EVALUASI PROGRAM PELATIHAN
(Training Program Evaluation)

Oleh :Dr. S. Eko Putro Widoyoko, M.Pd.

Abstrak

Kegiatan evaluasi program pelatihan tidak hanya dilaksanakan pada akhir kegiatan program, tetapi sebaiknya dilakukan sejak awal, yaitu mulai dari penyusunan rancangnan program pelatihan, pelaksanaan program pelatahan dan hasil dari pelatihan. Penilaian hasil pelatihan tidak cukup hanya pada hasil jangka pendek (output) tetapi dapat menjangkau hasil dalam jangka panjang (outcome and impact program). Ada berbagai macam model evaluasi program yang dapat dipilih untuk mengevaluasi program pelatihan. Model mana yang akan digunakan tergantung pada tujuan maupun kemampuan evaluator. Siapapun yang ditunjuk menjadi evaluator, agar hasil evaluasi dapat maksimal maka kompetensi evaluator harus dipertimbangkan. Kompetensi evaluator meliputi kompetensi manajerial, kompetensi teknis, kompetensi konseptual dan kompetensi bidang studi.

A. Pendahuluan

Dalam manajemen SDM terdapat beberapa fungsi pokok, dan fungsi evaluasi merupakan salah satu di antaranya, selain perencanaan, pengorganisasian dan pelaksanaan. Program pelatihan sebagai salah strategi pengembangan SDM memerlukan fungsi evaluasi untuk mengetahui efektivitas program yang bersangkutan. Pada umumnya orang beranggapan bahwa evaluasi program pelatihan diadakan pada akhir ahir pelaksanaan pelatihan. Anggapan yang demikian adalah kurang tepat, karena evaluasi merupakan salah satu mata rantai dalam sistem pelatihan yang jika dilihat dari waktu pelaksanaannya kegiatan penilaian dapat berada di awal proses perencanaan, di tengah proses pelaksanaan dan pada akhir penyelenggaraan pelatihan dan pasca kegiatan pelatihan. Penilaian yang dilaksanakan pada proses perencanaan disebut dengan analis kebutuhan (*need assessment*) yang berusaha untuk mengumpulkan informasi tentang kemampuan, ketrampilan maupun keahlian yang akan dikembangkan dalam pelatihan, karakteristik peserta pelatihan, kualitas materi pelatihan dilihat dari relevansi dan kebaharuan, kompetensi pelatih/instruktur/pengajar, tempat pelatihan beserta sarana dan prasarana yang dibutuhkan, akomodasi dan konsumsi serta jadwal kegiatan pelatihan. Penilaian yang dilaksanakan pada saat proses pelatihan disebut dengan monitoring yang bertujuan untuk mengumpulkan informasi tentang sejauh mana program yang telah disusun dapat diimplementasikan dengan baik. Dalam kegiatan monitoring ini berusaha untuk menilai kualitas proses pelatihan, baik dari aspek kinerja

instruktur, iklim kelas, sikap dan motivasi belajar atau berlatih para peserta pelatihan. Sedangkan penilaian pasca pelatihan bertujuan untuk mengetahui perubahan kinerja peserta setelah kembali ke tempat kerjanya masing-masing.

B. Tes, pengukuran, penilaian dan evaluasi

Ada tiga istilah yang sering digunakan dalam evaluasi, yaitu tes, pengukuran, dan penilaian. (*test, measurement,and assessment*). Tes merupakan salah satu cara untuk menaksir besarnya kemampuan seseorang secara tidak langsung, yaitu melalui respons seseorang terhadap stimulus atau pertanyaan (Djemari Mardapi, 1999: 2). Tes merupakan salah satu alat untuk melakukan pengukuran, yaitu alat untuk mengumpulkan informasi karakteristik suatu objek. Objek ini bisa berupa kemampuan peserta didik, sikap, minat, maupun motivasi. Respons peserta tes terhadap sejumlah pertanyaan menggambarkan kemampuan dalam bidang tertentu. Tes merupakan bagian tersempit dari evaluasi.

Pengukuran (*measurement*) dapat didefinisikan sebagai *the process by which information about the attributes or characteristics of thing are determinied and differentiated* (Oriondo,1998: 2). Guilford mendefinisi pengukuran dengan "*assigning numbers to, or quantifying, things according to a set of rules"* (Griffin & Nix, 1991: 3). Pengukuran dinyatakan sebagai proses penetapan angka terhadap individu atau karakteristiknya menurut aturan tertentu (Ebel & Frisbie. 1986: 14). Allen & Yen mendefinisikan pengukuran sebagai penetapan angka dengan cara yang sistematik untuk menyatakan keadaan individu (Djemari Mardapi, 2000: 1).Dengan demikian, esensi dari pengukuran adalah kuantifikasi atau penetapan angka tentang karakteristik atau keadaan individu menurut aturan-aturan tertentu. Keadaan individu ini bisa berupa kemampuan kognitif, afektif dan psikomotor. Pengukuran memiliki konsep yang lebih luas dari pada tes. Kita dapat mengukur karakateristik suatu objek tanpa menggunakan tes, misalnya dengan pengamatan, skala rating atau cara lain untuk memperoleh informasi dalam bentuk kuantitatif.

