

POHON TEORI EVALUASI KEBIJAKAN DAN PROGRAM (Metode, Nilai dan Menilai, Penggunaan)

Mami Hajaroh

Program Studi Kebijakan Pendidikan, Universitas Negeri Yogyakarta

mamihajaroh@uny.ac.id

Abstrak

Tulisan ini mendeskripsikan filosofi keilmuan evaluasi kebijakan dan program. Evaluasi sebagai bagian dari proses kebijakan merupakan tahap penting, yakni sebagai tahap penyediaan informasi tentang kebijakan dan program yang berjalan. Dalam prakteknya, evaluasi memerlukan teori-teori evaluasi. Teori yang dimaksudkan oleh para theorist identik dengan pendekatan atau model teori yang bersifat preskriptif, bukan tentang teori deskriptif atau empiris. Pohon teori evaluasi sebagai ilmu terapan dibangun oleh tiga akar ilmu yakni *epistemology*, *social inquiry* dan *social accountability* yang membangun tiga cabang ilmu evaluasi. Ketiga cabang ilmu evaluasi meliputi cabang metode (*method*), cabang menilai (*valuing*), dan cabang penggunaan (*use*). Ketiga cabang pohon ilmu evaluasi dilihat sebagai sesuatu yang saling berhubungan satu sama lain. Bila pohon ilmu evaluasi itu berada dalam ruang tiga dimensi maka ujung kanan adalah cabang penilaian akan berhubungan dengan sisi paling kiri yakni cabang penggunaan. Para ahli ilmu evaluasi ditempatkan pada setiap cabang dengan melihat kecenderungan utama dan khusus dari pendekatan ilmu yang mereka kembangkan.

Kata Kunci: Evaluasi, Pohon Teori, Pendekatan Evaluasi, Evaluasi Kebijakan

Abstract

This paper describes the scientific philosophy of evaluating policies and programs. Evaluation as part of the policy process is an important stage, namely the stage of providing information about the implementation of policies and programs. Practically, evaluation requires evaluation theories. The theory intended by theorists is identical to prescriptive approaches or theoretical models, it is not about descriptive or empirical theories. The tree of evaluation theory as applied science is built by three scientific roots, namely epistemology, social inquiry, and social accountability that build three branches of evaluation science. They are branch method, branch valuing, and branch usage. Those branches are interconnected one to another. If the evaluation science tree is in a three-dimensional space, the right end evaluation branch will relate to the far left side, namely the branch of usage. Evaluation experts are placed on each branch by looking at the main and specific trends of the science approach they develop.

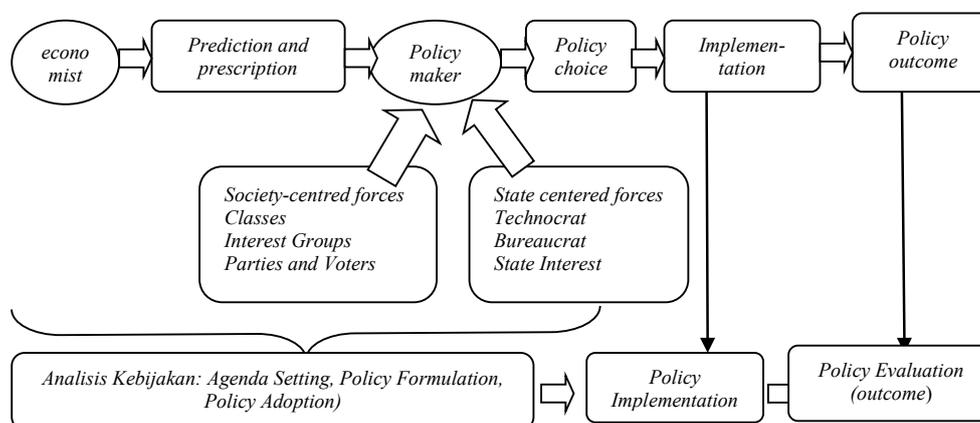
Keywords: Evaluation, Tree Theory, Evaluation Approach, Policy Evaluation

PENDAHULUAN

Evaluasi merupakan salah satu rangkaian kegiatan dalam proses implementasi kebijakan dan program pendidikan. Melalui evaluasi akan diperoleh informasi tentang apa yang telah dicapai dan mana yang belum tercapai dari sebuah program. Informasi hasil evaluasi selanjutnya digunakan sebagai bahan perbaikan program tersebut. Dengan kata lain evaluasi akan akan diketahui efektifitas, efisiensi, kualitas, kinerja, atau produktivitas suatu lembaga dalam melaksanakan programnya sekaligus untuk meningkatkannya. Arti penting dari pelaksanaan kegiatan evaluasi adalah untuk melihat kembali apakah suatu program dan kebijakan pendidikan itu telah dapat dilaksanakan secara efektif sesuai dengan perencanaan dan mencapai hasil sesuai yang diharapkan. Dari hasil kegiatan evaluasi akan dapat diketahui hal-hal yang telah dicapai dan apakah suatu program dapat memenuhi kriteria yang telah ditentukan. Berdasarkan hasil evaluasi kemudian diambil keputusan apakah program dan kebijakan tersebut akan: (1) diteruskan, (2) direvisi, (3) dihentikan, atau (4) dirumuskan kembali sehingga dapat ditemukan tujuan, sasaran dan alternatif baru yang sama sekali berbeda dengan sebelumnya.

Evaluasi kebijakan merupakan akhir dari proses yang dimulai dari merumuskan, menguji, mengadopsi kebijakan, implementasi dan evaluasi. Evaluasi merupakan cara untuk mengetahui *outcome* kebijakan dan program. Formulasi Proses kebijakan dalam model linier digambarkan sebagai berikut:

Gambar 1. Model Linier Formulasi Kebijakan (Grindle, M. and Thomas, J., 1990) dikombinasi dengan Proses Kebijakan.



Dalam pembaharuan kebijakan dan program pendidikan, kegiatan evaluasi terhadap program yang sedang maupun telah berjalan sebelumnya perlu dilakukan dengan baik. Untuk dapat menyusun program yang lebih baik, maka hasil evaluasi program sebelumnya merupakan acuan yang tidak dapat ditinggalkan.

Worthen dan Sanders (1973: 19) menyatakan bahwa: *Evaluation is the determination of the worth of thing. It includes obtaining information for use in judging the worth of a program, product, procedure, or objective, or the potential utility of alternative approaches designed to attain specified*

objectives. Evaluasi diartikan sebagai penentuan nilai (*worth*) terhadap sesuatu hal, yang meliputi aktifitas pengumpulan informasi yang digunakan untuk menentukan nilai keberhasilan suatu program, produk, prosedur, tujuan atau manfaat potensi pada desain alternatif pendekatan, untuk mempertahankan pendekatan yang khusus. Pendapat tersebut mengimplikasikan adanya kriteria tertentu yang digunakan untuk menentukan nilai atau harga (*worth*) sesuatu.

