



## Pengantar Evaluasi Program PENDIDIKAN

Evaluasi program dimaksudkan untuk memperoleh informasi yang akurat dan obyektif 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.

Buku ini memberikan bekal keilmuan dan kemampuan kepada mahasiswa khususnya dan praktisi pendidikan umumnya dalam rangka melakukan evaluasi program pendidikan. Untuk memenuhi keinginan tersebut maka di dalam buku ini dilengkapi materi-materi pembahasan terkait dengan model-model evaluasi program, perencanaan evaluasi program, instrumen evaluasi program, analisis dan interpretasi data dan dilengkapi juga pelaporan evaluasi program.

ISBN 978-602-6462-54-1



Dr. Rusydi Ananda, M.Pd  
Dr. Tien Rafida, M.Hum

Pengantar

## Evaluasi Program PENDIDIKAN

Editor:  
Dr. Candra Wijaya, M.Pd

**Perdana**  
Publishing

PENERBIT BUKU UMUM & PERGURUAN TINGGI  
Jl. Sosro No.16A Medan 20224, Tel 061-77151020  
Fax 071-7347756 Email: perdanapublishing@gmail.com

Dr. Rusydi Ananda, M.Pd  
Dr. Tien Rafida, M.Hum

Pengantar **Evaluasi Program Pendidikan**

**Perdana**  
Publishing

**Perdana**  
Publishing

# **PENGANTAR EVALUASI PROGRAM PENDIDIKAN**



# **PENGANTAR EVALUASI PROGRAM PENDIDIKAN**

**Dr. Rusydi Ananda, M.Pd**  
**Dr. Tien Rafida, M.Hum**

**Dr. Candra Wijaya, M.Pd., (Ed)**



Kelompok Penerbit Perdana Mulya Sarana

## **PENGANTAR EVALUASI PROGRAM PENDIDIKAN**

Penulis: Dr. Rusydi Ananda, M.Pd, dan  
Dr. Tien Rafida, M.Hum

Editor: Dr. Candra Wijaya, M.Pd

Copyright © 2017, pada penulis  
Hak cipta dilindungi undang-undang  
All rights reserved

Penata letak: Muhammad Yunus Nasution  
Perancang sampul: Aulia Grafika

Diterbitkan oleh:

**PERDANA PUBLISHING**

Kelompok Penerbit Perdana Mulya Sarana  
(ANGGOTA IKAPI No. 022/SUT/11)  
Jl. Sosro No. 16-A Medan 20224  
Telp. 061-77151020, 7347756 Faks. 061-7347756  
E-mail: [perdanapublishing@gmail.com](mailto:perdanapublishing@gmail.com)  
Contact person: 08126516306

Cetakan pertama: Februari 2017

**ISBN 978-602-6462-54-1**

Dilarang memperbanyak, menyalin, merekam sebagian  
atau seluruh bagian buku ini dalam bahasa atau  
bentuk apapun tanpa izin tertulis dari penerbit atau penulis

**BUKU INI DI DEDIKASIKAN KEPADA  
(ALM) AYAHANDA H. THAHARUDDIN, AG  
(ALM) AYAHANDA H. ARIFIN  
(ALM) IBUNDA ZUBAIDAH  
(ALM) IBUNDA HABIBAH**

Mereka mungkin bisa lupa  
Apa yang Anda katakan  
Tapi mereka takkan pernah melupakan  
Perasaan yang Anda timbulkan  
Dalam hati mereka

# KATA PENGANTAR

*Alhamdulillah*, puji syukur kehadiran Allah SWT, dengan rahmat dan izinNya, buku Pengantar Evaluasi Program Pendidikan sebagai menunjang kegiatan pembelajaran di Fakultas Ilmu Tarbiyah dan Keguruan khususnya pada program studi Manajemen Pendidikan Islam dapat diwujudkan.

Matakuliah Evaluasi Program Pendidikan merupakan salah satu matakuliah wajib bagi mahasiswa program studi Manajemen Pendidikan Islam. Matakuliah ini memberikan bekal keilmuan dan kemampuan kepada mahasiswa dalam rangka melakukan evaluasi suatu program khususnya program di bidang Pendidikan dan Pembelajaran.

Melalui buku Pengantar Evaluasi Program Pendidikan ini diharapkan dapat melengkapi buku-buku yang sudah ada, sekaligus sebagai bahan bacaan dan penambahan wawasan bagi mahasiswa maupun pembaca lainnya.

Penulis menyadari bahwa penulisan buku ini jauh dari kesempurnaan, oleh karena itu penyempurnaan buku Pengantar Evaluasi Program Pendidikan akan dilakukan seiring dengan perkembangan dan respon dari para pembaca.

Penulis

H. Rusydi Ananda  
Hj. Tien Rafida

# DAFTAR ISI

	Hal
KATA PENGANTAR .....	v
DAFTAR ISI .....	vii
<b>BAB I      PENDAHULUAN .....</b>	<b>1</b>
A. Pengertian .....	1
B. Tujuan dan Manfaat Evaluasi Program .....	7
C. Asumsi Dasar dan Prinsip Umum .....	10
D. Karakteristik Evaluasi Program .....	13
E. Komponen dan Indikator Program .....	14
F. Kriteria Dalam Evaluasi Program .....	18
<b>BAB II     EVALUATOR PROGRAM .....</b>	<b>23</b>
A. Pengertian .....	23
B. Jenis Evaluator .....	23
C. Pertimbangan Dalam Penentuan Evaluator .....	27
D. Peranan Evaluator .....	30
E. Syarat-Syarat Evaluator .....	31
F. Kompetensi Evaluator Program .....	34
<b>BAB III    MODEL-MODEL EVALUASI PROGRAM .....</b>	<b>36</b>
A. Klasifikasi Model Evaluasi Program .....	36
B. Model-Model Evaluasi Program .....	42
<b>BAB IV    PERENCANAAN EVALUASI PROGRAM .....</b>	<b>74</b>
A. Analisis Kebutuhan .....	74
B. <i>Scheduling</i> (Penjadwalan) .....	77
C. Penugasan dan Monitoring .....	78
D. <i>Budgeting</i> (Pembiayaan) .....	79
E. Proposal Evaluasi Program .....	80



<b>BAB V</b>	<b>INSTRUMEN EVALUASI PROGRAM .....</b>	<b>102</b>
	A. Sumber Data .....	102
	B. Instrumen Pengumpulan Data .....	104
	C. Penyusunan Instrumen .....	117
	D. Validitas dan Reliabilitas Instrumen .....	122
<b>BAB VI</b>	<b>ANALISIS DAN INTERPRESTASI DATA.....</b>	<b>139</b>
	A. Jenis Data .....	139
	B. Prinsip dan Kriteria Menganalisis Data .....	143
	C. Mengorganisasi Data .....	143
	D. Prosedur Analisis Data.....	146
	E. Pengolahan Data Wawancara .....	162
	F. Interpretasi Data.....	163
<b>BAB VII</b>	<b>KESIMPULAN DAN REKOMENDASI EVALUASI PROGRAM .....</b>	<b>166</b>
	A. Kesimpulan .....	166
	B. Rekomendasi.....	169
<b>BAB VIII</b>	<b>PELAPORAN DAN PEMANFAATAN LAPORAN EVALUASI PROGRAM .....</b>	<b>174</b>
	A. Pengertian .....	174
	B. Tujuan dan Manfaat Laporan Evaluasi.....	174
	C. Pendekatan dan Jenis Laporan.....	176
	D. Faktor-Faktor Dalam Perencanaan Laporan .....	183
	E. Format Laporan Evaluasi .....	186
	F. Kemenarikan Penyajian Laporan .....	204
	G. Pemanfaatan Laporan Evaluasi Program.....	208
	DAFTAR BACAAN .....	211
	LAMPIRAN 1. GLOSSARIUM .....	214
	TENTANG PENULIS .....	218

# BAB I

## PENDAHULUAN

### A. Pengertian

#### 1. Evaluasi

Evaluasi berasal dari kata “*evaluation*” (bahasa Inggris), kata tersebut diserap ke dalam perbendaharaan dalam bahasa Indonesia dengan tujuan mempertahankan kata aslinya dengan penyesuaian lafal Indonesia (Arikunto dan Jabar, 2009:1). Selanjutnya dijelaskan keduanya bahwa evaluasi adalah kegiatan untuk mengumpulkan informasi tentang bekerjanya sesuatu, yang selanjutnya informasi tersebut digunakan untuk menentukan alternatif yang tepat dalam mengambil keputusan.

Scriven dalam Stufflebeam dan Shinkfield (2007:369) mendefinisikan evaluasi yaitu: *evaluation is the process of determining the merit, worth, and value of things and evaluation are the products of the process*. Evaluasi adalah suatu proses menentukan manfaat, harga, dan nilai dari sesuatu dan evaluasi adalah produk dari proses tersebut. Dengan kata lain evaluasi adalah produk dari proses menentukan manfaat dan nilai dari sesuatu. Produk itu berbentuk temuan-temuan yang ditulis dalam bentuk laporan.

Kifer (1995:384) mendefinisikan evaluasi sebagai penyelidikan untuk menentukan nilai atau manfaat (*worth*) suatu program, produk, prosedur atau proyek. Selanjutnya Madaus dkk (1987:24) memaparkan evaluasi adalah studi yang dirancang dan dilaksanakan untuk menilai (*judge*) dan meningkatkan manfaat program yang dievaluasi.

Stufflebeam dan Shinkfield (2007:326) menyatakan bahwa: *evaluation is a systematic investigation of some object's value*. Evaluasi adalah suatu investigasi, penelitian, penyelidikan, atau pemeriksaan yang sistematis terhadap nilai suatu objek. Secara operasional Stufflebeam dan Shinkfield

(2007:326) memaparkan evaluasi adalah proses merencanakan, memperoleh, melaporkan, dan menggunakan informasi deskriptif dan mempertimbangkan beberapa manfaat objek, nilai signifikansi dan kejujuran dalam rangka memandu pengambilan keputusan, akuntabilitas, dukungan, menyebarkan praktek-praktek yang efektif serta meningkatkan pemahaman tentang fenomena-fenomena yang terlibat.

Mehren dan Lehmann (1978:5) menjelaskan evaluasi adalah suatu proses merencanakan, memperoleh dan menyediakan informasi yang sangat diperlukan untuk membuat alternatif-alternatif keputusan. Pengertian yang dikemukakan keduanya menunjukkan bahwa evaluasi itu merupakan suatu proses yang sengaja direncanakan untuk memperoleh informasi atau data dan berdasarkan informasi atau data tersebut dibuat suatu keputusan.

Selanjutnya menurut Alkin (1985:11) evaluasi adalah suatu aktivitas sistematis untuk mengumpulkan, menganalisis dan melaporkan informasi yang dapat digunakan untuk mengambil keputusan berkenaan dengan program atau proyek yang dievaluasi. Selanjutnya Guba dan Lincoln (1985:35) memaparkan evaluasi adalah: *a process for describing an evaluation and judging its merit and worth*. Evaluasi adalah proses atau kegiatan untuk menentukan manfaat nilai sesuatu.

The Joint Committee sebagaimana dikutip Stufflebeam dan Shinkfield (2007:9) menyatakan *evaluation is the systematic assessment of the worth or merit of an object*. Evaluasi adalah penilaian yang sistematis tentang nilai, harga atau manfaat dari suatu objek. Sistematis di sini menunjukkan bahwa evaluasi harus dilakukan secara resmi atau formal dan sistematis, bukan dilakukan sekedar formalitas dan asal-asalan.

National Study Committee on Evaluation 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* (Stark dan Thomas, 1994:12). Evaluasi merupakan suatu proses atau kegiatan pemilihan, pengumpulan, analisis, dan penyajian informasi yang dapat digunakan sebagai dasar pengambilan keputusan serta penyusunan program selanjutnya.

Briekerhoff dkk (1986:ix) menyatakan bahwa evaluasi merupakan proses yang menentukan sejauhmana tujuan pendidikan dapat dicapai dan dalam pelaksanaannya evaluasi tersebut fokus pada tujuh elemen yang harus diperhatikan yaitu:

- a. Penentuan fokus yang akan di evaluasi.
- b. Penyusunan desain evaluasi.
- c. Pengumpulan informasi.
- d. Analisis dan interpretasi informasi.
- e. Pembuatan laporan.
- f. Pengelolaan evaluasi.
- g. Evaluasi untuk evaluasi atau meta evaluasi.

Purwanto dan Suparman (1999:9) mendeskripsikan evaluasi adalah proses penerapan prosedur ilmiah untuk mengumpulkan informasi yang valid dan reliabel untuk membuat keputusan tentang program pendidikan dan pelatihan. Berdasarkan definisi tersebut ditemukan empat unsur pokok dalam evaluasi yaitu:

- a. Evaluasi selalu menerapkan suatu metode ilmiah baik berupa pengukuran ilmiah melalui penggunaan statistika maupun disiplin lain yang terkait.
- b. Kegiatan evaluasi selalu berusaha memperoleh informasi yang benar-benar valid dan reliabel dengan mempergunakan instrumen berupa tes, kuesioner, pedoman wawancara, pedoman pengamatan dan lain-lain.
- c. Hasil evaluasi adalah suatu informasi yang dapat berguna bagi pembuatan keputusan.
- d. Kegiatan evaluasi selalu diarahkan kepada suatu objek yang ada dalam suatu sistem pendidikan atau sistem pelatihan.