Penilaian (*assessment*) memiliki makna yang berbeda dengan evaluasi. The Task Group on Assessment and Testing (TGAT) mendeskripsikan asesmen sebagai semua cara yang digunakan untuk menilai unjuk kerja individu atau kelompok (Griffin & Nix, 1991: 3). Popham (1995: 3) mendefinisikan asesmen dalam konteks pendidikan sebagai sebuah usaha secara formal untuk menentukan status siswa berkenaan dengan berbagai

kepentingan pendidikan. Boyer & Ewel mendefinisikan asesmen sebagai proses yang menyediakan informasi tentang individu siswa, tentang kurikulum atau program, tentang institusi atau segala sesuatu yang berkaitan dengan sistem institusi. *"processes that provide information about individual students, about curricula or programs, about institutions, or about entire systems of institutions"* (Stark & Thomas,1994: 46). Berdasarkan berbagai uraian di atas dapat disimpulkan bahwa *assessment* atau penilaian dapat diartikan sebagai kegiatan menafsirkan data hasil pengukuran.

Evaluasi memiliki makna yang berbeda dengan penilaian, pengukuran maupun tes. Stufflebeam dan Shinkfield (1985: 159) menyatakan bahwa :

> *Evaluation is the process of delineating, obtaining, and providing descriptive and judgmental information about the worth and merit of some object's goals, design, implementation, and impact in order to guide decision making, serve needs for accountability, and promote understanding of the involved phenomena.*

Evaluasi merupakan suatu proses menyediakan informasi yang dapat dijadikan sebagai pertimbangan untuk menentukan harga dan jasa (*the worth and merit*) dari tujuan yang dicapai, desain, implementasi dan dampak untuk membantu membuat keputusan, membantu pertanggung jawaban dan meningkatkan pemahaman terhadap fenomena. Menurut rumusan tersebut, inti dari evaluasi adalah penyediaan informasi yang dapat dijadikan sebagai bahan pertimbangan dalam mengambil keputusan.

Komite Studi Nasional tentang Evaluasi (*National Study Committee on Evaluation*) dari UCLA (Stark & Thomas, 1994: 12), menyatakan bahwa:

> *Evaluation is the process of ascertaining the decision of concern, selecting appropriate information, and collecting and analyzing information in order to report summary data useful to decision makers in selecting among alternatives.*

Evaluasi merupakan suatu proses atau kegiatan pemilihan, pengumpulan, analisis dan penyajian informasi yang dapat digunakan sebagai dasar pengambilan keputusan serta penyusunan program selanjutnya. selanjutnya Griffin & Nix (1991:3) menyatakan:

> *Measurement, assessment and evaluation are hierarchial. The comparison of observation with the criteria is a measurement, the interpretation and description of the evidence is an assessment and the judgement of the value or implication of the behavior is an evaluation.*

Pengukuran, penilaian dan evaluasi bersifat hirarki. Evaluasi didahului dengan penilaian (*assessment*), sedangkan penilaian didahului dengan pengukuran. Pengukuran

diartikan sebagai kegiatan membandingkan hasil pengamatan dengan kriteria, penilaian (*assessment*) merupakan kegiatan menafsirkan dan mendeskripsikan hasil pengukuran, sedangkan evaluasi merupakan penetapan nilai atau implikasi perilaku.

Brikerhoff (1986:ix) menjelaskan bahwa evaluasi merupakan proses yang menentukan sejauh mana tujuan pendidikan dapat dicapai. Menurut Brikerhoff (1986:ix), dalam pelaksanaan evaluasi ada tujuh elemen yang harus dilakukan, yaitu: 1) penentuan fokus yang akan dievaluasi (*focusing the evaluation*), 2) penyusunan desain evaluasi (*designing the evaluation*), 3) pengumpulan informasi (*collecting information*), 4) analsis dan intepretasi informasi (*analyzing and interpreting*), 5) pembuatan laporang (*reporting information*), 6) pengelolaan evaluasi (*managing evaluation*), dan 7) evaluasi untuk evaluasi (*evaluating evaluation*).

Dalam pengertian tersebut menunjukkan bahwa dalam melakukan evaluasi, evaluator pada tahap awal harus menentukan focus yang akan dievaluasi dan desain yang akan digunakan. Hal ini berarti harus ada kejelasan apa yang akan dievaluasi yang secara implisit menenkankan adanya tujuan evaluasi, serta adanya perencanaan bagaimana melaksanakan evaluasi. Selanjutnya, dilakukan pengumpulan data, menganalisis dan membuat intepretasi terhadap data yang terkumpul serta membuat laporan. Selain itu, evaluator juga harus melakukan pengaturan terhadap evaluasi dan mengevaluasi apa yang telah dilakukan dalam melaksanakan evaluasi secara keseluruhan. Weiss (1972:4) menyatakan bahwa tujuan evaluasi adalah:

> *The purpose of evaluation research is to measure the effect of program against the goals it set out accomplish as a means of contributing to subsuquest decision making about the program and improving future programming.*

Ada empat hal yang ditekankan pada rumusan tersebut, yaitu: 1) menunjuk pada penggunaan metode penelitian, 2) menekankan pada hasil suatu program, 3) penggunaan kriteria untuk menilai, dan 4) kontribusi terhadap pengambilan keputusan dan perbaikan program di masa mendatang.