Untuk dapat melakukan evaluasi kebijakan dan program para evaluator memerlukan keilmuan evaluasi karena evaluasi merupakan satu proses akademis ilmiah yang semestinya mengikuti proses ilmiah. Para ilmuwan evaluasi mengembangkan teori dalam arti metode dan pendekatan dalam evaluasi sebagai dasar untuk melakukan evaluasi. Para teoris evaluasi ini diklasifikasikan ke dalam tiga cabang yang tumbuh dari pohon evaluasi.

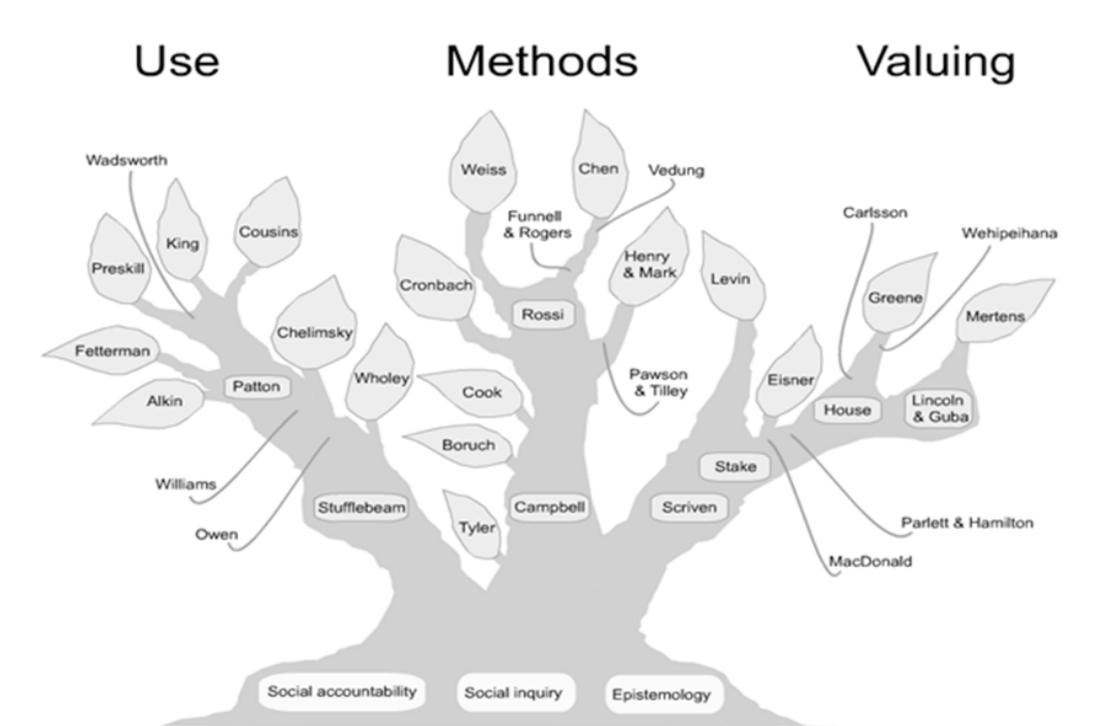
PEMBAHASAN

Pohon Teori Evaluasi

Alkin (2012:102) menempatkan evaluasi ke dalam Pohon Teori Evaluasi yang terdiri dari tiga cabang utama: metode, penilaian dan penggunaan. Teori yang dimaksudkan oleh para teoris evaluasi identik dengan pendekatan atau model bukan tentang teori deskriptif atau empiris yang menggambarkan apa yang akan terjadi dalam serangkaian tindakan tertentu. Menurut Alkin kompleksitas konteks kebijakan, program, dan tindakan evaluator yang sangat luas tidak akan pernah bisa membuat teori evaluasi deskriptif. Padahal di dalam penelitian evaluasi memerlukan teori-teori, oleh karena itu sebagai konsekuensi dari kekurangan penelitian evaluasi maka dikembangkan panduan tentang bagaimana melakukan evaluasi. Beberapa penulis evaluasi terkemuka menggambarkan mengenai bagaimana mereka meyakini bahwa evaluasi harus diperlakukan sebagai teori preskripsi. Preskripsi tentang apa yang diharuskan, ketentuan, petunjuk, dan peraturan yang harus dilakukan dalam evaluasi.

Tiga akar keilmuan yakni *epistemologi*, *social inquiry* dan *social accountability* membangun pohon teori evaluasi dengan tiga cabang. Ketiga cabang dalam teori evaluasi yakni penilaian (*valuing*), metode (*method*) dan Penggunaan (*use*). Pohon teori evaluasi dapat dilihat pada gambar dua cabang utama dari Pohon Teori Evaluasi adalah cabang metode karena evaluasi harus dipandu oleh metodologi penelitian. Menurut Alkin dan Christie (2004:17) cabang metode dari pohon teori evaluasi berkaitan dengan evaluasi sebagai riset, dan ini dikuasai oleh konstruksi pengetahuan. Tujuan dari konstruksi pengetahuan adalah meningkatkan pemahaman terhadap permasalahan publik dan menetapkan faktor-faktor, rencana solusi, dan teori implementasi. Ahli terkemuka teori dalam cabang metode, sekaligus sebagian dari mereka adalah peneliti metode evaluasi menurut Alkin dan Christie (2004: 17-32) adalah Campbell dan Stanley (1963), Cook dan Campbell (1979), Cronbach et al. (1981), Chen dan Rossi (1987), Chen (1990), Shadish, Cook dan Leviton (1991), Rossi dan Freeman (1993), dan Weiss (1998).

Gambar 2. Pohon Teori Evaluasi Alkin, M. C. (2012)



Alkin (2012: 104) menjelaskan di sebelah kanan cabang metode adalah cabang menilai (*Valuing*), yang awalnya terinspirasi oleh karya Michael Scriven. Scriven menyatakan bahwa bukanlah evaluasi bila tanpa penilaian (1967). Scriven berpendapat bahwa pekerjaan evaluasi adalah membuat nilai tentang objek yang sedang diteliti. Cabang ini terbelah dua, objektivis dan subjektivis, yang membedakan dua perspektif mendasar mengenai proses evaluasi. Objektivis dipengaruhi sub-cabang kiri yang banyak dipengaruhi oleh pandangan Scriven. Scriven berpendapat bahwa peran evaluator adalah melakukan penilaian, artinya peran penting evaluator adalah menilai. Dalam sub-cabang subyektivis ke kanan, para pendukung berpendapat bahwa realitas adalah proses yang terus berlangsung dan dinamis, dan kebenaran selalu relatif terhadap kerangka acuan tertentu. Dengan demikian, penilaian harus dilakukan dalam konteks memahami “makna subjektif” dari informasi evaluasi. Ahli terkemuka teori dalam cabang metode ini menurut Alkin dan Christie (2004) adalah Campbell dan Stanley (1963), Cook dan Campbell (1979), Cronbach et al. (1981), Chen dan Rossi (1987), Chen (1990), Shadish, Cook dan Leviton (1991), Rossi dan Freeman (1993), dan Weiss (1998).