Menurut Djaali dan Muljono (2004:1) evaluasi adalah suatu proses menilai sesuatu berdasarkan kriteria atau tujuan yang telah ditetapkan, yang selanjutnya diikuti dengan pengambilan keputusan atas objek yang di evaluasi. Selanjutnya menurut Mutrofin (2010:33) evaluasi adalah suatu kegiatan sistematis yang dilaksanakan untuk membantu audiensi agar dapat mempertimbangkan dan meningkatkan nilai suatu program atau kegiatan.

Sudjana (2008:9) memaknai evaluasi sebagai kegiatan mengumpulkan, mengolah dan menyajikan data untuk masukan dalam pengambilan keputusan mengenai program yang sedang dan/atau telah dilaksanakan. Produk evaluasi adalah tersusunnya nilai-nilai (*values*) seperti bermanfaat atau tidak bermanfaat, baik atau buruk, berhasil atau tidak berhasil, diperluas

atau dibatasi, dilanjutkan atau dihentikan, dan sebagainya, mengenai program yang sedang atau telah dilaksanakan.

Definisi-definisi terkait dengan evaluasi yang dikemukakan para ahli maka Mutrofin (2010:82) merangkum bahwa untuk mendeskripsikan evaluasi sebagai kerangka umum di dalamnya terdapat makna-makna sebagai berikut:

- a. Evaluasi sebagai *judgement professional*.
- b. Evaluasi sebagai pengukuran.
- c. Evaluasi sebagai analisis kesesuaian antara kinerja dengan tujuan, sasaran atau standar kerja.
- d. Evaluasi berorientasi pada keputusan.
- e. Evaluasi responsif atau bebas tujuan.

Senada dengan penjelasan yang dikemukakan Mutrofin di atas, maka Purwanto (2001:3-4) memaparkan bahwa dalam makna evaluasi itu terkandung 3 (tiga) aspek yang menjadi titik tekan. Ketiga titik tekan dalam makna evaluasi itu sebagai berikut:

- a. Kegiatan evaluasi merupakan proses yang sistematis. Ini berarti bahwa evaluasi merupakan kegiatan yang terencana dan dilakukan secara berkesinambungan. Evaluasi bukan hanya merupakan kegiatan akhir atau penutup dari suatu program tertentu, melainkan merupakan kegiatan yang dilakukan pada permulaan, selama program berlangsung dan pada akhir program setelah program itu dianggap selesai.
- b. Di dalam kegiatan evaluasi diperlukan berbagai informasi atau data yang menyangkut objek yang sedang dievaluasi. Berdasarkan data itulah selanjutnya diambil suatu keputusan sesuai dengan maksud dan tujuan evaluasi yang sedang dilaksanakan. Ketepatan keputusan hasil evaluasi sangat bergantung kepada kesahihan dan objektivitas data yang digunakan dalam pengambilan keputusan.
- c. Setiap kegiatan evaluasi tidak dapat dilepaskan dari tujuan-tujuan yang hendak dicapai. Tanpa menentukan atau merumuskan tujuan-tujuan terlebih dahulu, tidak mungkin menilai sejauhmana pencapaian hasil. Hal ini adalah karena setiap kegiatan penilaian memerlukan suatu kriteria tertentu sebagai acuan dalam menentukan batas ketercapaian objek yang dinilai.

Berdasarkan pemaparan di atas maka dapatlah dimaknai bahwa evaluasi terkait dengan 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. Dengan kata lain evaluasi pada hakikatnya adalah penyediaan informasi yang dapat dijadikan sebagai bahan pertimbangan dalam mengambil keputusan.

## 2. Program

Program didefinisikan sebagai suatu unit atau kesatuan kegiatan yang merupakan realisasi atau implementasi dari suatu kebijakan, berlangsung dalam proses yang berkesinambungan dan terjadi dalam suatu organisasi yang melibatkan sekelompok orang (Arikunto dan Jabar, 2009:4). Dalam hal ini ada tiga pengertian penting dan perlu ditekankan dalam menentukan program yaitu:

- a. Realisasi atau implementasi suatu kebijakan.
- b. Terjadi dalam waktu relatif lama dan bukan kegiatan tunggal tetapi jamak berkesinambungan.
- c. Terjadi dalam organisasi yang melibatkan sekelompok orang.

Menurut Joan sebagaimana dikutip Tayibnapis (2000:9) program adalah segala sesuatu yang dicobalakukan seseorang dengan harapan akan mendatangkan hasil atau pengaruh. Dalam hal ini suatu program dapat saja berbentuk nyata (*tangible*) seperti kurikulum, atau yang berbentuk abstrak (*intangible*) seperti prosedur. Sedangkan menurut Feuerstein (1990:209) program adalah sebuah rencana yang diputuskan terlebih dahulu, biasanya dengan sasaran-sasaran, metode, urutan dan konteks tertentu.

Menurut Suherman dan Sukjaya (1990:24) program adalah suatu rencana kegiatan yang dirumuskan secara operasional dengan memperhitungkan segala faktor yang berkaitan dengan pelaksanaan dan pencapaian program tersebut.

Berdasarkan pemaparan di atas maka dapatlah dimaknai bahwa program adalah suatu rencana yang melibatkan berbagai unit yang berisi kebijakan dan rangkaian kegiatan yang harus dilakukan dalam kurun waktu tertentu. Program dalam hal ini berupa aktivitas atau rangkaian aktivitas yang akan direncanakan.

### **3. Evaluasi Program**

Brikerhoff et-al (1983:2) mendefinisikan evaluasi program adalah suatu proses menemukan sejauhmana tujuan dan sasaran program atau proyek telah terealisasi, memberikan informasi untuk pengambilan keputusan, membandingkan kinerja dengan standar atau patokan untuk mengetahui adanya kesenjangan, penilaian harga dan kualitas dan penyelidikan sistematis tentang nilai atau kualitas suatu objek.

Evaluasi program menurut Tyler adalah proses untuk mengetahui apakah tujuan sudah dapat terealisasikan (Arikunto dan Jabar, 2009:5). Menurut Arikunto (2005:291) evaluasi program adalah kegiatan yang dimaksudkan untuk mengetahui seberapa tinggi tingkat keberhasilan dari kegiatan yang direncanakan.

Berdasarkan pemaparan di atas maka dapatlah dimakna bahwa evaluasi program adalah suatu unit atau kesatuan kegiatan yang bertujuan mengumpulkan informasi tentang realisasi atau implementasi dari suatu kebijakan, berlangsung dalam proses yang berkesinambungan, dan terjadi dalam suatu organisasi yang melibatkan sekelompok orang guna pengambilan keputusan.

### **4. Sponsor**

Sponsor ialah orang atau organisasi yang meminta evaluasi dan membayar untuk itu. Ilustrasinya sebagai berikut, bila Dekan Fakultas Ilmu Tarbiyah dan Keguruan (FITK) Universitas Islam Negeri (UIN) Sumatera Utara untuk melakukan evaluasi terhadap program pendidikan dan latihan profesi guru (PLPG) yang diselenggarakan oleh FITK UIN Sumatera Utara dan mengeluarkan pembiayaan atas evaluasi tersebut maka Dekan FITK UIN Sumatera Utara adalah sponsor.

### **5. Audiensi**

Audiensi adalah orang yang secara langsung atau tidak langsung berurusan dengan evaluasi program yang dilakukan. Audiensi berupa peminat, pemakai atau pelanggan informasi yang dikumpulkan selama program berjalan terdiri atas perencana program, manajer program dan staf yang menjakankan program. Peminat, pemakai atau pelanggan juga mungkin penerima pelayanan atau hasil evaluasi, misalnya program PLPG maka peminat, pemakai atau pelanggannya adalah guru, Kementerian Pendidikan maupun Kementerian Agama.

## 6. Evaluator Program

Evaluator program adalah seseorang yang melakukan evaluasi atau yang memungkinkan terjadinya evaluasi (Feuerstein, 1990:204). Menurut Purwanto dan Suparman (1999:67) 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 bergantung kepada pemilik program. Dengan demikian dapatlah dipahami bahwa evaluator program adalah individu ataupun tim yang melakukan pekerjaan dalam mengevaluasi suatu program. Terkait dengan evaluator program ini akan dibahas lebih detail pada bab dua.

## B. Tujuan dan Manfaat Evaluasi Program

Weiss (1972:4) menyatakan tujuan dilakukannya evaluasi program sebagai berikut:

1. Menunjuk pada penggunaan metode penelitian.
2. Menekankan pada hasil suatu program.
3. Penggunaan kriteria untuk menilai.
4. Kontribusi terhadap pengambilan keputusan dan perbaikan program di masa mendatang.

Menurut Kirkpatrick (1998:17) urgensi diperlukannya evaluasi program adalah:

1. Untuk menunjukkan eksistensi dari dana yang dikeluarkan terhadap pencapaian tujuan dan sasaran program yang dilakukan.
2. Untuk memutuskan apakah kegiatan yang dilakukan akan diteruskan akan dihentikan.
3. Untuk mengumpulkan informasi bagaimana cara untuk mengembangkan program di masa mendatang.

Tujuan evaluasi menurut Scriven mempunyai dua fungsi yaitu fungsi formatif dan fungsi sumatif. Fungsi formatif yaitu evaluasi dipakai untuk perbaikan dan pengembangan kegiatan yang sedang berjalan (program, orang, produk dan sebagainya) sedangkan fungsi sumatif yaitu evaluasi dipakai untuk pertanggungjawaban, keterangan, seleksi atau lanjutan. Dengan kata lain evaluasi bertujuan membantu pengembangan, implementasi



kebutuhan suatu program, perbaikan program, pertanggungjawaban, seleksi, motivasi, menambah pengetahuan, dan dukungan dari yang terlibat (Tayibnapis, 2000:4).

Mutrofin (2010:157) menyatakan tujuan evaluasi program adalah untuk mendapat informasi yang mungkin berguna pada saat memilih di antara berbagai kebijakan atau program alternatif untuk mencapai tujuan sosial. Selanjutnya menurut Tayibnapis (2000:59) tujuan evaluasi dapat bermacam-macam, antara lain sebagai pekerjaan rutin atau tanggung jawab rutin untuk membantu pekerjaan manajer dan karyawan dengan tujuan yang lebih banyak, dan informasi yang lebih lengkap dari yang sudah ada atau memberikan informasi untuk tim pembina atau penasehat, untuk klien, untuk dewan direktur atau pemberi dana atau sponsor.

Sukmadinata (2006:121) menjelaskan tujuan evaluasi program adalah:

1. Membantu perencanaan untuk pelaksanaan program.
2. Membantu dalam penentuan keputusan penyempurnaan atau perubahan program.
3. Membantu dalam penentuan keputusan keberlanjutan atau penghentian program.
4. Menemukan fakta-fakta dukungan dan penolakan terhadap program.
5. Memberikan sumbangan dalam pemahaman proses psikologis, sosial, politik dalam pelaksanaan program serta faktor-faktor yang mempengaruhi program.

Purwanto dan Suparman (1999:30-33) memaparkan tujuan evaluasi adalah:

### **1. Mengkomunikasikan program kepada masyarakat.**

Laporan hasil atau informasi dari evaluasi program yang dilakukan dapat memberikan pemahaman kepada khalayak tentang program atau tentang kinerja/performa. Oleh karena itu, mengkomunikasikan hasil evaluasi program yang lebih lengkap dari sekedar angka-angka kepada masyarakat memiliki keuntungan dan kebaikan terhadap program yang dievaluasi.

### **2. Menyediakan informasi bagi pembuat keputusan.**

Informasi yang dihasilkan dari evaluasi program akan berguna bagi setiap tahapan dari manajemen program mulai sejak perencanaan,

pelaksanaan ataupun ketika akan mengulangi dan melanjutkan program. Hasil evaluasi dapat dijadikan dasar bagi pembuatan keputusan, sehingga keputusan tersebut valid dibandingkan keputusan yang hanya berdasarkan kepada intuisi saja. Pembuat keputusan biasanya memerlukan informasi yang akurat agar dapat memutuskan sesuatu secara tepat dan informasi akurat tersebut antara lain dapat diperoleh dari kegiatan evaluasi yang dilaksanakan secara sistematis. Penyediaan informasi hasil evaluasi bagi pembuat keputusan tersebut tidak terbatas pada keputusan oleh pejabat pemegang otoritas dalam institusi itu saja, tetapi bisa meliputi pembuatan keputusan dalam berbagai level oleh pihak-pihak lain yang terkait.

### **3. Menyempurnakan program yang ada.**

Suatu evaluasi program yang dilaksanakan dengan baik dapat membantu upaya-upaya dalam rangka penyempurnaan jalannya program sehingga lebih efektif. Dengan instrumen yang ada, hasil yang dicapai dapat diukur dan diagnosis. Berbagai kelemahan dan kendala yang mungkin timbul dapat ditemukan dan dikenali, kemudian dianalisis serta ditentukan alternatif pemecahannya yang paling tepat. Komponen-komponen dalam sistem yang memiliki kekurangan dan kelemahan dapat dipelajari dan dicari solusinya. Berdasarkan hasil evaluasi akan dapat diperoleh informasi tentang dampak dari berbagai aspek program dan berhasil juga teridentifikasi berbagai faktor yang diperlukan atau perlu penyempurnaan.