Berdasarkan pendapat di atas dapat disimpulkan bahwa evaluasi merupakan proses yang sistematis dan berkelanjutan untuk mengumpulkan, mendeskripsikan, mengintepretasikan dan menyajikan informasi untuk dapat digunakan sebagai dasar membuat keputusan, menyusun kebijakan maupun menyusun program selanjutnya. Adapun tujuan evaluasi adalah untuk memperoleh informasi yang akurat dan objektif tentang suatu program. Informasi tersebut dapat berupa proses pelaksanaan program,

dampak/hasil yang dicapai, efisiensi serta pemanfaatan hasil evaluasi yang difokuskan untuk program itu sendiri, yaitu untuk mengambil keputusan apakah dilanjutkan, diperbaiki atau dihentikan. Selain itu, juga dipergunakan untuk kepentingan penyusunan program berikutnya maupun penyusunan kebijakan yang terkait dengan program.

C. Model-model Evaluasi Program Pelatihan

Ada banyak model evaluasi yang dikembangkan oleh para ahli yang dapat dipakai dalam mengevaluasi program pelatihan. Kirkpatrick, salah seorang ahli evaluasi program training dalam bidang pengembangan SDM selain menawarkan model evaluasi yang diberi nama *Kirkpatrick's training evaluation model* juga menunjuk model-model lain yang dapat dijadikan sebagai pilihan dalam mengadakan evaluasi terhadap sebuah program training. Model-model yang ditunjuk tersebut di antaranya adalah :

- Five Level ROI Model (Jack PhillPS')
- CIPP Model (Daniel Stufflebeam's)
- Responsive Evaluation Model (Robert Stake's)
- Congruence-Contingency Model (Robert Stake's)
- Five Levels of Evaluation (Kaufman's)
- CIRO (Context, Input, Reaction, Outcome)
- PERT (Program Evaluation and Review Technique)
- Goal-Free Evaluation Approach (Michael Scriven's)
- Discrepancy Model (Provus's)
- Illuminative Evaluation Model
  (http://www.businessballs.com/kirkpatricklearningevaluationmodel.htm. diambil pada tanggal 23 Nopember 2005)

Dari berbagai model tersebut di atas dalam tulisn ini hanya akan diuraikan secara singkat beberapa model.

1. Evaluasi model CIPP

Konsep evaluasi model CIPP (*Context, Input, Prosess and Product*) pertama kali ditawarkan oleh Stufflebeam pada tahun 1965 sebagai hasil usahanya mengevaluasi ESEA (the Elementary and Secondary Education Act). Konsep tersebut ditawarkan oleh Stufflebeam dengan pandangan bahwa tujuan penting evaluasi adalah bukan membuktikan tetapi untuk memperbaiki. *The CIPP approach is based on the view that the most important purpose of evaluation is not to prove but to improve* (Madaus, Scriven, Stufflebeam, 1993: 118). Evaluasi model CIPP dapat diterapkan dalam berbagai bidang, seperti pendidikan, manajemen, perusahaan sebagainya serta dalam berbagai jenjang baik itu proyek, program

maupun institusi. Dalam bidang pendidikan Stufflebeam menggolongkan sistem pendidikan atas 4 dimensi, yaitu *context, input, process dan product*, sehingga model evaluasi yang ditawarkan diberi nama CIPP model yang merupakan singkatan ke empat dimensi tersebut. Nana Sudjana & Ibrahim (2004: 246) menterjemahkan masing-masing dimensi tersebut dengan makna sebagai berikut:

1). *Context* : situasi atau latar belakang yang mempengaruhi jenis-jenis tujuan dan strategi pendidikan yang akan dikembangkan dalam sistem yang bersangkutan, seperti misalnya masalah pendidikan yang dirasakan, keadaan ekonomi negara, pandangan hidup masyarakat .

2). *Input* : sarana/modal/bahan dan rencana strategi yang ditetapkan untuk mencapai tujuan-tujuan pendidikan.

3). *Process* : pelaksanaan strategi dan penggunaan sarana/modal/ bahan di dalam kegiatan nyata di lapangan.

4). *Product* : hasil yang dicapai baik selama maupun pada akhir pengembangan sistem pendidikan yang bersangkutan.

2. Evaluasi model Brinkerhoff

Setiap desain evaluasi pada umumnya terdiri dari elemen-elemen yang sama, ada banyak cara untuk menggabungkan elemen tersebut, masing-masing ahli evaluasi atau evaluator mempunyai konsep yang berbeda dalam hal ini. Brinkerhoff & CS mengemukakan tiga golongan evaluasi yang disusun berdasarkan penggabungan elemen-elemen yang sama, seperti evaluator-evaluator yang lain, namun dalam komposisi dan versi mereka sendiri sebagai berikut :

a. Fixed vs Emergent Evaluation Design

Desain evaluasi yang tetap (*fixed*) ditentukan dan direncanakan secara sistematik sebelum implementasi dikerjakan. Desain dikembangkan berdasarkan tujuan program disertai seperangkat pertanyaan yang akan dijawab dengan informasi yang akan diperoleh dari sumber-sumber tertentu. Rencana analisis dibuat sebelumnya di mana sipemakai akan menerima informasi seperti yang telah ditentukan dalam tujuan. Walaupun desain *fixed* ini lebih terstuktur daripada desain *emergent*, desain *fixed* juga dapat disesuaikan dengan kebutuhan yang mungkin berubah. Kebanyakan evaluasi formal yang dibuat secara individu dibuat berdasarkan desain *fixed*, karena tujuan program telah

ditentukan dengan jelas sebelumnya, dibiayai dan melalui usulan atau proposal evaluasi.