Selanjutnya (Alkin, 2012: 104) mendeskripsikan bahwa cabang ketiga dari pohon teori evaluasi adalah penggunaan (*use*). Cabang yang ketiga ini fokus pada cara dan dimana informasi evaluasi akan digunakan, dan siapa yang akan menggunakan informasi. Alkin dan Christie (2004: 14) menuliskan bahwa Stufflebeam merupakan pionir pada cabang ini dan peneliti-peneliti lainnya adalah Patton (1997), Owen dan Rogers (1999), Alkin (2004), dan Shadish et al. (1991) yang mengidentifikasi langkah-langkah aktif yang diperlukan untuk peningkatan penggunaan temuan evaluasi.

Ketiga cabang pohon tidak dimaksudkan untuk dilihat sebagai cabang-cabang independen dari satu sama lain, tetapi ketiganya mencerminkan hubungan satu sama lain. Jika pohon teori itu berada dalam ruang tiga dimensi maka ujung kanan dari cabang penilaian akan berhubungan dengan sisi paling kiri yakni cabang penggunaan. Para ahli teori diposisikan pada setiap cabang tidak hanya mencerminkan penekanan pandangan mereka, tetapi juga kecenderungan utama dari pendekatan khusus yang mereka kembangkan. Sifat relasional dari ketiga cabang ini membantu dalam menggambarkan pendekatan atau model yang dikembangkan oleh para ahli teori.

Paradigma keilmuan (Merten and Wilson (2012: 56) yang digunakan dalam pohon teori evaluasi adalah post positivist pada cabang metode, pragmatic pada cabang penggunaan dan konstruktifis pada cabang menilai. Cabang metode evaluasi fokus desain dan data kuantitatif. Cabang penggunaan evaluasi fokus pada data yang ditemukan bermanfaat bagi stakeholder dan diadvokasi untuk menggunakan mix-method. Sedangkan cabang menilai menggunakan paradigma konstruktifis yang fokus terutama pada indentifikasi nilai-nilai dan perspektif yang *multiple* melalui metode kualitatif. Ketiga cabang teori evaluasi dapat dilihat tabel berikut:

Tabel 1. Paradigma Keilmuan Cabang Evaluasi

<i>Brach</i>	<i>Paradigm</i>	<i>Description</i>
<i>Method</i>	<i>Postpositivist</i>	<i>Focuses primarily on quantitative design and data</i>
<i>Use</i>	<i>Pragmatic</i>	<i>Focuses primarily on data that are found to be useful by stakeholders; advocated for the use of mix method</i>
<i>Value</i>	<i>Constructivist</i>	<i>Focuses primarily on identifying multiple values and perspectives through qualitative methods</i>
<i>Social Justice</i>	<i>Transformative</i>	<i>Focuses primarily on viewpoints of marginalized groups and interrogating</i>

Sumber: Merten and Wilson (2012. p.56)

Cabang Metode dalam Evaluasi

Cabang metode evaluasi merefleksikan akar dari penelitian sosial, dalam hal ini termasuk menggunakan metode penelitian secara ketat, dan secara luas berdasarkan asumsi-asumsi dari paradigma postpositivism dan positivism.

Tabel 2. Postpositivist Paradigm on The Method Branch

Description	Axiological assumption	Ontological assumption	Epistemological assumption	Methodological assumption
Focuses primarily on quantitative design and data	Respect Justice Beneficiary	One reality knowable within a certain level of probability	Distant Objectives	Scientific Method Hypothesis Quantitative Method

Sumber: Merten and Wilson (2012. p.60).

Alkin dan Christie (2004: 17-18) menganalisis tokoh-tokoh yang bekerja pada cabang metode evaluasi yakni **Ralph Tyler** sebagai titik awal dan utama untuk evaluasi program yang modern. Hasil karyanya dan gagasannya sangat luas, mempengaruhi pekerjaan banyak ahli teori evaluasi masa depan. Tyler mengembangkan gagasannya mengenai klasifikasi taksonomi hasil pembelajaran, kebutuhan untuk memvalidasi indikator, konsep evaluasi formatif, penguasaan konten, evaluasi yang berorientasi pada keputusan, kriteria yang direferensikan dan tes yang direferensikan oleh tujuan. Tyler menegaskan bahwa kurikulum yang akan dievaluasi didasarkan pada hipotesis. Tujuan evaluasi adalah untuk memvalidasi hipotesis program.

Fokus utama Tyler pada spesifikasi tujuan dan pengukuran hasil, menolak penerapan tes yang direferensikan oleh norma untuk evaluasi program. Sudut pandang Tyler dalam Alkin (2004:18) dikenal sebagai evaluasi berorientasi tujuan (atau sasaran-referensi). Pendekatan ini berfokus pada: (a) merumuskan pernyataan tujuan pendidikan, (b) mengklasifikasikan tujuan-tujuan tersebut ke dalam jenis-jenis utama, (c) mendefinisikan dan menyempurnakan masing-masing jenis tujuan dalam perilaku, (d) mengidentifikasi situasi di mana siswa diharapkan menampilkan jenis perilaku tersebut, (e) memilih dan mencoba metode yang tepat untuk memperoleh bukti mengenai setiap jenis tujuan, (f) berdasarkan uji coba awal memilih metode penilaian yang lebih tepat untuk pengembangan dan peningkatan lebih lanjut, dan (g) merancang sarana (alat) untuk menafsirkan dan menggunakan hasil evaluasi. Tyler menciptakan istilah evaluasi pendidikan pada tahun 1930-an untuk menggambarkan prosedur evaluasi yakni membandingkan hasil-hasil yang dimaksudkan (tujuan) dengan hasil yang sebenarnya (terukur dengan baik).

Gagasan **Donald Campbell** dalam Alkin dan Christie (2004: 21) yang paling menonjol mengenai desain eksperimental dan quasi-eksperimental untuk penelitian-penelitian sosial. Pada masa itu atas nama “ilmu pengetahuan yang baik,” ilmu-ilmu sosial menggunakan metode ilmiah yang ditemukan dalam ilmu fisika. Desain eksperimental masih sering disebut-sebut sebagai yang paling ketat dari semua desain penelitian atau sebagai “standar emas” yang digunakan untuk menilai semua desain

lainnya. Campbell mengakui kekurangan metode eksperimen digunakan dalam ilmu sosial karena dalam banyak konteks, eksperimen tidak dapat mencapai sesuatu yang diinginkan. Sebagai alternatif untuk eksperimen dikembangkan “desain quasi-eksperimental.” Ide-ide yang dikemukakan ini menjadi fondasi hampir pada semua metode penelitian ilmu sosial dan berdampak pada bertambahnya ide-ide baru yang terkait dengan desain penelitian. Pengenalan dan tanggapan terhadap desain quasi-eksperimental membantu munculnya iklim yang lebih menerima pemikiran alternatif mengenai penelitian ilmu sosial. Karya Campbell diarahkan pada peneliti ilmu sosial, dan relevansinya untuk evaluasi. Campbell juga menulis makalah yang menunjukkan potensi kesesuaian metode kualitatif sebagai pelengkap metode eksperimen kuantitatif. Metode kualitatif sangat membantu untuk area evaluasi seperti implementasi, identifikasi, dan interpretasi dampak. Namun hal ini menjadikan metode evaluasi semakin luas. Campbell adalah seorang eksperimentalis telah membentuk cara evaluasi dengan quasi-eksperimen.