### **4. Meningkatkan partisipasi dan pertumbuhan.**

Dengan adanya informasi hasil evaluasi atas suatu program, maka masyarakat akan lebih terpenggil untuk berpartisipasi dan ikut mendukung upaya-upaya peningkatan dan penyempurnaan program. Hasil evaluasi program yang dimasyarakatkan akan menggugah kepedulian masyarakat terhadap program, menarik perhatiannya dan akhirnya menumbuhkan rasa ikut memiliki (*sense of belonging*) terhadap program tersebut. Apabila hal ini telah terbina maka akan tercipta suatu kontrol eksternal yang ikut memacu dan mengawasi pertumbuhan kualitas dari program yang bersangkutan.

Secara khusus tujuan evaluasi program dalam pendidikan ditegaskan oleh Worthen dkk sebagaimana dikutip Tayibnapis (2000:3) yaitu:

1. Membuat kebijaksanaan dan keputusan.

2. Menilai hasil yang dicapai para peserta didik.
3. Menilai kurikulum.
4. Memberi kepercayaan kepada sekolah.
5. Memonitor dana yang telah diberikan.
6. Memperbaiki materi dari program pendidikan.

Berdasarkan pemaparan di atas maka dapatlah dipahami bahwa tujuan evaluasi adalah untuk memperoleh informasi yang akurat dan obyektif 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.

Berdasarkan pemaparan di atas dapatlah dimaknai bahwa evaluasi program bertujuan untuk mengetahui pencapaian tujuan program yang telah dilaksanakan. Selanjutnya, hasil evaluasi program dimanfaatkan sebagai dasar untuk melaksanakan kegiatan tindak lanjut atau untuk melakukan pengambilan keputusan berikutnya.

## **C. Asumsi Dasar dan Prinsip Umum**

### **1. Asumsi Dasar.**

Terdapat berbagai asumsi yang mendasari kegiatan evaluasi program, asumsi tersebut bersifat mendasar dan berkaitan dengan falsafah tertentu. Asumsi dasar tersebut dipaparkan Purwanto dan Suparman (1999:6) sebagai berikut:

- a. Evaluasi merupakan suatu kebutuhan dan mutlak diperlukan dalam suatu program. Dengan mengingat kepada manfaatnya, maka evaluasi merupakan suatu keharusan dan bagian tak terpisahkan dari kegiatan suatu program.
- b. Evaluasi merupakan salah satu fungsi penting dalam sistem suatu program. Evaluasi berkaitan dengan setiap komponen dalam sistem program dalam seluruh tahapan perancangan dan pengembangan program.
- c. Mengevaluasi program sosial adalah sulit, terutama berkaitan dengan standar yang digunakan. Namun demikian bukan berarti pekerjaan

evaluasi adalah pekerjaan yang tidak mungkin dilakukan dengan berhasil dan memuaskan.

## 2. Prinsip Umum.

Terdapat prinsip-prinsip umum yang harus menjadi perhatian dalam melaksanakan evaluasi suatu program. Prinsip-prinsip umum tersebut menjadi rambu-rambu yang dijadikan pedoman oleh pihak yang melakukan evaluasi sehingga diharapkan dengan mempedomani prinsip-prinsip umum ini, hasil evaluasi program yang dilakukan memiliki kredibilitas yang baik.

Prinsip-prinsip umum dalam melakukan evaluasi suatu program dirangkum oleh Cronbach dan Paton sebagaimana dikutip Arikunto dan Jabar (2009:78) sebagai berikut:

1. Evaluasi program adalah suatu seni. Tidak ada satupun saran untuk rancangan yang paling tepat bagi kegiatan evaluasi, tetapi untuk evaluasi program sebaiknya tidak menggunakan rancangan eksperimen.
2. Evaluator program tidak memiliki wewenang untuk memutuskan hasil program, tetapi sekedar memberikan bantuan data atau informasi kepada pengambil keputusan.
3. Tidak seorangpun di antara evaluator program berhak memberikan pertimbangan kepada pengambil keputusan. Evaluasi program adalah tanggungjawab sebuah tim.
4. Jika wilayah dari program yang dievaluasi luas, evaluator secara individual tidak berhak beranggapan bahwa dirinya memiliki hak atas bagian, tetapi seluruh tim memiliki hak yang sama atas semua bagian.
5. Rancangan evaluasi program bukan sesuatu yang sifatnya kaku dan statis, melainkan merupakan sesuatu yang berproses yaitu fleksibel, dapat dimodifikasi dan diperbaiki selama dalam proses kegiatan.
6. Sebuah program bukan hanya perlakuan tunggal tetapi juga jamak maksudnya suatu program memiliki banyak dimensi. Misalnya program pembelajaran maka memiliki dimensi-dimensi jamak yang dapat dievaluasi dari aspek pendidik, aspek peserta didik, aspek materi/kurikulum, aspek media/sumber dan sebagainya.
7. Aspek afektif dan psikomotorik sebaiknya tidak dihindari dalam proses pengumpulan data, perlu adanya keseimbangan antara data aspek kognitif, afektif dan psikomotorik. Kelemahan umum yang ada dalam

evaluasi adalah bahwa data yang dikumpulkan hanya kognitif atau dominan pada kognitif.

8. Mengevaluasi program sebaiknya tidak hanya memusatkan sasaran perhatian pada hasil atau dampak saja, tetapi semua gejala proses pelaksanaan perlu ditelusuri.

Senada dengan penjelasan Cronbach dan Paton di atas, Purwanto dan Suparman (1999:7-8) memaparkan 7 (tujuh) prinsip dasar evaluasi sebagai berikut:

1. Evaluasi harus dilakukan secara sistematis. Dengan demikian hasilnya dapat diharapkan dapat dipertanggungjawabkan dan dapat memenuhi kebutuhan berkaitan dengan program.
2. Evaluasi dilaksanakan sesuai dengan prinsip dasar dalam sistem instruksional dan berkaitan dengan seluruh aspek dalam sistem instruksional.
3. Evaluasi program harus dilakukan dengan sedapat mungkin menggunakan standar tertentu yang relevan dengan program yang dievaluasi.
4. Sumber kesalahan dapat diidentifikasi. Sumber kesalahan evaluasi terdapat pada beberapa komponen seperti:
  - a. Dalam instrumen evaluasi yang dipergunakan dalam pengumpulan data, seperti isinya yang kurang tepat (kurang valid), terlalu sulit, kurang pasti dan kurang reliabel.
  - b. Pada proses pengumpulan data baik yang menyangkut cara mengumpulkan atau cara mencatat dan memberi skor.
  - c. Kesalahan pada individu yang dievaluasi seperti kekurang-sungguhan dan kekurangjujuran individu tersebut.
5. Kesalahan dapat dikurangi (*minimized*). Mengetahui sumber-sumber kesalahan seperti diuraikan di atas adalah penting untuk mencegah terjadinya kesalahan tersebut baik pada saat menyusun instrumen evaluasi, proses pengumpulan data dan pendekatan dengan individu-individu yang dievaluasi. Berikut ini upaya-upaya pokok yang harus dilakukan yaitu:
  - a. Semua kegiatan evaluasi harus terkait dengan alasan melakukan evaluasi.
  - b. Harus mempunyai persepsi yang jelas dan konsisten tentang tujuan, proses, pengumpulan data dan pengolahan data serta pelaporannya.

- c. Mengumpulkan sampel informasi yang representatif sehingga hasil pengolahan informasi tersebut dapat mencerminkan keadaan seluruh populasi.
6. Kesalahan dapat dihitung. Kesalahan pada instrumen dapat dihitung melalui validitas dan reliabilitasnya. Validitas menyangkut kriteria instrumen dengan faktor yang seharusnya diukur, sedangkan reliabilitas meliputi stabilitas dan konsistensi internal instrumen.
7. Seberapapun tingkat kehati-hatian dalam mengumpulkan informasi kesalahan dapat saja terjadi.

## **D. Karakteristik Evaluasi Program**

Evaluasi program memiliki karakteristik yang berbeda dengan penelitian pada umumnya. Evaluasi program memiliki karakteristik yang unik dan tersendiri sebagai ciri khasnya, dalam hal ini terdapat 8 (delapan) karakteristik evaluasi program sebagaimana dipaparkan Arikunto dan Jabar (2009:8-9) sebagai berikut:

1. Proses kegiatan evaluasi program tidak menyimpang dari kaidah-kaidah yang berlaku bagi penelitian pada umumnya.
2. Dalam melaksanakan evaluasi program, peneliti harus berpikir secara sistematis yaitu memandang program yang diteliti sebagai sebuah kesatuan yang terdiri dari beberapa komponen atau unsur yang saling berkaitan satu sama lain dalam menunjang keberhasilan kinerja dari objek yang dievaluasi.
3. Agar dapat mengetahui secara rinci kondisi dari objek yang dievaluasi, perlu adanya identifikasi komponen yang berkedudukan sebagai faktor penentu bagi keberhasilan program.
4. Menggunakan standar, kriteria atau tolak ukur sebagai perbandingan dalam menentukan kondisi nyata dari data yang diperoleh dan untuk mengambil kesimpulan.
5. Kesimpulan atau hasil evaluasi program digunakan sebagai masukan atau rekomendasi bagi sebuah kebijakan atau rencana program yang telah ditentukan. Dengan kata lain, dalam melakukan kegiatan evaluasi program, peneliti harus berkiblat pada tujuan program kegiatan sebagai standar, kriteria atau tolak ukur.

6. Agar informasi yang diperoleh dapat menggambarkan kondisi nyata secara rinci untuk mengetahui bagian mana dari program yang belum terlaksana, maka perlu ada identifikasi komponen yang dilanjutkan dengan identifikasi subkomponen, sampai pada indikator dari program yang dievaluasi.
7. Standar, kriteria atau tolak ukur diterapkan pada indikator, yaitu bagian yang paling kecil dari program agar dapat dengan cermat diketahui letak kelemahan dari proses kegiatan.
8. Dari hasil evaluasi program harus dapat disusun sebuah rekomendasi secara rinci dan akurat sehingga dapat ditentukan tindak lanjut secara tepat.

## **E. Komponen dan Indikator Program**

Program merupakan sistem, sedangkan sistem adalah satu kesatuan dari beberapa bagian atau komponen program yang saling kait-mengait dan bekerja sama satu dengan lain untuk mencapai tujuan yang sudah ditetapkan dalam sistem. Dengan demikian program terdiri dari komponen-komponen yang saling berkaitan dan saling menunjang dalam rangka mencapai suatu tujuan (Arikunto dan Jabar, 2009:9).

Komponen program adalah bagian-bagian atau unsur-unsur yang membangun sebuah program yang saling terkait dan merupakan faktor-faktor penentu keberhasilan program. Oleh karena suatu program merupakan sebuah sistem maka komponen-komponen program tersebut dapat dipandang sebagai bagian sistem dan dikenal dengan istilah “subsistem”.

Selanjutnya istilah indikator berasal dari bahasa Inggris yaitu *to indicate* yang berarti menunjukkan atau tanda. Jadi indikator merupakan sesuatu yang dapat menunjukkan atau sebagai tanda dari suatu subkomponen dan sekaligus menunjukkan atau sebagai tanda suatu komponen. Dalam kegiatan evaluasi program, indikator merupakan petunjuk untuk mengetahui keberhasilan atau ketidakberhasilan suatu kegiatan. Perlu diketahui bahwa ketidakberhasilan suatu kegiatan dapat juga dipengaruhi oleh komponen atau subkomponen yang lain.

Ilustrasi dari penjelasan terkait dengan komponen, subkomponen dan indikator dari program yang akan dievaluasi maka dalam hal ini dikutip deskripsi yang disampaikan Arikunto dan Jabar (2009:10-12) terkait dengan

evaluasi program pembelajaran. Di mana dalam pembelajaran sebagai program memiliki komponen-komponen yang menjadi faktor penting keberlangsungannya, dalam hal ini faktor-faktor yang dimaksud sebagai berikut:

1. Pendidik.
2. Peserta didik.
3. Materi/kurikulum.
4. Sarana dan prasarana.
5. Pengelolaan.
6. Lingkungan.

Lebih detailnya dapat dilihat tabel berikut ini:

**Tabel 1.1. Ilustrasi Komponen, Subkomponen dan Indikator Program Pembelajaran**

Program	Komponen	Subkomponen	Indikator
Pembelajaran	Pendidik	Kemampuan merencanakan pembelajaran	<ul style="list-style-type: none"> <li>• Merumuskan tujuan pembelajaran</li> <li>• Pemilihan dan pengorganisasian materi</li> <li>• Pemilihan media/sumber</li> <li>• Pemilihan metode/strategi</li> <li>• Menentukan bentuk penilaian</li> </ul>
		Kemampuan melaksanakan pembelajaran	<ul style="list-style-type: none"> <li>• Memulai pembelajaran</li> <li>• Mengelola kegiatan belajar</li> <li>• Mengorganisir waktu</li> <li>• Mengorganisir siswa</li> <li>• Mengorganisir media</li> <li>• Menggunakan metode</li> <li>• Memberi penguatan</li> <li>• Melakukan penilaian</li> <li>• Menutup pembelajaran</li> </ul>
		Kemampuan menilai hasil pembelajaran	<ul style="list-style-type: none"> <li>• Membuat instrumen dan rubrik penilaian.</li> <li>• Melakukan penskoran</li> <li>• Melakukan administrasi penilaian</li> </ul>



	Peserta didik	Minat belajar	.....
			.....
		Motivasi belajar	.....
			.....
		Gaya belajar	.....
			.....
		Dst	Dst
	Kurikulum	.....	.....
		.....	.....
	Sarana dan prasarana	.....	.....
		.....	.....
	Pengelolaan	.....	.....
		.....	.....
	Lingkungan	.....	
		.....	