b. Formative vs Sumative Evaluation

Evaluasi formatif digunakan untuk memperoleh informasi yang dapat membantu memperbaiki program. Evaluasi formatif dilaksanakan pada saat implementasi program sedang berjalan. Fokus evaluasi berkisar pada kebutuhan yang dirumuskan oleh karyawan atau orang-orang program. Evaluator sering merupakan bagian dari pada program dan kerjasama dengan orang-orang program. Strategi pengumpulan informasi mungkin juga dipakai tetapi penekanan pada usaha memberikan informasi yang berguna secepatnya bagi perbaikan program.

Evaluasi sumatif dilaksanakan untuk menilai manfaat suatu program sehingga dari hasil evaluasi akan dapat ditentukan suatu program tertentu akan diteruskan atau dihentikan. Pada evaluasi sumatif difokuskan pada variable-variabel yang dianggap penting bagi sponsor program maupun fihak pembuat keputusan. Evaluator luar atau tim reviu sering dipakai karena evaluator internal dapat mempunyai kepentingan yang berbeda. waktu pelaksanaan evaluasi sumatif terletak pada akhir implementasi program. Strategi pengumpulan informasi akan memaksimalkan validitas eksternal dan internal yang mungkin dikumpulkan dalam waktu yang cukup lama.

c. Experimental and Quasi experimental Design vs Naural/Unotrusive

Beberapa evaluasi memakai metodologi penelitian klasik. Dalam hal seperti ini subyek penelitian diacak, perlakuan diberikan dan pengukuran dampak dilakukan. Tujuan dari penelitian untuk menilai manfaat suatu program yang dicobakan. Apabila siswa atau program dipilih secara acak, maka generalisasi dibuat pada populasi yang agak lebih luas. Dalam beberapa hal intervensi tidak mungkin dilakukan atau tidak dikehendaki. Apabila proses sudah diperbaiki, evaluator harus melihat dokumen-dokumen, seperti mempelajari nilai tes atau menganalisis penelitian yang dilakukan dan sebagainya. strategi pengumpulan data terutama menggunakan instrument formal seperti tes, suvey, kuesioner serta memakai metode penelitian yang terstandar.

3. Evaluasi model Kirkpatrick

Menurut Kirkpatrick evaluasi terhadap efektivitas program training mencakup empat level evaluasi, yaitu: level 1 – Reaction, level 2 – Learning, level 3 – Behavior, level 4 – Result

a. Evaluating Reaction

Mengevaluasi terhadap reaksi peserta training berarti mengukur kepuasan peserta (*customer satisfaction*). Program training dianggap efektif apabila proses training dirasa menyenangkan dan memuaskan bagi peserta training sehingga mereka tertarik termotivasi untuk belajar dan berlatih. Dengan kata lain peserta training akan termotivasi apabila proses training berjalan secara memuaskan bagi peserta yang pada akhirnya akan memunculkan reaksi dari peserta yang menyenangkan. Sebaliknya apabila peserta tidak merasa puas terhadap proses training yang diikutinya maka mereka tidak akan termotivasi untuk mengikuti training lebih lanjut. Dengan demikian dapat dimaknai bahwa keberhasilan proses kegiatan training tidak terlepas dari minat, perhatian dan motivasi peserta training dalam mengikuti jalannya kegiatan training. Orang akan belajar lebih baik manakala mereka memberi reaksi positif terhadap lingkungan belajar.

Kepuasan peserta training dapat dikaji dari beberapa aspek, yaitu materi yang diberikan, fasilitas yang tersedia, strategi penyampaian materi yang digunakan oleh instruktur, media pembelajaran yang tersedia, jadwal kegiatan sampai menu dan penyajian konsumsi yang disediakan.

b. Evaluating Learning

Menurut Kirkpatrick (1988: 20) *learning can be defined as the extend to which participans change attitudes, improving knowledge, and/or increase skill as a result of attending the program.* Ada tiga hal yang dapat instruktur ajarkan dalam program training, yaitu pengetahuan, sikap maupun ketrampilan. Peserta training dikatakan telah belajar apabila pada dirinya telah mengalamai perubahan sikap, perbaikan pengetahuan maupun peningkatan ketrampilan. Oleh karena itu untuk mengukur efektivitas program training maka ketiga aspek tersebut perlu untuk diukur. Tanpa adanya perubahan sikap, peningkatan pengetahuan maupun perbaikan ketrampilan pada peserta training maka program dapat dikatakan gagal. Penilaian *evaluating learning* ini ada yang

menyebut dengan penilaiah hasil (*output*) belajar. Oleh karena itu dalam pengukuran hasil belajar (*learning measurement*) berarti penentuan satu atau lebih hal berikut: a). Pengetahuan apa yang telah dipelajari ?, b). Sikap apa yang telah berubah ?, c). Ketrampilan apa yang telah dikembangkan atau diperbaiki ?.