Metode penelitian untuk evaluasi ditulis oleh **Edward Suchman**, dalam sebuah buku berjudul *Evaluative Research* pada tahun 1967. Judul ini menunjukkan pandangannya tentang evaluasi sebagai bentuk penelitian, dengan membedakan antara evaluasi sebagai penggunaan logika (akal sehat) dan penelitian evaluatif yang menggunakan metode dan teknik penelitian ilmiah. Dia menegaskan penggunaan yang tepat dari kata evaluatif sebagai kata sifat yang menentukan jenis penelitian. Suchman menganut pentingnya melakukan penelitian evaluatif dengan cara ilmiah. Untuk itu evaluator harus mengubah prosedur dengan bentuk kelompok kontrol atau pembanding agar evaluasi menjadi ilmiah. Selain mengakui kriteria ilmiah, evaluator juga harus mengakui kriteria administratif untuk menentukan kelayakan penelitian. Penelitian evaluatif diterapkan untuk menentukan sejauh mana suatu program atau prosedur tertentu mencapai hasil yang diinginkan. Keberhasilan evaluasi akan bergantung pada kegunaannya bagi administrator dalam meningkatkan layanan. Suchman mengakui peran penilaian dalam evaluasi, dan mencatat bahwa nilai-nilai dan asumsi mempengaruhi perumusan tujuan untuk penelitian evaluative. Pengaruh Suchman dalam metode evaluasi adalah identifikasinya atas lima kategori evaluasi, yakni: (1) upaya (kuantitas dan kualitas kegiatan yang terjadi), (2) kinerja (kriteria dampak yang mengukur hasil usaha), (3) kecukupan kinerja (sejauh mana kinerja memadai untuk jumlah total kebutuhan), (4) efisiensi (pemeriksaan jalur alternatif atau metode dalam hal biaya manusia dan moneter), dan (5) proses (bagaimana dan mengapa program bekerja atau tidak berfungsi).

Peter Rossi terkenal karena bukunya yang sangat populer dengan judul *Evaluation: A Systematic Approach*, tulisan yang sangat populer dan berpengaruh. Buku tersebut mendiskusikan tentang pengumpulan data kualitatif, pemanfaatan evaluasi, peran stakeholder, dan sebagainya. Tulisan-tulisan Rossi sebelumnya menekankan penggunaan desain eksperimen dan berkontribusi yang signifikan terhadap konseptualisasi evaluasi yang digerakkan oleh teori dan evaluasi komprehensif. Ia menggambarkan evaluasi sebagai penerapan sistematis prosedur penelitian sosial dalam menilai konseptualisasi dan desain, implementasi, dan kegunaan program intervensi sosial. Oleh karena itu

metode evaluasi yang diterapkan disesuaikan dengan tahap program. Misalnya, jika melakukan evaluasi program yang berjalan dengan baik maka evaluasi akan dirancang untuk fokus pada pekerjaan saat ini dari pada pengembangan program. Tetapi itu tidak berarti bahwa pengembangan program tidak akan ditangani oleh evaluasi; hanya akan kurang mendapat perhatian jika dibandingkan dengan evaluasi yang dimulainya bersamaan dengan dimulainya program.

Rossi, bersama dengan Weiss dan Chen, memberikan beberapa pemikiran mendasar tentang evaluasi yang digerakkan oleh teori. Evaluasi yang digerakkan oleh teori melibatkan konstruksi teori program yang terperinci yang kemudian digunakan untuk memandu evaluasi. Rossi mempertahankan bahwa pendekatan metode membantu mempertemukan dua jenis utama validitas yakni internal dan eksternal. Penggabungan ini memungkinkan evaluator untuk menangani validitas internal melalui pembuatan model program tertentu dan validitas eksternal implisit dengan menyelidiki cara-cara di mana program yang serupa dan yang berbeda, hasil evaluasi harus digeneralisasikan.

Cabang Nilai (*Value*) Dan Menilai (*Valuing*)

Memberikan nilai (*value*) dan harga (*worth*) merupakan pekerjaan evaluator (Gephart, 1981: 5). Akan tetapi hal ini masih memunculkan berbagai permasalahan yang berkaitan dengan *value* (nilai) dan *valuing* (penilaian) dalam evaluasi. Masalah itu berkaitan dengan beberapa pertanyaan tentang: apakah itu nilai? Apa karakteristik yang membuat sesuatu itu bernilai? apa struktur nilai itu? Apa peranan usaha dari *value* dan *valuing*? Kesulitan lain adalah ketidakmampuan mengkonstruksi nilai atau dengan kata lain tidak ada metodologi yang cukup exist untuk menentukan nilai. Meskipun dengan masalah tersebut menentukan nilai, bolehkah dilembagakan sebagai tugas profesi evaluator?

Gephart (1981: 5) selanjutnya menjelaskan bahwa *Value* (nilai) dan *valuing* (menilai) merupakan jantung dari evaluasi. Hanya saja *value* dan *valuing* sebagai *core* dari evaluasi tidak secara luas diformulasikan dalam proses evaluasi di masa silam. Lebih banyak evaluasi diformulasikan sebagai usaha yang dilakukan untuk menjelaskan proses program berdasarkan sebuah pencarian kebenaran, tidak untuk nilai. Pergeseran kebutuhan evaluasi dari pencarian kebenaran kepada *value* dan *valuing* perlu untuk dipahami tidak hanya diketahui saja oleh mereka yang melakukan evaluasi, apalagi ketika *value* dan *valuing* di klaim sebagai bagian terdepan dari sebuah profesi evaluator. Tujuan dari evaluasi kebijakan dan program adalah untuk meningkatkan efektifitas, efisiensi ataupun kinerja sebuah program bukan untuk membuktikan.

Hal yang menarik adalah ketika terminologi sentral dari konstruk dan proses evaluasi adalah nilai dan menilai akan tetapi hal ini masih sedikit didiskusikan dalam literatur evaluasi. Juga menarik bahwa nilai/*value* dan menilai (*valuing*) adalah sesuatu yang sangat simpel dan sangat kompleks. Semua adalah nilai (*value*), nilai kebebasan, nilai keindahan matahari terbenam, nilai pengetahuan, nilai orang-orang dimasa lalu, nilai hubungan dan nilai pemikiran, akan tetapi kita memiliki kesulitan

dengan menilai. Menemukan ide dari nilai menjadi sangat kompleks, abstrak dan *elusive*. Ini sesuatu yang sering menjadi sumber dari usaha memahami dan menyelesaikan proses evaluasi.