Contoh ilustrasi lain yang dapat diberikan adalah program bimbingan konseling di sekolah. Dalam hal ini bimbingan konseling adalah program, sedangkan komponennya adalah (1) guru bimbingan konseling (2) program pelayanan bimbingan konseling, (3) pengelolaan administrasi bimbingan konseling, dan (4) sarana/prasarana bimbingan konseling. Detailnya dapat dilihat pada tabel berikut:

**Tabel 1.2. Ilustrasi Komponen, Subkomponen dan Indikator Program Bimbingan Konseling**

Program	Komponen	Subkomponen	Indikator
BK	Guru BK	Latar Pendidikan	• S1 BK • S1 Non BK
		Pengalaman	• < 5 tahun • > 5 tahun
		Diklat BK	• Pernah mengikuti • Tidak pernah

	Program pelayanan BK	Perencanaan program BK	<ul style="list-style-type: none"> <li>• Merencanakan jenis layanan BK yang variatif.</li> <li>• Merencanakan langkah-langkah program BK</li> <li>• Merencanakan program BK dan relevansi dengan perkembangan zaman</li> </ul>
		Melaksanakan program layanan BK	<ul style="list-style-type: none"> <li>• Memperhatikan karakteristik siswa</li> <li>• Menempatkan siswa sebagai subjek</li> <li>• Menciptakan suasana layanan BK yang menyenangkan</li> </ul>
		Menilai proses dan pelaksanaan layanan BK	<ul style="list-style-type: none"> <li>• Menyiapkan dokumen penilaian BK secara detil</li> <li>• Melakukan penilaian BK secara periodik terhadap siswa.</li> <li>• Menilai secara internal terhadap layanan BK yang dilakukan.</li> </ul>
		Tindak lanjut berdasarkan hasil penilaian	<ul style="list-style-type: none"> <li>• Berdiskusi dengan teman sejawat tentang evaluasi BK yang efisien dan tepat untuk menuntaskan masalah yang dihadapi siswa.</li> <li>• Memberikan respon langsung terhadap persoalan yang dihadapi siswa</li> <li>• Menjalin hubungan yang baik dengan orang tua siswa tentang masalah yang dihadapi siswa</li> </ul>
	Pengelolaan Administrasi BK	Arsip dokumen layanan BK	<ul style="list-style-type: none"> <li>• Dokumen layanan BK di arsipkan dalam dokumen file</li> <li>• Dokumen file terjamin keamanannya</li> <li>• Aksesibilitas dalam menemukan dokumen file</li> </ul>
		Media pengarsipan	<ul style="list-style-type: none"> <li>• Media tulis</li> <li>• Media digital</li> </ul>
	Sarana/ prasarana	Hard ware	<ul style="list-style-type: none"> <li>• Ruangan untuk layanan BK</li> <li>• Fasilitas yang terdapat dalam ruangan seperti AC/kipas angin, computer</li> </ul>
		Soft ware	<ul style="list-style-type: none"> <li>• AUM (alat ungkap masalah)</li> <li>• Alat perekam dan sebagainya</li> </ul>

## **F. Kriteria Dalam Evaluasi Program**

### **1. Pengertian**

Kriteria atau dikenal dengan istilah tolok ukur atau standar adalah sesuatu yang digunakan sebagai patokan atau batas minimal untuk sesuatu yang diukur (Arikunto dan Jabar, 2009:30). Dalam hal ini kriteria menunjukkan gradasi atau tingkatan dan ditunjukkan dalam bentuk kata keadaan atau predikat.

### **2. Urgensi Kriteria.**

Urgensi kriteria dalam evaluasi program dijelaskan Arikunto dan Jabar (2009:32) sebagai berikut:

- a. Dengan adanya kriteria atau tolok ukur, evaluator dapat lebih mantap dalam melakukan evaluasi terhadap objek yang akan dinilai karena ada patokan yang diikuti.
- b. Kriteria atau tolok ukur yang sudah dibuat dapat digunakan untuk menjawab atau mempertanggungjawabkan hasil penilaian yang sudah dilakukan, jika ada pihak yang ingin menelusuri lebih jauh atau ingin mengkaji ulang.
- c. Kriteria atau tolok ukur digunakan untuk mengekang masuknya unsur subjektif yang ada pada diri evaluator. Dengan adanya kriteria maka dalam melakukan evaluasi, evaluator dituntun oleh kriteria, mengikuti butir demi butir, tidak mendasarkan diri atas pendapat pribadi yang mungkin sekali dicemari oleh seleranya.
- d. Dengan adanya kriteria atau tolok ukur maka hasil evaluasi akan sama meskipun dilakukan dalam waktu yang berbeda dan dalam kondisi fisik evaluator yang berbeda pula. Misalnya evaluator sedang dalam kondisi badan yang masih segar atau dalam keadaan lelah hasilnya akan sama.
- e. Kriteria atau tolok ukur memberikan arahan kepada evaluator apabila banyaknya evaluator lebih dari satu orang. Kriteria atau tolok ukur yang baik akan ditafsirkan sama oleh siapa saja yang menggunakannya.

### 3. Sumber Penyusunan Kriteria

Sumber penyusunan kriteria dalam evaluasi program dipaparkan oleh Arikunto dan Jabar (2009:33-34) sebagai berikut:

**a. Sumber pertama.**

Apabila yang dievaluasi merupakan suatu implementasi kebijakan maka yang dijadikan sebagai kriteria atau tolok ukur adalah peraturan atau ketentuan yang sudah dikeluarkan berkenaan dengan kebijakan yang bersangkutan. Apabila penentu kebijakan tidak mengeluarkan ketentuan secara khusus maka penyusun kriteria dapat menggunakan ketentuan yang pernah berlaku umum yang sudah dikeluarkan pengambil kebijakan terdahulu dan belum pernah dicabut masa berlakunya.

**b. Sumber kedua.**

Dalam mengeluarkan kebijakan biasanya disertai dengan buku pedoman atau petunjuk pelaksanaan. Di dalam petunjuk pelaksanaan tertuang informasi yang lengkap antara lain dasar pertimbangan dikeluarkannya kebijakan, prinsip, tujuan, sasaran, dan rambu-rambu pelaksanaannya. Butir-butir yang tertera di dalamnya terutama dalam tujuan kebijakan, mencerminkan harapan dari kebijakan. Oleh karena itu, pedoman atau petunjuk pelaksanaan itulah yang distatuskan sebagai sumber kriteria.

**c. Sumber ketiga.**

Apabila tidak ada ketentuan atau petunjuk pelaksanaan yang dapat digunakan oleh penyusun sebagai sumber kriteria maka penyusun menggunakan konsep atau teori-teori yang terdapat dalam literatur ilmiah.

**d. Sumber keempat**

Jika tidak ada ketentuan, peraturan atau petunjuk pelaksanaan, dan juga tidak ada teori yang diacu, penyusun disarankan untuk menggunakan hasil penelitian. Dalam hal ini sebaiknya tidak langsung mengacu pada hasil penelitian yang baru saja diselesaikan seorang peneliti, tetapi disarankan sekurang-kurangnya hasil penelitian yang sudah dipublikasikan atau diseminarkan. Jika ada yang sudah disajikan kepada khalayak ramai yaitu disimpan di perpustakaan.

**e. Sumber kelima.**

Apabila penyusun tidak menemukan acuan yang tertulis dan mantap, maka dapat meminta bantuan pertimbangan kepada orang yang dipandang mempunyai kelebihan dalam bidang yang sedang dievaluasi

sehingga terjadi langkah yang dikenal dengan istilah *expert judgment*.

**f. Sumber keenam.**

Apabila sumber acuan tidak ada, sedangkan ahli yang dapat diandalkan sebagai orang yang lebih memahami masalah dibanding penyusun juga sukar dicari atau dihubungi maka penyusun dapat menentukan kriteria secara bersama dengan anggota tim atau beberapa orang yang mempunyai wawasan tentang program yang akan dievaluasi. Kriteria yang tersusun dari diskusi ini merupakan hasil kesepakatan bersama.

**g. Sumber ketujuh.**

Dalam keadaan yang sangat terpaksa karena acuan tidak ada, ahli juga tidak ada, sedangkan untuk menyelenggarakan diskusi terlalu sulit maka jalan terakhir adalah melakukan pemikiran sendiri. Dalam keterpaksaan seperti ini penyusun kriteria hanya mengandalkan pemikiran sendiri yang akan digunakan untuk mengevaluasi program. Jika ternyata sesudah digunakan dalam mengevaluasi masih menjumpai kesulitan, penyusun harus meninjau kembali dan wajib memperbaikinya berkali-kali sampai mencapai suatu rumusan yang sesuai dengan kondisi yang diinginkan.

## **4. Jenis Kriteria.**

Jenis kriteria atau tolok ukur yang digunakan dalam evaluasi program dibedakan atas dua jenis yaitu kriteria kuantitatif dan kriteria kualitatif. Berikut penjelasannya:

**a. Kriteria kuantitatif.**

Kriteria kuantitatif dapat dibedakan menjadi dua yaitu: kriteria kuantitatif tanpa pertimbangan dan kriteria kuantitatif dengan pertimbangan (Arikunto dan Jabar, 2009:34).

*Kriteria kuantitatif tanpa pertimbangan* adalah kriteria yang disusun hanya dengan memperhatikan rentangan bilangan tanpa mempertimbangkan apa-apa dilakukan dengan membagi rentangan bilangan. Contoh: kondisi maksimal yang diharapkan untuk hasil tes diperhitungkan 100. Jika penyusun menggunakan lima kategori nilai maka antara nilai 0 sampai 100 dibagi rata sehingga menghasilkan kategori sebagai berikut:

- Nilai 5 (baik sekali) yaitu skor 81 – 100.
- Nilai 4 (baik) yaitu skor 61 – 80.
- Nilai 3 (cukup) yaitu skor 41 – 60.
- Nilai 2 (kurang) yaitu skor 21 – 40.
- Nilai 1 (kurang sekali) yaitu skor < 21.

Kategori tidak saja dalam bentuk baik sekali sampai kurang sekali, tetapi dapat juga tinggi sekali sampai rendah sekali, sering kali sampai jarang sekali. Selain itu dapat juga menggunakan istilah lain yang menunjukkan kualitas suatu keadaan, sifat atau kondisi seperti banyak sekali, sibuk sekali, dan sebagainya. Untuk pertimbangan atau pendapat, maka dapat menggunakan kata sangat setuju, setuju dan seterusnya.

*Kriteria kuantitatif dengan pertimbangan* yaitu kriteria kuantitatif dikategorikan yang dibuat karena adanya pertimbangan tertentu berdasarkan sudut pandang dan pertimbangan evaluator.

#### **b. Kriteria kualitatif.**

Kriteria kualitatif adalah kriteria yang dibuat tidak menggunakan angka-angka, dalam hal ini yang dipertimbangkan adalah indikator dan yang dikenai kriteria adalah komponen. Kriteria kualitatif dibedakan ada dua jenis yaitu kriteria kualitatif tanpa pertimbangan dan kriteria kualitatif dengan pertimbangan.

*Kriteria kualitatif tanpa pertimbangan*, dalam hal ini penyusun kriteria tinggal menghitung indikator dalam komponen yang dapat memenuhi persyaratan. Dari penjelasan tersebut dapatlah dimaknai bahwa komponen adalah unsur pembentuk kriteria program dan indikator adalah unsur pembentuk kriteria komponen.

*Kriteria kualitatif dengan pertimbangan*. Dalam menyusun kriteria terlebih dahulu evaluator perlu merundingkan jenis kriteria mana yang akan digunakan, yaitu memilih kriteria tanpa pertimbangan atau dengan pertimbangan. Jika yang dipilih adalah kriteria dengan pertimbangan maka tentukan indikator mana yang harus diprioritaskan atau dianggap lebih penting dari yang lain. Kriteria kualitatif dengan pertimbangan disusun melalui dua cara yaitu: (1) dengan mengurutkan indikator, dan (2) menggunakan pembobotan.

Kriteria kualitatif dengan pertimbangan mengurutkan indikator dilakukan dengan urutan prioritas maka dihasilkan kriteria kualitatif dengan pertimbangan sebagai berikut:

- Nilai 5, jika memenuhi semua indikator (4 indikator).
- Nilai 4 jika memenuhi 3 indikator.
- Nilai 3 jika memenuhi 2 indikator.
- Nilai 2 jika memenuhi 1 indikator.
- Nilai 1 jika tidak memenuhi satupun indikator.