c. Evaluating Behavior

Evaluasi pada level ke 3 (evaluasi tingkah laku) ini berbeda dengan evaluasi terhadap sikap pada level ke 2. Penilaian sikap pada evaluasi level 2 difokuskan pada perubahan sikap yang terjadi pada saat kegiatan training dilakukan sehingga lebih bersifat internal, sedangkan penilaian tingkah laku difokuskan pada perubahan tingkah laku setelah peserta kembali ke tempat kerja. Apakah perubahan sikap yang telah terjadi setelah mengikuti training juga akan diimplementasikan setelah peserta kembali ke tempat kerja, sehingga penilaian tingkah laku ini lebih bersifat eksternal. Perubahan perilaku apa yang terjadi di tempat kerja setelah peserta mengikuti program training. Dengan kata lain yang perlu dinilai adalah apakah peserta merasa senang setelah mengikuti training dan kembali ke tempat kerja?. Bagaimana peserta dapat mentrasfer pengetahuan, sikap dan ketrampilan yang diperoleh selama training untuk diimplementasikan di tempat kerjanya. Karena yang dinilai adalah perubahan perilaku setelah kembali ke tempat kerja maka evaluasi level 3 ini dapat disebut sebagai evaluasi terhadap *outcomes* dari kegiatan training.

d. Evaluating Result

Evaluasi hasil dalam level ke 4 ini difokuskan pada hasil akhir (*final result*) yang terjadi karena peserta telah mengikuti suatu program. Termasuk dalam kategori hasil akhir dari suatu program training di antaranya adalah kenaikan produksi, peningkatan kualitas, penurunan biaya, penurunan kuantitas terjadinya kecelakaan kerja, penurunan *turnover* dan kenaikan keuntungan. Beberapa program mempunyai tujuan meningkatkan moral kerja maupun membangun teamwork yang lebih baik. Dengan kata lain adalah evaluasi terhadap *impact* program.

4. Evaluasi model Stake (Model Countenance)

Stake menekankan adanya dua dasar kegiatan dalam evaluasi, yaitu *description* dan *judgement* dan membedakan adanya tiga tahap dalam program pelatihan, yaitu *antecedent (context), transaction (process*) dan *outcomes*. Stake

mengatakan bahwa apabila kita menilai suatu program pelatihan, kita melakukan perbandingan yang relatif antara program dengan program yang lain, atau perbandingan yang absolut yaitu membandingkan suatu program dengan standar tertentu. Penekanan yang umum atau hal yang penting dalam model ini adalah bahwa evaluator yang membuat penilaian tentang program yang dievaluasi. Stake mengatakan bahwa *description* di satu pihak berbeda dengan *judgement* di lain fihak. Dalam model ini *antecendent* (masukan) *transaction* (proses) dan *outcomes* (hasil) data di bandingkan tidak hanya untuk menentukan apakah ada perbedaan antara tujuan dengan keadaan yang sebenarnya, tetapi juga dibandingkan dengan standar yang absolut untuk menilai manfaat program (Farida Yusuf Tayibnapis, 2000: 22).

D. Langkah-langkah Evaluasi Program Pelatihan

Dalam mengadakan evaluasi terhadap program pelatihan secara sistematis pada umumnya menempuh 4 langkah, yaitu: 1). penyusunan desain evaluasi, 2). pengembangan instrument pengumpulan data, 3). pengumpulan data (*assessment*), menafsirkan dan membuat judgement, serta 4). menyusun laporan hasil evaluasi (Purwanto dan Atwi Suparman, 1999: 73).

1. Menyusunan Desain Evaluasi

   Langkah pertama dalam evaluasi adalah menyusun rencana evaluasi yang menghasilkan desain evaluasi. Pada langkah ini evaluator mempersiapkan segala sesuatu sesuatu yang berkaitan dengan pelaksanaan evaluasi, mulai menentukan tujuan evaluasi, model yang akan digunakan, informasi yang akan dicari serta metode pengumpulan dan analaisis data. Apabila langkah pertama dapat menghasilkan desin evaluasi yang cukup komprehensif dan rinci, maka sudah dapat dijadikan sebagai acuan kegiatan evaluasi yang akan dilaksanakan. Rancangan atau desain evaluasi biasanya disusun oleh evaluator setelah melakukan diskusi dan ada kesepakatan dengan pihak yang akan membiayai kegiatan evaluasi atau sponsor. Namun adakalanya rancaranga disusun oleh evaluator untuk dijadikan bahan mengadakan negoisasi dengan sponsor.

2. Mengembangkan Instrumen Pengumpulan Data

   Setelah metode pengumpulan data ditentukan, langkah selanjutnya adalah menentukan bentuk instrument yang akan digunakan serta kepada siapa instrumen tersebut ditujukan (responden). Kemudian setelah itu perlu dikembangkan butir-

butir dalam instrument. Berbagai pertimbangan mengenai berapa banyak informasi yang akan dikumpulkan, instrument dikembangkan sendiri, mengadopsi ataupun menggunakan instrument baku dari instrument yang sudah ada sebelumnya. Untuk memperoleh data yang valid maka instrument yang digunakan harus memperhatikan masalah validitas dan reliabilitas. Selain hal tersebut, masalah efisiensi dan efektivitas harus tetap diperhatikan. Jenis-jenis instrument yang paling sering digunakan untuk mengumpulkan data dalam evaluasi program pelatihan adalah dalam bentuk tes, angket,ceklis pengamatan, wawancara atau evaluator sendiri sebagai instrument.

