangkan Jaringan Evaluasi, sebuah organisasi profesional bagi para penilai dari program-program pendidikan, kesehatan, pemerintahan, dan sosial. Dan ia mengembangkan surat kabar organisasi ini menjadi majalah yang sangat dihormati, *Evaluation News*.

Tulisan ini telah menjadi sebuah ringkasan singkat tentang beberapa kontribusi penting dari Michael Scriven kepada evaluasi. Kontribusi-kontribusi ini telah menyebar luas selama dua dekade, dan telah mempengaruhi praktek evaluasi secara kuat. Para mahasiswa evaluasi yang sering dapat mempelajari tulisan-tulisan orisinalnya dan mengikuti kontribusinya di masa mendatang.

## **Pertanyaan-pertanyaan Tanpa Jawaban**

Di dalam ujian pengetahuan ini anda harus berusaha keras dan diberikan sedikit bantuan untuk menilai jawaban-jawaban anda. Pada bagian ini akan akan mengandalkan terutama pada kemampuan anda sendiri. Bagian ini memberikan beberapa pertanyaan, tetapi jawabannya tidak diberikan. Pertanyaan-pertanyaan semuanya penting. Anda akan menemukan bahwa sebagian dari pertanyaan-pertanyaan tersebut menunjukkan kontroversial. Karena itu, di dalam ujian ini, sebagai seorang penilaian tipe Scriven yang sesungguhnya, anda harus dapat merumuskan penilaian anda sendiri, mengembangkan dan mempersiapkan diri untuk mempertahankan jawaban-jawaban anda atas pertanyaan-pertanyaan tersebut.

Kami meminta kepada anda untuk memperhatikan dan memberikan tanggapan kepada pertanyaan-pertanyaan berikut. Tuliskan jawaban-jawaban anda sehingga anda akan siap menanggapi pertanyaan-pertanyaan tersebut dan membicarakannya bersama dengan orang lain.

1. Apakah semua evaluasi bersifat komparatif? Mengapa atau mengapa tidak?
2. Apakah evaluasi formatif dan evaluasi sumatif merupakan konsep-konsep yang berbeda menurut pemahaman dan operasionalnya?

3. Jika seorang penilai mempertimbangkan suatu program di satu tempat dan pada waktu yang tepat, apakah ia kehilangan kebebasan dan obyektivitasnya tentang evaluasi programnya di masa depan? Mengapa atau mengapa tidak? Jika ya, haruskah seorang penilai menghentikan hubungannya dengan program tersebut?
4. Gambarkan sebuah program pelatihan yang diselenggarakan pada musim panas untuk guru-guru tingkat kedua yang bekerja di sekolah-sekolah dengan konsentrasi yang tinggi. Menurut konsepsualisasi evaluasi Scriven, bagaimana anda akan mengevaluasi program tersebut?
5. Sekarang, menilai suatu program, proyek, atau produk yang anda anggap tepat, menurut pendekatan terhadap evaluasi dari Scriven. Jelaskan apa yang akan dievaluasi. Jelaskan secara singkat model anda untuk mengevaluasi hal itu. Kemudian periksa hasil-hasil evaluasi anda.
6. Sebagai tugas anda yang terakhir, lakukan evaluasi unit ini menurut pendekatan terhadap evaluasi dari Scriven.