Jika yang dikenai kriteria itu bukan indikator, tetapi subindikator maka yang digunakan untuk mempertimbangkan penentuan kriteria adalah subindikator atau rincian dari indikator. Dalam hal ini kriteria yang akan digunakan ditentukan atas dasar subindikator yang sudah diidentifikasi terlebih dahulu. Namun yang perlu diingat evaluator bahwa tidaklah sebuah indikator itu dapat dirinci lagi ke dalam subindikator, dalam keadaan seperti ini indikator merupakan satu-satunya dasar pembuatan kriteria.

Kriteria kualitatif dengan pertimbangan pembobotan dalam hal ini jika dalam menentukan kriteria dengan pertimbangan indikator, nilai dari tiap-tiap indikator tidaklah sama, kemudian letak, kedudukan dan pemenuhan persyaratannya dibedakan dengan menentukan urutan, dalam pertimbangan pembobotan indikator-indikator yang ada diberi nilai dengan bobot berbeda.

Penentuan peranan subindikator dalam mendukung nilai-nilai indikator harus disertai dengan alasan-alasan yang tepat. Jika sudah ditentukan pembobotan maka evaluator tinggal memilih akan menggunakan skala dalam menilai objek, dapat skala 1 – 3, skala 1 – 5 atau skala 1 – 100.

Jika nilai indikator disingkat NI, bobot subindikator disingkat BSI, nilai subindikator disingkat NSI dan jumlah bobot disingkat JB maka rumus nilai akhir indikator sebagai berikut:

$$NI = \frac{BSI \times NSI}{JB}$$

Jika nilai komponen disingkat NK, bobot indikator disingkat BI, nilai indikator disingkat NI, dan jumlah bobot disingkat JB, maka rumus akhir komponen sebagai berikut:

$$NK = \frac{BI \times NI}{JB}$$

## BAB II

# EVALUATOR PROGRAM

### A. Pengertian

**F**euerstein (1990:204) memaparkan evaluator program adalah seseorang yang melakukan evaluasi atau yang memungkinkan terjadinya evaluasi. Hal senada dijelaskan oleh Purwanto dan Suparman (1999:50) bahwa evaluator program orang yang dipercaya oleh pemilik program dan orang-orang yang berkepentingan dengan program (*stakeholder*) untuk melaksanakan evaluasi.

Berdasarkan pemaparan di atas maka dapatlah dimaknai bahwa evaluator program adalah pihak dalam hal ini individu (biasanya berupa tim) yang melakukan evaluasi terhadap suatu program yang tersebut bertanggung jawab secara penuh terhadap hasil penilaian terhadap program yang dievaluasi.

Namun perlu diingat bahwa evaluator program tidak memiliki hak untuk mengambil keputusan tentang program, tetapi sekedar memberikan rekomendasi kepada pengambil keputusan, selanjutnya pihak pengambil keputusan itulah yang menentukan tindak lanjut.

### B. Jenis Evaluator

Apabila ditelisik berdasarkan asal atau dari mana evaluator program, maka dapat diklasifikasi atas 2 (dua) jenis yaitu:

#### 1. Evaluator internal.

Evaluator internal adalah orang dalam program atau orang yang sangat mengetahui hal ihwal program yang dievaluasi (Feuerstein, 1990:15). Selanjutnya dijelaskan oleh Feuerstein bahwa evaluator internal sudah



mengetahui fungsi-fungsi, tujuan-tujuan, problem-problem, kekuatan-kekuatan dan kelemahan-kelemahan program.

Hal senada dijelaskan oleh Arikunto dan Jabar (2009:23) evaluator internal adalah petugas evaluasi program yang sekaligus merupakan salah satu dari petugas atau anggota pelaksana program yang akan dievaluasi

Merujuk kepada penjelasan di atas dapatlah dimaknai bahwa evaluator internal adalah individu yang menjadi evaluator suatu program yang sekaligus merupakan salah seorang dari anggota dalam program tersebut. Individu yang berasal dari satuan program yang dievaluasi menjadi evaluasi internal memiliki kelebihan dan kelemahan tersendiri. Contoh Fakultas Ilmu Tarbiyah dan Ilmu Keguruan (FITK) UIN Sumatera yang melaksanakan program pendidikan dan latihan profesi guru (PLPG), di samping pihak Fakultas membentuk kepanitiaan PLPG juga memberikan mandat kepada beberapa orang yang bukan dalam kepanitiaan untuk mengamati pelaksanaan program PLPG. Tim evaluator ini dapat ditunjuk sejak awal bersama panitia PLPG dan dapat juga dibentuk kemudian.

Feuerstein (1990:16) memaparkan kelebihan dan kekurangan evaluator internal sebagai berikut:

- a. Terlalu banyak mengetahui program.
- b. Sangat sulit untuk bersikap objektif.
- c. Merupakan bagian dari struktur kekuasaan dan kewenangan yang ada.
- d. Mungkin didorong oleh harapan-harapan pribadi yang akan diperoleh.
- e. Mungkin tidak terlatih secara khusus dalam metode evaluasi. Tidak banyak (mungkin hanya sedikit lebih banyak) pengalaman mengikuti training dibanding orang lain yang terlibat dalam program.
- f. Akrab dengan dan mengerti program tersebut dan dapat menafsirkan perilaku-perilaku dan sikap-sikap pribadi.
- g. Sudah dikenal orang yang terlibat dalam program sehingga tidak menimbulkan gangguan atau hambatan. Rekomendasi-rekomendasi akhir mungkin kurang menimbulkan kekhawatiran.

Menurut Arikunto dan Jabar (2009:23) kelebihan evaluator internal adalah:

- a. Evaluator internal memahami seluk-beluk secara baik program yang akan di evaluasi sehingga kekhawatiran untuk tidak atau kurang tepatnya sasaran tidak perlu ada, dengan kata lain evaluasi tepat pada sasaran.

- b. Oleh karena evaluator adalah orang dalam, pengambil keputusan tidak perlu banyak mengeluarkan dana/honor untuk membayar evaluator program.

Kelemahan dari penggunaan evaluator internal dalam mengevaluasi suatu program sebagai berikut:

- a. Adanya unsur subjektivitas dari evaluator, sehingga berusaha menyampaikan aspek positif dari program yang dievaluasi dan menginginkan agar kebijakan tersebut dapat diimplementasikan dengan baik pula, dengan kata lain evaluator internal dapat dikhawatirkan akan bertindak subjektif.
- b. Oleh karena sudah memahami seluk-beluk program, jika evaluator yang ditunjuk kurang sabar, kegiatan evaluasi akan dilaksanakan dengan tergesa-gesa sehingga kurang cermat.

## **2. Evaluator Eksternal.**

Evaluator eksternal adalah seseorang yang mampu mengamati sebuah program secara jelas karena dia tidak terlibat secara pribadi dan dengan demikian dia tidak akan memiliki sesuatu yang bersifat subjektif untuk diperoleh atau dibuang dari evaluasi (Feuerstein, 1990:15).

Menurut Arikunto dan Jabar (2009:24) evaluator eksternal adalah orang-orang yang tidak terkait dengan kebijakan dan implementasi program, mereka berada di luar dan diminta oleh pengambil keputusan untuk mengevaluasi keberhasilan program atau keterlaksanaan kebijakan yang sudah diputuskan.

Berdasarkan pemaparan di atas maka dapat dipahami bahwa evaluator eksternal atau evaluator luar adalah individu yang tidak terkait dengan kebijakan dan implementasi program. Individu tersebut berada di luar dan diminta oleh pengambil keputusan untuk mengevaluasi keberhasilan program atau keterlaksanaan kebijakan yang sudah diputuskan.

Penggunaan individu-individu yang menjadi evaluator eksternal dalam melakukan evaluasi suatu program memiliki kelebihan dan kelemahan tersendiri. Feuerstein (1990:16) memaparkan kelebihan dan kekurangan evaluator eksternal sebagai berikut:

- a. Dapat melihat program dengan penglihatan segar.
- b. Tidak terlihat secara personal, sehingga lebih mudah bersikap objektif.
- c. Tidak termasuk dalam struktur kekuasaan yang ada.

- d. Tidak memperoleh apa-apa dari program, tetapi mungkin memperoleh penghargaan dari evaluasi.
- e. Terlatih dalam metode evaluasi. Mungkin sudah berpengalaman dalam melakukan evaluasi yang lain. Dianggap sebagai seorang ahli dalam program.
- f. Mungkin tidak mengerti program dan orang yang terlibat di dalamnya.
- g. Dapat menimbulkan kegelisahan karena staf program dan partisipan tidak mengetahui secara pasti motivasi seorang evaluator.

Selanjutnya Arikunto dan Jabar (2009:24) memaparkan kelebihan dan kelemahan evaluator eksternal sebagai berikut:

- a. Oleh dikarenakan tidak berkepentingan atas keberhasilan program maka evaluator eksternal dapat bertindak secara objektif selama melaksanakan evaluasi dan mengambil kesimpulan. Apapun hasil evaluasi, tidak akan ada respon emosional dari evaluator karena tidak ada keinginan untuk memperlihatkan bahwa program tersebut berhasil. Kesimpulan yang dibuat akan lebih sesuai dengan keadaan dan kenyataan.
- b. Seorang ahli yang dibayar, biasanya akan mempertahankan kredibilitas kemampuannya, dengan begitu evaluator eksternal akan bekerja secara serius dan hati-hati.

Kelemahan penggunaan evaluator eksternal dalam melakukan evaluasi suatu program adalah:

- a. Evaluator eksternal adalah orang baru yang sebelumnya tidak mengenal kebijakan tentang program yang akan dievaluasi. Mereka berusaha mengenal dan mempelajari seluk-beluk program tersebut setelah mendapat permintaan untuk mengevaluasi. Mungkin sekali pada waktu mendapat penjelasan atau mempelajari isi kebijakan, ada hal-hal yang kurang jelas. Hal itu wajar karena evaluator eksternal tidak ikut dalam proses kegiatannya. Dampak dari ketidakjelasan pemahaman tersebut memungkinkan kesimpulan yang diambil kurang tepat.
- b. Pemborosan, pengambil keputusan/kebijakan (dalam hal ini bertindak sebagai sponsor) harus mengeluarkan dana/honor yang cukup banyak untuk membayar evaluator eksternal tersebut.

Berdasarkan pemaparan di atas dapatlah dimaknai bahwa evaluator program dapat berasal dari kalangan internal (evaluator dan pelaksana program) dan kalangan eksternal (orang di luar pelaksana program tetapi orang yang terkait dengan kebijakan dan implementasi program).

Selanjutnya mencermati kelebihan dan kelemahan evaluator internal maupun evaluator eksternal di atas, maka timbulkan pertanyaan bagaimana yang lebih baik dalam melaksanakan evaluasi suatu program apakah menggunakan evaluator internal atau evaluator eksternal? Menurut hemat penulis, sebaiknya dalam melakukan evaluasi terhadap suatu program maka lebih tepat dan baik mengkombinasikan penggunaan evaluator internal dan evaluator eksternal. Dengan demikian evaluator internal sebagai pihak yang telah mengenal secara mendalam tentang program yang dievaluasi dapat menjelaskan kepada pihak evaluator eksternal sehingga diperkirakan tidak akan terjadi manipulasi hasil. Hal ini menguntungkan bagi pengambil keputusan atau pelaksana program yang dievaluasi.

### **C. Pertimbangan Dalam Penentuan Evaluator.**

Terdapat berbagai pertimbangan yang harus diperhatikan oleh pemilik program dalam menentukan evaluator program apakah evaluator internal, evaluator eksternal ataupun kombinasi dari evaluator internal dan evaluator eksternal. Pertimbangan yang harus diperhatikan tersebut dijelaskan oleh Purwanto dan Suparman (1999: 52-53) sebagai berikut:

#### **1. Pertimbangan Antara Evaluator Orang Dalam dan Orang Luar.**

Sebaiknya evaluator berasal dari orang dalam atau orang luar. Apakah kelebihan dan kekurangan masing-masing? Orang dalam adalah orang yang berasal dari bagian atau institusi penyelenggara program dan biasanya telah ikut dalam proses pengembangan dan pelaksanaan program. Sedangkan yang dimaksud orang luar adalah mereka yang berperan sebagai evaluator berasal dari luar bagian atau institusi penyelenggara program.

Apabila evaluator ditentukan berasal dari orang dalam, kelebihanannya adalah evaluator tersebut sudah mengetahui organisasi dengan baik dan mengetahui reputasi, status, kredibilitas organisasi tempatnya bekerja. Orang dalam juga memiliki hubungan yang baik dengan staf, memahami saluran komunikasi dalam organisasi, telah memahami program dan telah

memiliki minat terhadap keberhasilan program. Apabila evaluator orang dalam maka kelemahannya adalah terjadinya bias karena konflik kepentingan, mungkin evaluator tidak memiliki keterampilan evaluasi atau pekerjaan evaluasi yang dilaksanakan terganggu oleh tugas lain dan akibatnya tidak dapat menepati waktu.