3. Mengumpulkan Data, Analisis dan Judgement

   Langkah ketiga merupakan tahapan pelaksanaan dari apa yang telah dirancang pada langkah pertama dan kedua. Pada langkah ketiga ini evaluator terjun ke "lapangan" untuk mengimplementasikan desain yang telah dibuat, mulai dari mengumpulkan dan menganalisis data, menginterpretasikan, dan menyajikan dalam bentuk yang mudah untuk dipahami dan komunikatif. Pengumpulan data dapat dari populasi maupun dengan menggunakan sampel. Apabila menggunakan sampel maka harus representatif mewakili populasi, oleh karena harus memperhatikan tekhnik sampling yang baik. Berdasarkan data yang dikumpulkan kemudian dianalisis dan dibuat judgement berdasarkan kriteria maupun standar yang telah ditetapkan sebelumnya. Dari hasil judgement kemudian disusun rekomendasi kepada penyelenggara kegiatan pelatihan maupun fihak-fihak lain yang mempunyai kepentingan dengan kegiatan pelatihan.

   Langkah ketiga ini merupakan proses esensial dari kegiatan evaluasi program pelatihan di mana terjadi dialog antara evaluator dengan obyek evaluasi. Hal yang harus diperhatikan oleh evaluator pada tahap ini adalah masalah etika dan penguasaan 'setting' atau latar di mana evaluasi dilaksanakan.

4. Menyusun Laporan Hasil Evaluasi

   Menyusun laporan merupakan langkah terakhir kegiatan evaluasi program pelatihan. Laporan disusun sesuai dengan kesepakatan kontrak yang ditandatangani. Misalnya dalam kontrak disepakati bahwa laporan dibuat dua jenis laporan dengan sasaran atau penerima laporan yang berbeda. dapat disepakati pula bahwa penyampaian laporan secara tertulis dan ada kesempatan presentasi. Langkah

terakhir ini erat kaitannya dengan tujuan diadakannya evaluasi. Oleh karena itu gaya dan format penyampaian laporan harus disesuaikan dengan penerima laporan.

E. Evaluator Program Pelatihan

Keberhasilan kegiatan evaluasi program pelatihan akan sangat ditentukan oleh siapa yang melakukan evaluasi atau evaluator. Evaluator adalah orang yang dipercaya oleh pemilik program dan orang-orang yang berkepentingan dengan program (*stakeholder*) untuk melaksanakan evaluasi. Penentuan siapa yang akan menjadi evaluator ini sangat tergantung kepada pemilik program. Berikut ini disajikan hal-hal yang berkaitan dengan alternatif evaluator, pertimbangan penentuan, dan kompetensi evaluator.

1. Alternatif Penentuan Evaluator

   Penentuan tentang siapa yang akan berperan sebagai evaluator sangat penting dan menentukan dalam kegiatan evaluasi. Membuat keputusan tentang siapa yang akan mengambil bagian sebagai evaluator terkadang mengandung konflik pilihan yang dilematis. Pertanyaan-pertanyaan yang muncul ketika memikirkan siapa yang akan berperan sebagai evaluato adalah: Apakah evaluator berasal dari dalam atau dari luar organisasi?. Apakah evaluator merupakan sebuah tim atau individu?. Apakah evaluator merupakan tenaga paruh waktu (*part-time*) atau bekerja penuh (*full-time*)?. Apakah evaluator merupakan tenaga professional atau amatir?. Jawaban atas setiap pertanyaan tersebut mengarah kepada pemilihan dan penentuan evaluator.

2. Pertimbangan dalam Penentuan Evaluator

   Berikut ini berbagai pertimbangan yang dapat dijadikan pedoman sebelum menentukan evaluator. Pertimbangan ini berkaitan dengan masalah keuntungan dan kerugian atau kelebihan dan kekurangannya.

   a. Pertimbangan antara evaluator orang dalam dan orang luar

      Orang dalam adalah orang yang berasal dari bagian atau institusi penyelenggaran program pelatihan, dan biasanya mereka telah ikut dalam proses pengembangan dan pelaksanaan program pelatihan. Sedangkan yang dimaksud dengan orang luar adalah mereka yang berperan sebagai evaluator berasal dari luar bagian atau institusi penyelenggara program pelatihan. Kelebihan orang

dalam adalah mereka sudah mengetahui organisasi dengan baik, dapat mengetahui reputasi, status dan kredibilitas organisasi tempatnya bekerja. Ia memiliki hubungan yang baik dengan staf, memahami saluran komunikasi dalam organisasi. Kelemahan orang dari dalam adalah terjadinya bias, karena konflik kepentingan, mungkin evaluator tidak memiliki ketrampilan evaluasi, atau pekerjaan evaluasi yang dilaksanakannya terganggu oleh tugas lain dan akibatnya ia tidak dapat menepati waktu. Sebaliknya apabila evaluator ditentukan dari orang luar, maka kelebihannya mereka dapat bersikap netral, dapat bertindak sebagai pengamat independent, obyektif sebagai pengamat, dan lebih kompeten dalam teknik evaluasi. Kekurangan evaluator dari luar adalah mereka kurang akrab dengan kebiasaan organisasi, tidak mengenal tatacara yang ada di organisasi yang dimasuki, dan tidak menutup kemungkinan pemilihannya hanya berdasarkan rekomendasi.