Nilai (*value*) dan menilai (*valuing*) berperan sentral dalam evaluasi. Nilai didasarkan pada konteks yang secara kualitas dan alamiah ditetapkan dalam sebuah setting. Problem pemikiran yang ditemui adalah relativitas terminology “*absolute value*”, “*inherent value*”, dan “*fundamental value*”.

Andriessen (2005: 4) menuliskan empat cara menentukan nilai. Menilai membutuhkan suatu objek untuk dinilai, kerangka kerja untuk penilaian, dan kriteria yang mencerminkan kegunaan atau keinginan dari objek. Dibawah ini empat cara dalam menentukan nilai:

1. Mendefinisikan kriteria nilai dalam istilah moneter, dalam hal ini metode untuk menentukan nilai adalah metode penilaian keuangan.
2. Menggunakan kriteria non-moneter dan menerjemahkannya ke dalam fenomena yang dapat diamati, dalam hal ini metodenya adalah metode pengukuran nilai.
3. Jika kriteria tidak dapat diterjemahkan ke dalam fenomena yang dapat diamati tetapi sebaliknya tergantung pada penilaian pribadi oleh evaluator, maka metodenya adalah metode penilaian nilai.
4. Jika kerangka kerja tidak termasuk kriteria untuk nilai tetapi melibatkan skala metrik yang berhubungan dengan fenomena yang dapat diamati, maka metode ini adalah metode pengukuran.

Berbagai Teori tentang *Valuing* (menilai) dalam Evaluasi

Para teorist yakin bahwa perbedaan antara evaluator dengan peneliti lain adalah evaluator itu melakukan aktifitas menilai (*valuing*) pada temuan mereka. Pada cabang ini pengarusutamanya adalah **Michael Scriven** (1967) yang dengan brilian menunjukkan sebagian besar gambaran kerangka utama konseptual dan prosedur menilai (*valuing*). Kontribusi utama Scriven adalah mendefinisikan secara tetap peran evaluator dalam membuat *judgment* nilai. Scriven mencatat: buruk katakan buruk dan baik katakan baik dan ini adalah tugas evaluator untuk memutuskannya. Kegagalan besar evaluator adalah memberikan informasi dengan simpel pada pembuat keputusan dan menghindari kewajiban untuk memberikan *final judgment*, dan menunjukkan evaluator yang tidak profesional.

Evaluator dalam menilai harus memenuhi aturan dalam melayani kepentingan publik (*public interest*). Maksudnya, tanggung jawab evaluator tidak hanya sampai pada klien, *user* atau stakeholder, akan tetapi sampai pada semua konsumen potensial. Peran Evaluator dalam menilai sama dengan membuat laporan untuk konsumen dimana evaluator menentukan criteria-kriteria tertentu untuk memberikan *judgment* yang dibuat dan menyajikannya berdasarkan semua hal yang dilihat. Evaluator bertanggung jawab untuk mengidentifikasi alternatif-alternatif kunci dalam membuat *judgment* nilai. Scriven memperkenalkan metode alternatifnya untuk desain eksperimen yakni *quasi eksperiment* yang disebutnya dengan metode modus operanadi (MO). MO merupakan assosiasi gambaran dari *event*, proses yang berurutan waktu yang dapat digambarkan sebagai karakteristik *causal chain* yang

menghubungkan penyebab dengan dampak. Scriven adalah *leading theorist* dari perspektif nilai dan menyediakan “*spiritual guidance*” dan arah bagi evaluator lain dalam dalam cabang menilai.

Pendekatan evaluasi **Elliot Eisner** sangat berbeda dengan Scriven walaupun keduanya sama-sama menghargai pentingnya menilai dan *judgment*. Eisner memberi perhatian pada membuat *judgment* nilai tentang kualitas beberapa object, situasi dan proses. Eisner pertama kali mengemukakan pandangannya pada apa yang disebut “*educational Connoisseurship*” dalam *Journal of Aesthetic Education* (1976). Pandangan evaluatifnya merespon pandangan negativanya dalam pendekatan tradisional yang focus pada *outcome* pendidikan yang diukur dengan tes terstandar menggunakan prinsip-prinsip tes psikologi atau dengan prosedur *criterion-referenced test*. Eisner menolak *technological Scientism*, termasuk menolak penggunaan desain eksperimen dan quasi eksperimen yang tergantung pada metode kuantitatif. Dia mencatat bahwa pola-pola berpikir tidak dapat diukur secara kuantitatif. Dia setuju teknik kuantitatif memberikan informasi yang berguna, tetapi evaluasi memerlukan sesuatu yang *sophisticated*, interpretif tidak hanya terpisah dalam pola-pola yang signifikan, tetapi juga memahami makna dari apa yang diketahui.

Pandangan **Robert Stake** pada evaluasi responsive memiliki dampak yang substansial pada teori evaluasi yang digunakan pada cabang menilai. Dia sangat sangat mendukung pada studi kasus dan produk evaluator sebagai “*thick description*”. Komponen esensial dari evaluasi responsif Stake adalah: a. meyakini bahwa ada nilai tidak benar pada tiap sesuatu (pengetahuan terikat pada konteks) b. meyakini bahwa perspektif stakeholder adalah elemen integral dalam evaluasi dan, c. meyakini bahwa studi kasus adalah metode terbaik untuk merepresentasikan keyakinan dan nilai dari stakeholder dan dari laporan hasil evaluasi. Menurut dia dalam studi kasus memerlukan konteks dan deskripsi aktifitas tetapi pada beberapa *audience* kekurangan cara untuk memberikan *judgment* kualitas. Dia berpikir bahwa aktifitas *judgment* adalah bagian dari deskripsi, bagian dari aktifitas observasi. Dengan kata lain melihat dan *judgment* evaluasi adalah bagian dari aktifitas dan tugas dari evaluasi. Stake juga mengemukakan adanya *multiple* realitas dan perspektif kebutuhan stakeholder yang direpresentasikan dalam evaluasi, dia yakin bahwa stakeholder tidak berpartisipasi dalam evaluasi sebagaimana partisipasi para ahli teori. Bagaimanapun *judgment* tidak berdiri sendiri. Tugas evaluator mendengarkan permohonan partisipan dengan sengaja, kadang-kadang negosiasi, untuk memutuskan apakah partisipan tertarik.