Sebaliknya apabila evaluator ditentukan berasal dari orang luar maka kelebihanannya adalah tidak mempunyai pendapat tentang organisasi tersebut sebelumnya (netral) dan bisa bertindak sebagai pengamat independen, objektif sebagai pengamat, dan lebih kompeten dalam teknik evaluasi. Sedangkan apabila evaluator dari orang luar maka kelemahannya adalah kurang akrab dengan kebiasaan organisasi, tidak mengenal tatacara yang ada di organisasi yang dievaluasi, bahkan mungkin ada yang berlawanan dengan kebiasaannya, memerlukan waktu untuk memahami program dan pemilihan biasanya hanya berdasarkan pada rekomendasi.

## **2. Pertimbangan Antara Evaluator Tim dan Individual.**

Manakah yang lebih baik, evaluator terdiri dari beberapa orang yang bekerja dalam tim atau masing-masing bertanggungjawab secara individual? Bagaimana sebaiknya evaluator bekerja dalam tim atau secara individual? Apa masalahnya jika evaluator adalah suatu tim atau jika individual? Apakah evaluator adalah individu atau perorangan maka kelebihanannya adalah adanya kejelasan tentang siapa yang harus bertanggungjawab. Sedangkan kelemahannya evaluator individual adalah keberhasilan atau kegagalan evaluasi tergantung pada satu orang.

Sebenarnya hampir mustahil pekerjaan evaluasi program hanya diselesaikan oleh satu orang tanpa bantuan orang lain. Apabila evaluator ditentukan tim maka kelebihanannya adalah adanya pembagian tanggungjawab yang jelas dan evaluator terdiri atas gabungan orang dengan berbagai keahlian sehingga saling melengkapi. Sementara itu kelemahannya adalah perlu waktu untuk pembentukan tim, peralatan dan pertimbangan politis dan dibutuhkan biaya yang tidak sedikit.

### 3. Pertimbangan Antara Evaluator Bekerja Penuh dan Bekerja Paruh Waktu.

Sebaiknya evaluator ditugaskan secara penuh ataukah bekerja secara paruh waktu? Bagaimana pula dengan masalah hubungan atau kontrak kerja evaluator? Manakah yang lebih baik, evaluator yang bekerja penuh (*full time*) ataukah bekerja paruh waktu (*part time*)? Masing-masing pilihan ada kelebihan dan kelemahannya.

Kelebihan apabila evaluator bekerja penuh adalah pekerjaan teroganisir dan terkait dengan logis, dan ketepatan dan arus informasi tidak tergantung pada evaluator. Kelemahan dari evaluator apabila bekerja penuh adalah mahal, mengurangi partisipasi dalam kegiatan evaluasi dan evaluator tampak seperti orang luar.

Apabila evaluator bekerja paruh waktu, kelebihanannya adalah dapat melibatkan berbagai keahlian dalam waktu tidak terlalu lama dan dimungkinkannya penggunaan tenaga ahli dari luar. Sementara itu kelemahannya adalah kunjungan yang singkat tidak memungkinkan untuk mempelajari permasalahan secara menyeluruh dan perlu biaya dan peralatan yang cukup banyak untuk penjadwalan.

### 4. Pertimbangan Antara Evaluator Amatir dan Profesional.

Apakah evaluator tenaga amatir atau profesional? Apakah kelebihan dan kelemahannya masing-masing? Terakhir, masalah pilihan antara tenaga amatir dan profesional dan bagaimana resikonya? Perlu ditekankan di sini bahwa yang dimaksud dengan profesional adalah mereka yang menjadikan pekerjaan evaluasi atau penelitian sebagai pekerjaan pokok sehari-hari dan telah menekuni pekerjaan evaluasi dalam waktu yang cukup lama. Orang-orang di luar kriteria tersebut dianggap sebagai amatir.

Apabila kita memilih tenaga amatir sebagai evaluator maka kelebihanannya adalah meskipun amatir evaluator biasanya dapat memahami isi dan objek evaluasi dengan baik dan dapat memilih berbagai ketrampilan evaluasi berdasarkan pengalaman. Kelemahannya, evaluator amatir karena kurangnya pengetahuan tentang objek akibatnya menurunkan objektivitas evaluasi, kemampuan evaluasinya terbatas dan memiliki keterbatasan dalam pilihan rancangan evaluasi.

Sebaliknya apabila menggunakan tenaga profesional maka kelebihanannya adalah evaluator dapat menjalankan evaluasi berdasarkan pengalaman dan keterampilan teknis dan evaluator memiliki berbagai pilihan cara evaluasi berdasarkan pengalaman. Sedangkan kelemahannya adalah tenaga profesional (biasanya orang luar) tidak dapat diterima oleh orang dalam, keterampilan evaluator dalam mengevaluasi tidak dihargai, kecenderungan menggunakan metode tertentu, dan menghalangi pemilihan metode atau rancangan lain.

#### **D. Peranan Evaluator.**

Evaluator program memiliki berbagai peran yaitu sebagai hakim, terdakwa, juri, pengacara, saksi ahli, detektif, pekerja sosial atau reporter keadilan. Tentu saja pengambilan peran harus disesuaikan dengan waktu, tempat dan jenis tindakannya dalam suatu kegiatan evaluasi yang utuh (Anderson et-al dalam Purwanto dan Suparman, 1999:57).

Evaluator sebagai hakim, peran ini relatif pasif, evaluator tidak aktif mengembangkan rancangan evaluasi dan tidak mengumpulkan data. Evaluator lebih banyak melihat pada informasi yang disajikan orang lain kepadanya, sehingga yang dilakukannya adalah menganalisis dan memikirkan ulang evaluasi yang telah dilaksanakan berdasarkan itu dibuat suatu kesimpulan. Evaluator yang mengasumsikan diri sebagai hakim harus menghindari kesan gegabah atau congkak, ia harus tetap hati-hati, dan tidak membuat orang lain tersinggung dan kurang terhormat.

Terkadang evaluator berperan bagaikan detektif pada saat ia melakukan kegiatan pengumpulan data, misalnya dengan cara mengadakan pengamatan partisipatif. Bahkan evaluator harus bertindak adil dan objektif bagaikan peran seorang hakim atau juri dalam pengadilan, terutama ketika evaluator harus mengemukakan dan melaporkan penilaiannya.

Evaluator program menurut Tayibnapis (2000:136-137) memiliki peranan strategis sebagai berikut:

1. Sebagai penolong dan penasihat terhadap perencana dan pengembang program. Pada waktu program baru mulai dikerjakan, mungkin evaluator akan dipanggil untuk menerangkan dan memonitor kegiatan program. Memeriksa kemajuan dan pencapaian program, perubahan sikap, melihat masalah-masalah yang potensial, dan melihat bagian-bagian yang memerlukan perbaikan. Dalam hal ini evaluator program berperan sebagai evaluator formatif.

2. Mungkin evaluator bertanggungjawab dan bertugas membuat pernyataan singkat tentang pengaruh umum dan pencapaian program. Dalam hal ini evaluator harus menyiapkan laporan tertulis yang harus diserahkan kepada pemimpin atau direktur program. Laporan berisi tentang penjelasan program, pencapaian tujuan umum program, pencatatan hasil-hasil yang diharapkan, dan pembuatan perbandingan dengan program-program alternatif. Dalam hal ini evaluator berperan sebagai evaluator sumatif.

## E. Syarat-Syarat Evaluator

Untuk dapat menjadi evaluator program haruslah memenuhi persyaratan-persyaratan yang ketat. Menurut Schnee sebagaimana dikutip Arikunto (1988:15-17) karakteristik evaluator program sebagai berikut:

1. Evaluator hendaknya merupakan otonom. Evaluator hendaknya orang luar yang sama sekali tidak ada ikatan dengan pengambil kebijakan maupun pengelola dan pelaksana program. Di samping itu juga harus lepas dari tekanan politik.
2. Ada hubungan baik dengan responden dalam arti dapat memahami sedalam-dalamnya watak, kebiasaan dan cara hidup klien yang akan dijadikan sumber data evaluasi.
3. Tanggap akan masalah politik dan sosial karena tujuan evaluasi adalah pengembangan program.
4. Evaluator berkualitas tinggi, dalam arti jauh dari keahlian biasa. Evaluator adalah orang yang mempunyai *self concept* yang tinggi, tidak mudah terombang-ambing.
5. Menguasai teknik untuk memilih desain dan metodologi penelitian yang tepat untuk program yang di evaluasi.
6. Bersikap terbuka terhadap kritik. Untuk mengurangi dan menahan diri dari bias, maka evaluator memberi peluang kepada orang luar untuk melihat apa yang sedang dan telah dilakukan.
7. Menyadari kekurangan dan keterbatasannya serta bersikap jujur, menyampaikan atau menerangkan kelemahan dan keterbatasan tentang evaluasi yang dilakukan.
8. Bersikap pasrah kepada umum mengenai penemuan positif dan negatif. Evaluator harus berpandangan luas dan bersikap tenang apabila menemukan



data yang tidak mendukung program dan berpendapat bahwa penemuan negatif sama pentingnya dengan penemuan positif.

9. Bersedia menyebarkan hasil evaluasi. Untuk program yang penting dan menentukan, hasil evaluasi hanya pantas dilaporkan kepada pengambil keputusan dalam sikap tertutup atau pertemuan khusus. Namun untuk program yang biasa dan dipandang bahwa masyarakat dapat menarik manfaat dari penilaiannya, sebaiknya hasil evaluasi disebarluaskan, khususnya bagi pihak-pihak yang membutuhkan.
10. Hasil penilaian yang tidak secara eksplisit dinyatakan sebagai informasi terbuka, sebaiknya tidak disebarluaskan (merupakan sesuatu yang konfidensial).
11. Tidak mudah membuat kontrak. Evaluator yang tidak memenuhi persyaratan-persyaratan yang telah disebutkan sebaiknya tidak dengan mudah menyanggupi menerima tugas karena secara etis dan moral akan merupakan sesuatu yang kurang dapat dibenarkan.

Persyaratan untuk menjadi seorang evaluator yang kompeten dan dapat diandalkan menurut Tayibnapis (2000:8) harus mempunyai kombinasi berbagai ciri antara lain:

1. Mengetahui dan mengerti teknik pengukuran dan metode penelitian.
2. Mengerti tentang kondisi sosial dan hakikat objek evaluasi.
3. Mempunyai kemampuan *human relation*.
4. Jujur.
5. Bertanggung jawab.

Selanjutnya persyaratan untuk menjadi evaluator dijelaskan oleh Arikunto dan Jabar (2009:22-23) sebagai berikut:

1. Mampu melaksanakan, persyaratan pertama ini harus dipenuhi oleh evaluator adalah individu yang memiliki kemampuan untuk melaksanakan evaluasi yang didukung oleh teori dan ketrampilan praktek.
2. Cermat, dalam hal ini individu yang menjadi evaluator dapat melihat celah-celah dan detail dari program serta bagian program yang akan dievaluasi.
3. Objektif, tidak mudah dipengaruhi oleh keinginan pribadi, agar dapat mengumpulkan data sesuai dengan keadaannya, selanjutnya dapat mengambil kesimpulan sebagaimana diatur oleh ketentuan yang harus diikuti.

4. Sabar dan tekun, agar di dalam melaksanakan tugas dimulai dari membuat rancangan kegiatan dalam bentuk menyusun proposal, menyusun instrumen, mengumpulkan data dan menyusun laporan, tidak gegabah dan tergesa-gesa.
5. Hati-hati dan bertanggung jawab yaitu melakukan pekerjaan evaluasi dengan penuh pertimbangan, namun apabila masih ada kekeliruan yang diperbuat, berani menanggung resiko atas segala kesalahannya.

Di dalam sumber lainnya, Arikunto (1988:14-15) menyatakan syarat untuk menjadi seorang evaluator program yaitu:

### **1. Memahami materi.**

Evaluator memahami tentang seluk-beluk program yang dievaluasi, antara lain:

- a. Tujuan program yang sudah ditentukan sebelum mulai kegiatan.
- b. Komponen-komponen program.
- c. Variabel yang diujicobakan atau dilaksanakan.
- d. Jangka waktu dan penjadwalan kegiatan.
- e. Mekanisme pelaksanaan program.
- f. Pelaksana program
- g. Sistem memonitoring kegiatan program.

### **2. Menguasai teknik.**

Evaluator harus menguasai cara-cara atau teknik yang digunakan di dalam melaksanakan evaluasi program. Oleh karena evaluasi program tidak lain adalah penelitian evaluasi maka evaluator program harus menguasai metodologi penelitian, meliputi:

- a. Cara membuat perencanaan penelitian.
- b. Teknik menentukan populasi dan sampel.
- c. Teknik menyusun instrumen penelitian.
- d. Prosedur dan teknik pengumpulan data.
- e. Penguasaan teknik pengolahan data.
- f. Cara menyusun laporan penelitian.

### **3. Objektif dan cermat.**

Evaluator adalah sekelompok orang yang mengemban tugas penting yang dalam tugasnya ditopang oleh data yang dikumpulkan secara cermat dan objektif. Berdasarkan atas data tersebut mereka diharapkan,

mengklasifikasikan, mentabulasikan, mengolah dan sebagainya secara cermat dan objektif pula. Khususnya di dalam menentukan pengambilan strategi penyusunan laporan, evaluator tidak boleh memandang satu atau dua aspek sebagai hal yang istimewa, dan tidak boleh pula memihak. Baik pelaku evaluasi dari dalam (internal) maupun luar (eksternal), tidak dibenarkan “mengambil muka” dari orang/lembaga yang meminta bantuan atau menugaskannya untuk mengevaluasi.