b. Pertimbangan antara evaluator tim dan individual

Kelebihan evaluator individual atau perorangan adalah adanya kejelasan tentang siapa yang harus bertanggungjawab, sedangkan kekurangannya adalah keberhasilan atau kegagalan evaluasi tergantung pada satu orang. Sebenarnya hampir mustahil pekerjaan evaluasi program pelatihan hanya diselesaikan oleh satu orang tanpa bantuan orang lain. Apabila evaluator ditentukan oleh tim maka kelebihannya adalah adanya pembagian tugas dan tanggung jawab yang jelas. Evaluator terdiri atas gabungan orang dengan berbagai keahlian sehingga bisa saling melengkapi. Kelemahan evaluator tim adalah perlu waktu untuk pembentukan tim, membutuhkan biaya yang tidak sedikit.

c. Pertimbangan antara evaluator parti-time dan full-time

Evaluator parti-time dan full-time masing-masing memiliki kelebihan dan kelemahan. Kelebihan apabila evaluator bekerja penuh waktu (*full-time*) adalah pekerjaan terorganisir dengan baik, ketepatan dan arus informasi tidak tergantung pada evaluator, sedangkan kelemahannya adalah biaya relatif lebih mahal, mengurangi kesempatan partisipasi dalam kegiatan evaluasi. Apabila evaluator bekerja paruh waktu (*part-time*) adalah dapat melibatkan berbagai keahlian dalam waktu yang tidak terlalu lama dan dimungkinkan penggunaan tenaga ahli dari luar, sementara kelemahannya adalah waktu kunjungan singkat

tidak memungkinkan untuk mempelajari permasalahan secara menyeluruah dan perlu biaya dan peralatan yang banyak untuk penjadwalan.

d. Pertimbangan antara evaluator amatir dan professional

Pengertian evaluator professional di sini adalah mereka yang menjadikan pekerjaan evaluasi atau penelitian sebagai pekerjaan pokok sehari-hari dan telah menekuni pekerjaan evaluasi dalam waktu yang lama. Orang-orang di luar kriteria tersebut dianggap sebagai amatir. Kelebihan evaluator amatir, terutama yang sudah berpengalaman, meskipun amatir evaluator biasanya dapat memahami isi dan obyek evaluasi dengan baik dan dapat memilih berbagai ketrampilan evaluasi berdasarkan pengalaman. Kelemahan evaluator amatir adalah karena kurangnya pengetahuan tentang evaluasi akibatnya dapat menurunkan obyektivitas evaluasi, kemampuan evaluasinya terbatas dan mereka memiliki keterbatasan dalam pilihan rancangan evaluasi. Kelebihan evaluator professional adalah evaluator dapat melaksanakan evaluasinya berdasarkan pengalaman dan pengetahuan teknis dan evaluator memiliki berbagai pilihan model evaluasi berdasarkan pengetahuan maupun pengalamannya. Kelemahan evaluator professional (biasanya orang luar), tidak selamanya dapat diterima oleh orang dalam, kecenderungan menggunakan metode tertentu, dan menghalangi pemilihan metode atau rancangan orang lain.

3. Kompetensi Evaluator

Evaluator haruslah dipilih dari orang yang benar-benar memiliki kompetensi di bidangnya. Ketidakbebasan dalam penentuan evaluator harus dihindari, sebab hal itu akan berpengaruh negatif terhadap hasil evaluasi. Ketidakbebasan karena konflik kepentingan atau *conflict of interest* lebih besar pengaruhnya terhadap hasil ketimbang ketidakmampuan dalam bidang teknis. Kompetensi evaluator dapat dikelompokkan menjadi empat jenis, yaitu: kompetensi manajemerial, kompetensi teknis, kompetensi koseptual dan kompetensi bidang studi (Purwanto dan Atwi Suparman, 1999: 55).

a. Kompetensi Manajerial. Kompetensi manajerial (*magerial skill*) merupakan ketrampilan dalam mengelola dan mengendalikan seluruh kegiatan evaluasi sehingga dapat berlangsung dengan baik. Ketrampilan manajerial ini meliputi: ketrampilan mengorganisir, memimpin, mengkoordinir, mengarahkan,

mengawasi, ketrampilan berkomunikasi, ketrampilan interpersonal, analsis system, membuat perjanjian atau kontrak, menjelaskan wawasan politik, ketrampilan menerapkan etika profesi dan sebagainya.