Barry MacDonald telah ditempatkan pada cabang menilai. Ia mengenali beragam perspektif-perspektif yang dimiliki oleh stakeholder dan percaya tugas penilai adalah untuk menyajikan nilai-nilai dari stakeholder yang berbeda. MacDonald melukiskan evaluator sebagai ‘perunding dari bermacam-macam perspektif’. Evaluator akan mengumpulkan dan mengkomunikasikan alternatif-alternatif definisi, perspektif, dan *judgment* berpegang pada orang-orang di dalam dan disekitar program. Dia menggambarkan evaluasi sebagai peristiwa atau kejadian dalam konteks birokrasi, otokrasi atau demokrasi. Evaluasi birokrasi dilakukan untuk kepentingan orang-orang dalam kekuasaan (*power*). Dalam konteks ini evaluator menerima nilai-nilai dari mereka yang memegang dan menawarkan

informasi yang membantu mereka menyelesaikan tujuan kebijakan mereka. Evaluasi-evaluasi otokratis serupa dengan evaluasi birokratis dengan pengecualian pada independensi evaluator. Dalam konteks ini evaluator fokus pada isu-isu jasa dan tindakan dari *expert advisor*. Teknik dari studi ini harus menghasilkan bukti ilmiah. Evaluasi demokratis dilaksanakan untuk kepentingan informasi bagi masyarakat. Dalam hal ini penilai mengenali pluralisme nilai dan mencari representasi merata dari *interest* dalam formulasi nilai. Aktivitas utamanya adalah mengumpulkan definisi-definisi dan reaksi pada program. MacDonald mendukung penggunaan evaluasi demokratis karena evaluasi ini dapat melukiskan berbagai kenyataan satu program dengan kebenaran dan keadilan, meninggalkan pembuatan keputusan dengan banyak aturan dari berbagai perspektif dan *judgment* untuk pertimbangan ketika mereka membuat keputusan. Evaluasi demokratis mengizinkan evaluator untuk mengadopsi sebuah pendirian (secara politik). Dalam peran ini evaluator dapat memfasilitasi negosiasi antar sudut pandang, perspektif dan nilai-nilai yang berbeda.

Ernest House memahami bahwa layanan evaluasi bertujuan menyediakan informasi pada pembuat keputusan sehingga mereka dapat menentukan legitimasi alokasi pentingnya sumber daya. House memperlihatkan bahwa evaluasi tidak pernah netral nilai. Evaluator berhadapan dengan tugas memahami kebutuhan dan posisi kelompok stakeholder yang bervariasi, terutama orang miskin dan tidak memiliki kekuasaan (*poor dan powerless*) dan menjaga keseimbangan informasi dari persepsi keadilan. Dalam melakukan penyeimbangan ini evaluator membentuk bermacam informasi yang dikumpulkan juga analisisnya. Esensinya, evaluator menempatkan diri mereka dalam peran juru bicara orang-orang atau representasi dari *poor dan powerless*. Penting untuk dicatat bahwa oleh House, peran dari penilai bukan untuk menggambarkan nilai-nilai dalam terminologi baik atau buruk seperti Scriven tetapi dalam terminologi benar, terbuka, dan adil (*right, fair, and just*).

Scriven dan House menempatkan komponen menilai dalam posisi unggul tetapi mereka melakukan dengan menekankan pada substansi yang berbeda. Hal ini merupakan informasi untuk menguji apa yang oleh House dikenal sebagai "*ethical fallacies*" dalam evaluasi, yaitu: *clientism (taking the client's interest as the ultimate consideration)*; *contractualism (adhering inflexibly to the contract)*; *managerialism (placing the interest of the managers above all else)*; *methodologicalism (believing that proper methodology solves all ethical problems)*; *pluralism/elitism (including only the powerful stakeholders' interests in the evaluation)*; and *relativism (taking all viewpoints as having equal merit.)*

Egon Guba and Yvonna Lincoln ditempatkan sebagai generasi keempat pada cabang menilai. Dalam pandangannya, evaluator adalah *penilai (valuer)*. Guba dan Lincoln memandang bahwa stakeholder sebagai individu primer yang dilibatkan dalam penempatan nilai. Point-point pandangannya diturunkan pada keyakinan bahwa sebagai ganti dari adanya satu realitas, terdapat multiple realitas yang didasarkan pada persepsi dan interpretasi dari individu yang terlibat dalam program evaluasi. Guba dan Lincoln yakin bahwa peran evaluator adalah memfasilitasi negosiasi antar individu-individu dalam merefleksikan realitas yang jamak. Generasi keempat ini mendasarkan pada paradigma konstruktivisme

yakni menempatkan eksistensi realitas tunggal, dan individu-individu mengkonstruksi persepsinya pada realitas itu. Evaluator adalah orkestrator dari proses negosiasi menuju pada puncak konsensus pada informasi yang terbaik dan konstruksi yang lebih *sophisticated*.

Cabang Penggunaan Evaluasi

Cabang yang ketiga fokus pada cara di mana informasi hasil evaluasi akan digunakan, dan siapa yang akan menggunakan informasi (Alkin dan Christie: 2004). Para teoretikus yang berorientasi pada keputusan merasa penting untuk melakukan evaluasi yang dirancang khusus untuk membantu para stakeholder program, dimana evaluasi menjadi kunci dalam pengambilan keputusan. Model CIPP oleh Stufflebeam adalah salah satu model yang paling terkenal dari teori-teori dalam cabang penggunaan evaluasi.

Stufflebeam, bersama dengan Guba mengembangkan model CIPP (Context, Input, Process dan Product) sebagai pendekatan untuk evaluasi yang difokuskan pada proses pengambilan keputusan. Evaluasi konteks (*context*) melibatkan identifikasi kebutuhan untuk memutuskan tujuan program. Evaluasi masukan (*input*) mengarahkan pada keputusan tentang strategi dan desain kebijakan dan program. Evaluasi proses (*process*) mengidentifikasi kekurangan dalam program saat implementasi. Evaluasi produk (*product*) mengukur hasil kebijakan dan program untuk keputusan mengenai kelanjutan atau fokus kembali pada program. Stufflebeam menggambarkan model evaluasi CIPP sebagai proses siklus. Strategi utamanya adalah bekerja dengan evaluasi yang dirancang dengan cermat dengan mempertahankan fleksibilitas. Dalam hal ini evaluator harus melihat desain sebagai proses, bukan produk. Evaluasi harus memberikan aliran informasi yang berkelanjutan kepada pengambil keputusan untuk memastikan bahwa program secara berkelanjutan meningkatkan layanan mereka. Untuk meningkatkan layanan, evaluasi harus membantu pengambil keputusan dalam mengalokasikan sumber daya program yang terbaik dalam melayani klien.

Stufflebeam menekankan evaluasinya pada Standar Evaluasi Program merupakan standar profesional yang harus digunakan oleh evaluator untuk melakukan pekerjaan mereka. Standar Evaluasi Program fokus pada empat domain yang terkait dengan praktik yakni utilitas, kelayakan, kepatutan, dan akurasi. Standar utilitas dimaksudkan untuk memastikan bahwa evaluasi akan melayani kebutuhan informasi dari pengguna yang dituju. Standar kelayakan dimaksudkan untuk memastikan bahwa evaluasi akan realistis, bijaksana, diplomatik, dan hemat. Standar kepatutan dimaksudkan untuk memastikan bahwa evaluasi akan dilakukan secara legal, etis, dan dengan hormat untuk kesejahteraan mereka yang terlibat dalam evaluasi, serta mereka yang terkena dampak dari hasilnya. Standar akurasi dimaksudkan untuk memastikan bahwa evaluasi akan mengungkapkan dan menyampaikan informasi yang secara teknis memadai tentang fitur yang menentukan nilai atau manfaat dari program yang sedang dievaluasi. Evaluasi juga harus didasarkan pada kontrak yang jelas dengan standar kelayakan kontrak formal. Evaluator dan klien memiliki perjanjian formal tertulis yang merinci tentang apa yang

harus dilakukan, bagaimana, oleh siapa, dan kapan sehingga keduanya berkewajiban untuk mematuhi semua ketentuan perjanjian atau secara formal untuk menegosiasikan kembali.