#### **4. Jujur dan dapat dipercaya.**

Evaluator merupakan tempat pengambil keputusan menumpahkan seluruh kepercayaannya. Mengapa pengambilan keputusan meminta bantuan untuk mengevaluasi program yang dipandang penting untuk dievaluasi? Terdapat dua hal yang menjadi alasannya yaitu:

- a. Mereka menghindari adanya bias (kesalahan pengamatan atau kesalahan persepsi).
- b. Dalam mempertanggungjawabkan tindakannya kepada masyarakat luas, tidak akan ada rasa risih karena adanya kemungkinan tidak jujur.

Atas dasar alasan penyerahan tugas mengevaluasi tersebut kepada evaluator, maka menjadi suatu beban mental yang berat pada tim evaluator untuk tidak menyalahgunakan kepercayaan yang diberikan kepadanya. Sebagai timbal baliknya mereka harus dapat menunjukkan tingkat keterpercayaan yang tinggi kepada pemberi tugas.

Berdasarkan pemaparan di atas maka dapatlah dimaknai bahwa evaluator program haruslah individu yang memiliki syarat-syarat mumpuni, di antaranya mampu melaksanakan, cermat, objektif, sabar dan tekun, serta hati-hati dan bertanggung jawab.

## **F. Kompetensi Evaluator Program**

Evaluator program sebagai orang yang melakukan evaluasi terhadap suatu program, maka sudah barang tentu haruslah memiliki kompetensi untuk melakukan evaluasi. Kompetensi yang harus dimiliki oleh seorang evaluator program menurut Purwanto dan Suparman (1999:55-56) sebagai berikut:

### **1. Kompetensi manajemen.**

Kompetensi manajemen merupakan keterampilan dalam mengelola dan mengendalikan seluruh kegiatan evaluasi sehingga dapat berlangsung sebaik-baiknya, secara efektif dan efisien. Keterampilan manajemen itu terdiri atas sub-sub kompetensi yaitu: melakukan supervisi, menjelaskan wawasan politik, menerapkan etika profesi, keterampilan berkomunikasi, keterampilan interpersonal, analisis sistem, membuat perjanjian atau kontrak, membuat pembiayaan, dan menentukan tujuan. Selain itu masih ada keterampilan tambahan yang perlu juga dikuasai seperti keterampilan mengorganisir, memimpin, mengarahkan dan membimbing staf, terutama untuk kegiatan-kegiatan yang memerlukan tim.

### **2. Kompetensi teknis.**

Kompetensi teknis yaitu ketrampilan melakukan kegiatan evaluasi langkah demi langkah, dari perencanaan sampai selesai tuntas. Keterampilan ini meliputi sub-sub kompetensi yaitu: memilih atau mengembangkan instrumen, mengadministrasikan tes, melakukan analisis statistik, menerapkan metode survey, menerapkan teknik pengamatan, menerapkan psikometri, menerapkan rancangan eksperimen, melakukan kendali mutu data, menggunakan aplikasi komputer, menerapkan metodologi studi kasus, melakukan analisis biaya, membuat intepretasi, membuat rekomendasi dan menulis laporan serta mempresentasikan laporan.

### **3. Kompetensi konseptual.**

Kompetensi konseptual adalah keterampilan tingkat tinggi yang berkaitan dengan kemampuan menganalisis, dan pemecahan masalah. Keterampilan konseptual yang harus dikuasai evaluator meliputi sub-sub kompetensi yaitu menentukan pilihan (alternatif), menyusun rencana awal, meng-kategorikan dan menganalisis masalah, melihat dan menunjukkan hubungan dan membuat kesimpulan.

### **4. Kompetensi bidang ilmu.**

Kompetensi bidang ilmu merupakan keahlian dan kemampuan dalam bidang disiplin ilmu yang terkait dengan evaluasi. Keahlian itu meliputi berpengalaman kerja di bidang yang dievaluasi, berpengetahuan tentang sumber literatur, memahami pentingnya konsepsi dalam bidang yang relevan dan mengenal pakar-pakar di bidangnya.

# BAB III

## MODEL-MODEL EVALUASI PROGRAM

### A. Klasifikasi Model Evaluasi Program

Worther dan Sander (1996:240) melakukan investigasi terhadap model evaluasi program yang pernah ada sejak tahun 1967-1987. Hasil investigasinya menemukan lebih dari 50 model evaluasi. Model-model evaluasi tersebut masing-masing memiliki karakteristik yang berbeda yang membedakan satu model dengan model lainnya sehingga satu model akan lebih tepat diterapkan pada satu *setting* tertentu dengan tujuan tertentu daripada model evaluasi yang lain.

Issac dan Michael (1987:6-7) mengklasifikasikan 6 (enam) model evaluasi program dengan pendekatan dan tujuan yang berbeda antara masing-masing model. Klasifikasi didasarkan atas 12 (duabelas) karakteristik perbedaan dan persamaan dari masing-masing model evaluasi yaitu: definisi, tujuan, penekanan, peran evaluator, keterkaitan dengan tujuan, keterkaitan dengan pembuatan rancangan, tipe evaluasi, konstruk, kriteria penilaian, implikasi terhadap rancangan, kontribusi dan keterbatasan. Klasifikasi 6 (enam) model tersebut adalah:

1. *Goal oriented evaluation model.*  
Evaluasi dilakukan secara berkesinambungan dan kontiniu yang bertujuan untuk menilai sejauhmana program telah tercapai.
2. *Decision oriented evaluation model.*  
Evaluasi diorientasikan untuk memberikan masukan dan pertimbangan dalam pengambilan keputusan.
3. *Transactional evaluation model.*  
Evaluasi ditujukan untuk menggambarkan proses program dan perspektif nilai dari tokoh-tokoh penting dalam masyarakat.

4. *Evaluation research model.*

Evaluasi dilakukan untuk menjelaskan pengaruh kependidikan dan pertimbangan strategi pembelajaran.

5. *Goal-free evaluation model.*

Evaluasi tidak mengacu pada tujuan program, namun fokus mengevaluasi pengaruh program baik yang diharapkan maupun yang tidak diharapkan namun terjadi.

6. *Adversary evaluation model.*

Evaluasi yang bertujuan mengumpulkan kasus-kasus menonjol untuk diinterpretasi nilai program dari dua sisi dengan menggunakan informasi yang sama tentang program.

House (1978:45-48) mengklasifikasikan model evaluasi berdasarkan asumsi filosofis yang mendasari model evaluasi tersebut dikonstruksi. Satu model evaluasi dibedakan dengan model evaluasi lainnya didasarkan atas asumsi dasarnya sehingga bisa dilihat bagaimana model evaluasi secara logis serupa atau berbeda dengan model evaluasi lainnya.

Berdasarkan hasil kajiannya maka House mengklasifikasikan 8 (delapan) model evaluasi yaitu:

1. *System analysis.*

Model evaluasi ini melihat hubungan antar sub-sistem dalam suatu program dan menggunakan pengukuran *output* secara kualitatif.

2. *Behavior objectives.*

Model evaluasi yang tujuannya program dirumuskan ke dalam bentuk perilaku spesifik yang terukur. Tujuan program dirumuskan dalam bentuk perilaku spesifik kemudian diukur dengan tes acuan norma atau tes acuan patokan.

3. *Decision making.*

Model evaluasi ini bertujuan sebagai dasar pembuatan kebijakan. Model ini menggunakan informasi dan data sebagai bahan pertimbangan dalam pengambilan keputusan terhadap program yang dievaluasi.

4. *Goal free.*

Model evaluasi ini tidak diorientasikan untuk mengevaluasi tujuan program namun lebih menitikberatkan evaluasi dampak program berdasar pandangan pihak pengguna program (klien-konsumen).

5. *Art criticism.*

Model evaluasi yang digali dari seni tradisional dan sastra, dengan tujuan melakukan kritik terhadap program yang dikaji.

6. *Accreditation.*

Model ini menggunakan standar eksternal yang diharapkan untuk mengevaluasi suatu program oleh tim profesional dari luar dengan menggunakan standar dari luar pula. Tim evaluator berperan sebagai *reviewer* yang menilai baik-buruk berdasarkan standar dari luar program.

7. *Transaction.*

Model yang memfokuskan evaluasi pada proses pembelajaran di ruang kelas, sekolah atau program yang sedang berlangsung. Model ini banyak menggunakan metode informal dalam melakukan investigasi.

8. *Adversary.*

Model evaluasi yang menekankan pada keragaman argumentasi dari evaluator tentang program yang diteliti.

Kifer (1995) mengklasifikasikan evaluasi program kepada 4 (empat) kelompok model evaluasi yaitu:

## 1. Model evaluasi tradisional.

Model evaluasi ini merupakan model evaluasi pertama yang disebut sebagai evaluasi yang mendasarkan atas konsistensi antara tujuan, aktivitas dan hasil akhir. Model evaluasi ini dikenal dengan *a goal attainment model* yakni tujuan umum dijabarkan secara operasional dalam terminologi khusus dalam bentuk perilaku yang terukur. Tokoh pengembang model evaluasi ini adalah Ralph Tyler pada tahun 1949.

## 2. Model evaluasi studi kasus dan etnografi.

Model evaluasi studi kasus dan etnografi ini menggunakan pendekatan kualitatif, tidak seperti model evaluasi tradisional berorientasi pada kebijakan dengan menggunakan kuantitatif. Model evaluasi ini menekankan pada pemahaman tentang evaluasi respon seseorang atas berbagai hal yang diminati, seringkali disebut *stakeholder* evaluasi dengan menggunakan metode antropologi untuk mengumpulkan fakta tentang objek yang dievaluasi.

Karakteristik pendekatan ini melibatkan *observer participant* yang menanyakan kepada informan kunci tentang apa yang terjadi, memberikan

informasi tentang program yang diimplementasikan. Model evaluasi ini berpandangan bahwa pengumpulan data didasarkan kepercayaan terhadap persepsi dan pengalaman observer. Model evaluasi studi kasus dan etnografi ini dikembangkan oleh Stake pada tahun 1977.

3. Model evaluasi *goal free* dan integratif.

Model evaluasi *goal free* (tanpa tujuan atau bebas tujuan) ini mengkaji semua dampak atau hasil akhir secara integratif. Model evaluasi *goal free* dan integratif ini dikembangkan oleh Scriven pada tahun 1983.

4. Model evaluasi berorientasi kebijakan.

Model evaluasi ini mengkaji seluruh aspek yang terdapat objek yang dikaji, dalam hal ini informasi dan data diperoleh dari berbagai sumber untuk membuat kebijakan. Model evaluasi ini seringkali disebut dengan istilah CIPP (*context, input, process, product*).

Selanjutnya klasifikasi model evaluasi program ditinjau dari maksud dan tujuannya menurut Purwanto dan Suparman (1999:16-17) maka dikelompokkan menjadi 6 (enam) kelompok yaitu:

1. Evaluasi berorientasi tujuan (*goal-oriented evaluation*). Tujuan dari evaluasi ini pada tujuan untuk melakukan pengukuran terhadap kemajuan dan efektivitas inovasinya. Hasil evaluasi tersebut menunjukkan seberapa tinggi hasil belajar yang dicapai peserta setelah mengikuti program yang ditentukan. Tokoh evaluasi ini adalah Bloom dan Provus.
2. Evaluasi berorientasi keputusan (*decision-oriented evaluation*). Tujuan dari evaluasi adalah menghasilkan rekomendasi bagi pembuat keputusan tentang apa yang harus dilakukan oleh pengambil keputusan sehubungan dengan program yang dievaluasi. Tokoh evaluasi ini adalah Stufflebeam.
3. Evaluasi transaksional. Model ini biasanya terkonsentrasi pada proses pendidikan/program itu sendiri dan menggunakan berbagai metode informal dalam investigasi dan menggunakan studi kasus sebagai metode utama. Salah satu evaluasi transaksional ini adalah pendekatan responsif dalam evaluasi atau *responsive approach to evaluation Stake's*, termasuk dalam evaluasi transaksional ini adalah *contentance model* yang juga dikembangkan oleh Stake. Evaluasi kasus sebagai suatu unik dan didasarkan kepada persepsi dan pengetahuan evaluator dan pelaksanaannya cenderung menggunakan pendekatan naturalistik. Tokoh dalam model ini adalah Stake dan Rippey.



4. Penelitian evaluasi. Evaluasi model ini berfokus pada upaya untuk memperoleh penjelasan tentang pengaruh pendidikan dan pelatihan terhadap perbaikan kinerja individu atau organisasi. Penjelasan tentang pengaruh tersebut harus didasarkan kepada kajian teori ilmiah. Digunakannya kajian teori ini menjadi ciri khas *evaluation research*. Tokoh penelitian evaluasi ini adalah Campbell dan Colley.
5. Evaluasi bebas tujuan (*goal free evaluation*). Evaluasi harus mengukur pengaruh program dan didasarkan pada kriteria program. Secara esensial evaluasi diartikan sebagai pengumpulan data secara umum tentang pengaruh aktual.