b. Ketrampilan Teknis. Kompetensi/ketrampilan Teknis yakni ketrampilan melakukan kegiatan evaluasi langkah demi langkah, dari perencanaan sampai pembuatan laporan evaluasi secara tuntas. Termasuk ketrampilan teknis ini di antaranya adalah: ketrampilan mengembangkan instrument, melaksanakan tes dan pengukuran, melakukan analisis statistik, menguasai berbagi metode pengumpulan data, menguasai aplikasi komputer, menguasai berbagai *soft-ware* seperti excel, SPSS, Amos, Lisrel dan berbagai *soft-ware* dalam bidang statistik lainnya, menerapkan metodolgi penelitian evaluasi, membuat interpretasi, membuat rekomendasi dan menulis laporan serta mempresentasikan laporan.
c. Kompetensi Konseptual, yaitu ketrampilan tingkat tinggi yang berkaitan dengan kemampuan menganalisis dan pemecahan masalah. Ketrampilan konspetual (*conceptual skill*) yang harus dikuasasi evaluator di antaranya adalah kemampuan menentukan pilihan (alternative), menyusun rencana awal, mengklasifikasikan dan menganalisis masalah, melihat dan menunjukkan hubungan antar variabel dan membuat kesimpulan.
d. Kompetensi bidang studi, yaitu kemampuan di bidang disiplin ilmu yang terkait dengan kegiatan evaluasi. Keahlian ini meliputi: pengalaman kerja di bidang evaluasi, berpengatahuan tentang sumber literatur yang berkaitan dengan obyek yang dievaluasi, menguasai konsep-konsep maupun model- model evaluasi.

F. Penutup

Kegiatan penilaian dalam evaluasi program pelatihan tidak hanya dilaksanakan pada akhir kegiatan program, tetapi sebaiknya dilakukan sejak awal, yaitu dari penyusunan rancangnan program pelatihan, pelaksanaan program pelatahan dan hasil dari pelatihan. Penilaian hasil pelatihan tidak cukup hanya pada hasil jangka pendek (*output*) tetapi dapat menjangkau hasil dalam jangka panjang (*outcome and impact program*). Ada berbagai macam model evaluasi program yang dapat dipilih untuk mengevaluasi program pelatihan. Model mana yang akan digunakan tergantung pada tujuan maupun kemampuan evaluator.

Evaluator dalam kegiatan evaluasi program pelatihan dapat dari orang dalam maupun dari orang luar, dapat bersifat indidual maupun tim, dari tenaga paruh waktu maupun penuh waktu, orang professional maupun tenaga amatir. Siapapun yang ditunjuk menjadi evaluator, agar hasil evaluasi dapat maksimal maka kompetensi evaluator harus dipertimbangkan. Kompetensi evaluator meliputi kompetensi manajerial, kompetensi teknis, kompetensi konseptual dan kompetensi bidang studi.

Daftar Pustaka :

- Djemari Mardapi. (1999). *Pengukuran, penilaian dan evaluasi*. Makalah disampaikan pada Penataran evaluasi pembelajaran matematika SLTP untuk guru inti matematika di MGMP SLTP tanggal 8 – 23 Nopember 1999 di PPPG Matematika Yogyakarta.

- Djemari Mardapi. (2000). *Evaluasi pendidikan*. Makalah disampaikan pada Konvensi Pendidikan Nasional tanggal 19 – 23 September 2000 di Universitas Negeri Jakarta.

- Ebel, R.L. & Frisbie, D.A. (1986). *Essential of educational measurement*. New Jerseey: Prentice- Hall, Inc.

- Farida Yusuf Tayibnapis. (2000). *Evaluasi program*. Jakarta: Rineka Cipta

- Griffin, P. & Nix, P. (1991). *Educational assessment and reporting*. Sydney: Harcout Brace Javanovich, Publisher.

- Kirkpatrick, D.L. (1998). *Evaluating training programs, The four levels*, *Second edition*. San Francisco: Berrett-Koehler Publisher, Inc.

- Kirkpatrick, D.L. (2005). *Kirkpatrick's training evaluation model.* Diambil pada tanggal 23 Sepember 2005, dari http://www. businessballs. com/ Kirkpatrick learningevaluationmodel.htm.

- Madaus, G.F. & Scriven, M.S. & Stuffebeam, D.L. (1993). *Evaluation models, viewpoints on educational and human services evaluation*. Boston: Kluwer-Nijhoff Publishing.

- Nana Sudjana & Ibrahim. (2004). *Penelitian dan penilaian pendidikan*. Bandung: Sinar Baru Algesindo.

- Naugle. (2005) . *Kirkpatrick's evaluation model as a means of evaluating teacher performance*. Diambil pada tanggal 15 September 2005, dari http://www. findarticles.com/p/articles.

- Oriondo, L. L. & Antonio, E. M.D. (1998). *Evaluating educational outcomes (Test, measurment and evaluation)*. Florentino St: Rex Printing Company, Inc.

- Purwanto dan Atwi Suparman, (1999). *Evaluasi program diklat*. Jakarta: Sekolah Tinggi Ilmu Administrasi, Lembaga Administrasi Negara.

- Popham, W. J. (1995). *Classroom assessment*. Boston: Allyn and Bacon.

- Stark, J.S. & Thomas, A. (1994). *Assessment and program evaluation*. Needham Heights: Simon & Schuster Custom Publishing.

- Stufflebeam, D.L. & Shinkfield, A.J. (1985). *Systematic evaluation*. Boston: Kluwer Nijhof Publishing.