Dengan menggunakan pendekatan Stufflebeam ini evaluator melibatkan panel stakeholder yang representatif untuk membantu menentukan pertanyaan evaluasi, membentuk rencana evaluasi, meninjau draf laporan dan menyebarkan temuan. Keberhasilan evaluasi bergantung pada interaksi reguler antara evaluator dengan stakeholder yang diyakini bahwa tanpa mereka pendekatan evaluasi akan gagal. Pendekatan evaluasi Stufflebeam melibatkan para stakeholder yang biasanya berada dalam posisi pengambilan keputusan dalam memfokuskan evaluasi dan dalam memastikan evaluasi membahas pertanyaan mereka yang paling penting, menyediakan informasi yang tepat waktu dan relevan untuk membantu pengambilan keputusan dan menghasilkan catatan akuntabilitas. Informasi formatif dan sumatif disediakan dalam panel stakeholder untuk meningkatkan penggunaan hasil evaluasi.

Michael Patton pada cabang penggunaan evaluasi menekankan pada prosedur yang akan meningkatkan penggunaan evaluasi ke spektrum yang lebih luas dari para stakeholder. Penelitian utama tentang pemanfaatan evaluasi mempengaruhi perluasan dari evaluasi. Evaluasi tidak hanya berkaitan dengan pengambil keputusan, tetapi juga memikirkan tentang kewajiban evaluator untuk membantu memastikan pemanfaatan dari evaluasi. Sudut pandang ini menegaskan bahwa evaluator harus proaktif dan tidak puas dengan evaluasi yang mungkin hanya diletakkan di rak, oleh karena itu evaluator harus membuat agar evaluasi lebih berpotensi digunakan.

Penjelasan teoritis yang paling menonjol dari pemanfaatan (atau penggunaan) yang dikembangkan oleh Michael Patton berbeda dengan para ahli teori evaluasi sebelumnya. Patton menyatakan bahwa evaluator harus mencari individu yang cenderung menjadi pengguna sebenarnya dari evaluasi. Patton menyajikan diagram alur evaluasi yang berfokus pada penggunaan *utilization-focused evaluation* (UFE). Empat fase utama UFE adalah: a) mengidentifikasi pengguna yang dituju; komitmen pengguna untuk fokus yang dituju dari evaluasi dan untuk pemanfaatan evaluasi; b) keterlibatan dalam metode, desain, dan pengukuran; c) keterlibatan, aktif dan secara langsung menafsirkan temuan dan membuat penilaian; dan d) membuat keputusan tentang penyebaran lebih lanjut. Kemungkinan evaluasi yang digunakan akan sangat meningkat sesuai dengan identifikasi orang-orang yang memiliki saham dalam evaluasi dan yang secara pribadi peduli dengan temuan yang dihasilkannya. Selain itu, evaluator harus secara aktif terlibat dalam mengembangkan komitmen pengguna untuk pemanfaatan potensial.

Marvin Alkin diidentifikasi sebagai theorist yang fokus pada evaluasi dan masalah pengambilan keputusan. Model awal Alkin memiliki banyak kesamaan dengan model CIPP Stufflebeam. Perbedaan utama antara Alkin dan Stufflebeam adalah bahwa proses dan produk memiliki dimensi sumatif dan formatif. Dengan demikian, orang dapat melihat proses secara sumatif (dokumentasi program) atau produk secara formatif (yang berarti hasil). Pandangan Alkin menempatkannya sebagai theorist evaluasi yang berorientasi pada pengguna, karena ia menyebutkan bahwa untuk meningkatkan kemungkinan

pemanfaatan evaluasi harus fokus pada pada pengguna potensial yang teridentifikasi. Dalam karya teoritisnya, Alkin sangat menolak keterlibatan evaluator sebagai agen yang menilai. Sebaliknya, ia lebih suka bekerja dengan pengguna primer di awal proses evaluasi untuk menetapkan sistem nilai untuk menilai data hasil yang potensial. Dalam sesi interaktif, ia menyajikan berbagai hasil potensial yang disimulasikan dan mencari penilaian (nilai) atas implikasi masing-masing. Alkin mengakui bahwa ada kondisi di mana ia tidak dapat dipertahankan untuk melibatkan pengguna utama yang dituju dalam proses prasangka ini. Dalam keadaan seperti itu, ia lebih suka menyajikan data evaluasi se-fakta mungkin tanpa memaksakan penilaian nilai, kecuali ada kasus ekstrim yang menuntut penilaian terjadi.

Jean King memperluas cabang penggunaan dengan fokusnya pada penerapan model evaluasi partisipatif. King mendefinisikan evaluasi sebagai “sebuah proses penyelidikan sistematis untuk memberikan informasi yang baik tentang karakteristik, kegiatan, atau hasil dari program atau kebijakan untuk tujuan yang berharga”. Definisi ini menggabungkan pentingnya penggunaan dengan menandakan bahwa evaluasi harus digunakan untuk “tujuan yang bernilai”. King menyebutkan bahwa evaluasi berkolaboratif dengan pengguna dengan kepentingan untuk meningkatkan peluang pemanfaatan. Teori King juga berkaitan dengan menciptakan lingkungan partisipatif sepanjang proses evaluasi. King berpendapat bahwa untuk keberhasilan evaluasi partisipatif harus ada: (a) struktur kekuasaan yang menerima, (b) berbagi makna pengalaman di antara peserta, (c) relawan dan pemimpin, (d) waktu yang cukup, (e) sumber daya yang cukup, dan (f) tingkat kepercayaan antarpribadi dan organisasi yang besar. King menganjurkan upaya keterlibatan dalam komunikasi dan diskusi untuk menciptakan makna bersama dalam merumuskan metode penelitian, mengumpulkan, dan menafsirkan data. Membangun kepercayaan antara evaluator dan stakeholder sebagai salah satu persyaratan mendasar untuk kesuksesanevaluasi partisipatif. Olehkarena itu para evaluator harus memperhatikan dengan cermat dinamika antarpribadi yang terjadi selama evaluasi. Evaluasi tanpa interaksi interpersonal yang efektif sebagai mesin tanpa pelumasan yang tepat. Ia juga mendefinisikan praktik evaluasi interaktif sebagai tindakan yang disengaja untuk melibatkan orang-orang dalam membuat keputusan, mengambil tindakan, dan merefleksikan sambil melakukan studi evaluasi. Evaluator harus fokus pada memainkan peran pengambil keputusan, aktor, dan praktisi reflektif. Tiga peran yang disarankan oleh King mengakui pentingnya faktor interpersonal ketika melakukan evaluasi. Pedoman yang disarankan King ditujukan untuk meningkatkan peluang pemanfaatan dan evaluasi partisipatif.

John M.Owen memusatkan pada bidang evaluasi pengembangan dan perubahan organisasi. Di bidang ini, perbedaan antara konsultan pengembangan organisasi dan evaluator sering diperdebatkan. Bagi Owen, pemanfaatan evaluasi adalah memberikan pertimbangan yang sangat penting ketika merencanakan, bernegosiasi, merancang, dan melakukan evaluasi. Untuk itu evaluator agar beradaptasi dengan kebutuhan dan kepedulian stakeholder. Fleksibilitas inilah yang membedakannya dari evaluator dengan pengembangan organisasi lainnya. Menurut Owen, fleksibilitas dan keinginan untuk bekerja secara kolaboratif dengan para pemangku kepentingan akan meningkatkan potensi

penggunaan evaluasi. Owen mendukung partisipasi stakeholder dalam berbagai tahap evaluasi tetapi bahwa keterlibatan stakeholder yang paling penting adalah selama tahap perencanaan dan negosiasi. Evaluator harus siap untuk mengakui kerangka kepentingan klien, sehingga pengetahuan yang mereka hasil evaluasi memiliki arti penting untuk keputusan berikutnya. Owen menjelaskan banyak teknik, pedoman, dan kerangka kerja untuk perencanaan dan negosiasi evaluasi yang fleksibel dengan para stakeholder.

David Fetterman adalah penulis *Empowerment Evaluation* pada tahun 1996 yang ia gambarkan sebagai proses yang mendorong *self-determination* di antara para penerima evaluasi program, termasuk di dalamnya program pelatihan, fasilitasi, advokasi, iluminasi, dan pembebasan. Tujuan evaluasi pemberdayaan untuk menumbuhkan *self-determination* dari pada *dependency* peserta program termasuk klien yang pada dasarnya melakukan evaluasi mereka sendiri. Evaluator luar sering berfungsi sebagai pelatih atau fasilitator tambahan, memberikan pengetahuan dan alat untuk penilaian diri dan akuntabilitas berkelanjutan kepada klien. Fetterman merasa bahwa peserta pelatihan mengevaluasi program diri mereka sendiri dan melatih mereka melalui desain evaluasi, hal ini merupakan bentuk pemberdayaan yang efektif.

Fetterman menjelaskan mengenai dua bentuk umum evaluasi pemberdayaan yang keduanya hanya memiliki sedikit perbedaan. Perbedaan utama antara keduanya adalah sejauh mana evaluator berpartisipasi dalam proses evaluasi. Pertama, evaluator mengajari peserta program untuk melakukan evaluasi program mereka sendiri, dan membuat mereka lebih mandiri. Kedua, evaluator berfungsi sebagai pelatih untuk membantu orang lain melakukan evaluasi mereka sendiri. Dengan evaluator menggunakan pendekatan ini, memungkinkan peserta untuk membentuk arah evaluasi, menyarankan solusi ideal untuk masalah mereka, dan kemudian mengambil peran aktif dalam membuat perubahan sosial. Menurut Fetterman, titik akhir evaluasi bukanlah penilaian terhadap nilai program. Dalam pandangannya, nilai dan menilai tidak statis. Evaluasi sebagai proses yang berkelanjutan melalui internalisasi dan pelembagaan proses evaluasi diri dan praktik, pendekatan yang dinamis dan responsif terhadap evaluasi dapat dikembangkan untuk mengakomodasi pergeseran populasi, tujuan, penilaian nilai dan kekuatan eksternal”. Oleh karena itu penting untuk membedakan antara evaluasi partisipatif dan evaluasi pemberdayaan. Evaluasi partisipatif didasari oleh konsepsi keterlibatan stakeholder dan evaluasi pemberdayaan terlihat sangat mirip yakni stakeholder memutuskan dan membantu dalam melakukan semua aspek evaluasi, termasuk desain, implementasi, analisis, dan interpretasi.

SIMPULAN

Melalui pohon teori evaluasi kita dapat memahami secara komprehensif keilmuan evaluasi sebagai ilmu terapan. Ilmu evaluasi bersifat multidisiplin maka teori-teori dibangun dari tiga bidang ilmu yakni *epistemology*, *social inquiry* dan *social accountability*. Ketiga akar ilmu ini membangun pohon ilmu evaluasi dengan tiga cabang evaluasi meliputi cabang metode (*method*), cabang menilai (*valuing*),

dan cabang penggunaan (*Use*). Di setiap cabang terdapat ahli-ahli evaluasi yang saling memperkokoh tumbuh dan kuatnya cabang ilmu.

Dalam melakukan praktek evaluasi penting untuk mendasarkan pada ketiga cabang tersebut. Evaluasi mengharuskan menggunakan metode ilmiah sebagaimana riset dalam ilmu-ilmu sosial. Peneliti evaluasi dalam melakukan evaluasi kebijakan secara simultan harus memberikan nilai dan menilai terhadap hasil-hasil riset yang merupakan inti dari evaluasi. Pekerjaan menilai inilah yang membedakan antara peneliti konvensional dan peneliti evaluasi. Penilaian yang diberikan oleh penelitian evaluasi berguna bagi keputusan kebijakan dan program, apakah kebijakan dilanjutkan, diperbaiki atau dihentikan. Dalam melakukan evaluasi kebijakan dan program pendidikan juga tidak terlepas dari penggunaan konsep-konsep dan teori-teori evaluasi pada umumnya, hanya lebih fokus pada konten pendidikan.

DAFTAR PUSTAKA

- Andriessen. 2005. *Value, Valuation, and Valorisation*. Published by <http://www.innovatedutch.com>
- Blaine R. Worthen, James R. Sanders. 1987. *Educational evaluation: alternative approaches and practical guidelines*. New York and London: Longman Inc.
- Grindle, M. and Thomas, J. (1990) 'After the Decision: Implementing Policy Reforms in Developing Countries'. *World Development*. Vol. 18 (8).
- Marvin C. Alkin, and Christie, Christina. A. 2004. *An Evaluation Theory Tree*. Published by www.sagepub.com/upm-data/5074_Alkin_Chapter_2.pdf
- Marvin C. Alkin. 2012. *Evaluation Roots: An International Perspective*. University of California, Los Angeles. International Development Research Centre. Published by journals.sfu.ca/jmde/index.php/jmde_1/article
- Gephart, William J. 1981. *Value, Valuing, and Evaluation, Research on Evaluation Program*. Northwest Regional Educational Laboratory.
- Mertens, Donna M., and Wilson, Amy T. 2012. *Program Evaluation Theory and Practice, First Edition: A Comprehensive Guide*. Guilford Press: USA