Evaluasi juga menilai pentingnya pengaruh tersebut dalam mencapai kebutuhan yang ditentukan. Ada empat alasan untuk melakukan evaluasi bebas tujuan yaitu:

- a. Untuk menghindari resiko dari keterbatasan tujuan program dan menghindari hilangnya resiko dari keterbatasan tujuan program dan menghindari hilangnya hasil-hasil kegiatan yang tidak terantisipasi.
- b. Untuk mengubah konotasi negatif dari dampak yang tidak dikehendaki.
- c. Untuk mengurangi bias pemikiran dalam evaluasi.
- d. Menjaga objektivitas dan independensi evaluator.

Tokoh dalam model evaluasi ini adalah Sriven.

6. Evaluasi *adversary* (*adversary models of evaluation*). Evaluasi harus menampilkan kasus terbaik bagi setiap permasalahan yang timbul dalam program. Oleh karena itu evaluasi ini menggunakan berbagai jenis sumber data, dan berusaha menggali penilaian berbagai pihak tentang segi positif dan negative dari program. Tokoh model evaluasi ini adalah Levine dan Owens.

Berdasarkan pendekatan dalam melakukan evaluasi, maka model evaluasi program diklasifikasikan oleh Brinkerhoff et-al kepada 9 (sembilan) kelompok sebagaimana dikutip Purwanto dan Suparman (1999:17-18) sebagai berikut:

1. Pendekatan kesepadanan dan ketaatan atau *congruency and compliance*. Dalam pendekatan ini kemajuan program dan aktivitas dicatat dan dibandingkan dengan rencana (desain, maksud/tujuan), beberapa standar eksternal atau kriteria. Tujuannya antara lain untuk membantu manajemen memelihara jalannya program agar sesuai dengan aturan,

mendokumentasi bahwa rencana dan proposal telah memadai, mendemonstrasikan dan memenuhi ketentuan atau aturan. Model yang relevan dengan pendekatan ini antara lain: model evaluasi kesenjangan, *program evaluation and review technique* (PERT) dan *management by objective* (MBO).

2. Pendekatan pembuatan keputusan (*decision making*). Dalam pendekatan ini informasi dikumpulkan dengan sebaik-baiknya dan selengkap mungkin agar dapat diolah dan dianalisis sehingga dapat dijadikan dasar bagi kegiatan pembuatan keputusan.
3. Pendekatan responsif (*responsive*). Menurut pendekatan ini evaluasi harus mampu menjawab permasalahan yang muncul atau yang diprediksi akan muncul dalam kegiatan.
4. Pendekatan *objectives based*. Termasuk dalam kelompok ini adalah evaluasi yang dikembangkan oleh Popham's yaitu *instructional objectives approach*.
5. Pendekatan *naturalistic*. Jenis-jenis evaluasi yang termasuk paling cocok dengan pendekatan ini adalah evaluasi transaksional, evaluasi bebas tujuan dan *adversaty evaluation*.
6. Pendekatan *expert judgement*. Berdasarkan pendekatan ini evaluasi harus dilaksanakan oleh evaluator yang benar-benar memiliki kompetensi dan kemampuan dibidangnya.
7. Pendekatan eksperimental. Data tentang hasil secara hati-hati dicari dan diukur di bawah kondisi perlakuan yang terkontrol, sesudah menggunakan kelompok kontrol atau metode statistik untuk mengukur dan mengontrol kesalahan. Tujuan adalah untuk membandingkan pengaruh suatu pendekatan dengan yang lain, mendemonstrasikan hubungan sebab-akibat, memberikan bukti bahwa program mempengaruhi hasil, mengidentifikasi keterkaitan dan hubungan di antara variabel-variabel kunci dalam program, dan memvalidasi perolehan dari program. Model yang relevan adalah Campbell dan Stanley.
8. Pendekatan *cost analysis*. Pembiayaan program ditetapkan dan dianalisis untuk menentukan jumlah yang dialokasikan untuk kegiatan apa dan untuk tujuan mana? Tujuan; keterkaitan antara peningkatan hasil dengan peningkatan biaya, memfasilitasi upaya-upaya replikasi. Model ini adalah *cost effectiveness analysis* dari Levin's.

9. Pendekatan pengembangan organisasi (*organizational development*). Informasi tentang staf dan masalah-masalah proyek, harapan-harapan dan kemajuan secara regular dikumpulkan, kemudian dikembalikan kepada staf. Tujuan evaluasi dengan pendekatan ini adalah untuk membantu meningkatkan pengetahuan dan kepastian tentang apa yang terjadi bagaimana kejadiannya dan mengapa, membantu staf agar lebih efektif, produktif dan puas, mengidentifikasi kebutuhan pengembangan staf dan pengembangan organisasi, memfasilitasi pertumbuhan staf dan proyek. Model evaluasi ini adalah *discrepancy evaluation* dari Provus, *pelatihan evaluation* dari Brinkerhoff dan *action research* dari Shumsky.

## B. Model-Model Evaluasi Program

Terkait dengan model-model evaluasi program maka dalam kajian literatur terdapat berbagai ragam model evaluasi yang dapat digunakan oleh evaluator sebagai acuan dalam melakukan evaluasi suatu program. Di antara model-model evaluasi program tersebut diantaranya: Goal-Free Evaluation Approach (Scriven), Formative and Summative model (Scriven), Five level ROI Model (Jack Phillips), Context, Input, Process, Product atau CIPP Model (Stufflebeam), Four levels evaluation model (Kirkpatrick), Responsive evaluation model (Stake), Context, Input, Reaction, Outcome atau CIRO model, Congruence-Contingency model (Stake), Five Levels of Evaluation model (Kaufmann), Program Evaluation and Review Technique atau PERT model, Alkin model, CSE-UCLA Model, Provus Discrepancy model, Illuminative evaluation model dan lainnya.

Untuk memilih berbagai model evaluasi program kiranya pendekatan *eclectic* dapat dijadikan rujukan. Pendekatan *eclectic* yaitu memilih berbagai model dari beberapa pilihan yang terbaik sesuai dengan kebutuhan, sesuai dengan situasi dan sesuai dengan kondisi setempat (Tayibnapis, 2000:7).

Pemilihan suatu model evaluasi akan tergantung pada kemampuan evaluator, tujuan evaluasi serta untuk siapa evaluasi itu dilaksanakan. Sistem evaluasi yang dilakukan harus difokuskan dengan jelas pada proses perbaikan daripada pertanggungjawaban untuk produk akhir. Sistem ini harus dioperasikan dekat dengan titik intervensi (obyek dalam hal ini program) untuk perubahan.

## 1. Stufflebeam's Model (CIPP Model).

Model ini dikembangkan oleh Stufflebeam, model CIPP yang merupakan sebuah singkatan dari huruf awal empat buah kata, yaitu *Context*, *Input*, *Process*, and *Product*. Keempat kata yang disebutkan dalam singkatan CIPP tersebut merupakan sasaran evaluasi, yang tidak lain adalah komponen dari proses sebuah program kegiatan. Dengan kata lain, model CIPP adalah model evaluasi yang memandang program yang dievaluasi sebagai sebuah sistem.

Keunikan model ini adalah pada setiap tipe evaluasi terkait pada perangkat pengambil keputusan (*decission*) yang menyangkut perencanaan dan operasional sebuah program. Keunggulan model CIPP memberikan suatu format evaluasi yang komprehensif/menyeluruh pada setiap tahapan evaluasi yaitu tahap konteks, masukan, proses, dan produk.

Model CIPP ini bertitik tolak pada pandangan bahwa keberhasilan program pendidikan dipengaruhi oleh berbagai faktor, seperti: karakteristik peserta didik dan lingkungan, tujuan program dan peralatan yang digunakan, prosedur dan mekanisme pelaksanaan program itu sendiri. Dalam hal ini Stufflebeam melihat tujuan evaluasi sebagai:

- a. Penetapan dan penyediaan informasi yang bermanfaat untuk menilai keputusan alternatif.
- b. Membantu audience untuk menilai dan mengembangkan manfaat program pendidikan atau obyek.
- c. Membantu pengembangan kebijakan dan program.

Aspek yang dievaluasi dan prosedur pelaksanaan evaluasi model CIPP menurut Stufflebeam dalam Oliva (1992:491) dapat dilihat pada tabel berikut:

**Tabel III. Aspek dan Prosedur Pelaksanaan Evaluasi**

Aspek	Context	Input	Process	Output
Objek (sasaran)	Mendefinisikan operasional <i>context</i> , mengidentifikasi dan memperkirakan kebutuhan dan mendiagnosa masalah, memprediksi kebutuhan dan peluang	Mengidentifikasi dan memperkirakan kapabilitas sistem, strategi input yang sekarang tersedia, dan mendesain untuk implementasi strategi	Mengidentifikasi dan memperkirakan di dalam proses tentang kerusakan di dalam desain prosedur atau implementasi, menyediakan informasi sebelum program diputuskan dan memperbaiki dokumen even prosedural dan aktivitas	Menghubungkan informasi outcomes dengan obyek dan informasi <i>context</i> , <i>input</i> , dan <i>process</i>
Metode	Mendeskripsikan <i>context</i> , membandingkan dengan yang sebenarnya dan mengawasi input dan output, membandingkan kemungkinan dan ketidakmungkinan sistem kerja, dan menganalisa penyebab ketidakmungkinan dan ketidaksesuaian kenyataan dengan tujuan (harapan)	Mendeskripsikan dan menganalisis SDM dan sumber daya material yang tersedia, solusi strategis, dan desain prosedur untuk relevansi, kemungkinan kegiatan yang dapat dilaksanakan, dan kebutuhan ekonomi dalam rangkaian kegiatan	Memonitoring setiap aktivitas yang berpotensi terdapat tantangan secara prosedural, dan memberikan tanda untuk antisipasi, untuk memperoleh informasi yang spesifik untuk memutuskan suatu program, dan mendeskripsikan proses yang aktual	Mendefinisikan operasional dan mengukur kriteria asosiasi dengan obyektif dan membandingkan hasil pengukuran dengan standar sebelum dilakukan antisipasi, dan menginterpretasi outcomes berdasarkan dokumen informasi <i>context</i> , <i>input</i> , dan <i>process</i>
Hubungan pengambilan keputusan dengan proses perubahan	Memutuskan dalam hal menyajikan perangkat, tujuan asosiasi, dengan mendiskusikan kebutuhan dan peluang, dan sasaran asosiasi untuk perubahan perencanaan kebutuhan	Memilih SDM sebagai pendukung, solusi strategis, dan desain prosedural untuk perubahan struktur kerja (aktivitas)	Untuk implementasi dan memperbaiki desain program dan prosedur untuk keefektifan proses kontrol	Untuk memutuskan dalam kegiatan secara kontinu, menghentikan (mengakhiri), modifikasi, mengatur kembali fokus perubahan aktivitas dengan tahapan materi yang lain dalam proses perubahan untuk mengatur kembali aktivitas perubahan

Empat aspek dalam model evaluasi CIPP yaitu *context*, *input*, *process*, dan *output* membantu pengambil keputusan untuk menjawab empat pertanyaan dasar mengenai:

- 1) Apa yang harus dilakukan (*What should we do?*) mengumpulkan dan menganalisa *need assessment* data untuk menentukan tujuan, prioritas dan sasaran.
- 2) Bagaimana kita melaksanakannya (*How should we do it?*) sumber daya dan langkah-langkah yang diperlukan untuk mencapai sasaran dan tujuan dan mungkin meliputi identifikasi program eksternal dan material dalam mengumpulkan informasi.
- 3) Apakah dikerjakan sesuai rencana (*Are we doing it as planned?*) Ini menyediakan informasi bagi pengambil keputusan tentang seberapa baik program diterapkan. Dengan secara terus-menerus monitoring program, pengambil keputusan mempelajari seberapa baik pelaksanaan telah sesuai petunjuk dan rencana, konflik yang timbul, dukungan staff dan moral, kekuatan dan kelemahan material, dan permasalahan penganggaran.
- 4) Apakah berhasil (*Did it work?*); Dengan mengukur *outcome* dan membandingkannya pada hasil yang diharapkan, pengambil-keputusan menjadi lebih mampu memutuskan jika program harus dilanjutkan, dimodifikasi, atau dihentikan sama sekali.

#### **a. Context Evaluation (Evaluasi Konteks)**

Evaluasi konteks membantu merencanakan keputusan, menentukan kebutuhan yang akan dicapai oleh program dan merumuskan tujuan program (Tayibnapis, 2000:14). Tujuan evaluasi konteks yang utama adalah untuk mengetahui kekuatan dan kelemahan yang dimiliki. Dengan mengetahui kekuatan dan kelemahan ini, evaluator akan dapat memberikan arah perbaikan yang diperlukan.

Evaluasi konteks juga terkait dengan upaya untuk menggambarkan dan merinci lingkungan kebutuhan yang tidak terpenuhi, populasi dan sampel yang dilayani, dan tujuan proyek (Arikunto dan Jabar, 2009:48). Lebih lanjut keduanya mengilustrasikan dengan memberi contoh evaluasi program makanan tambahan anak sekolah (PMTAS) dalam pengajuan pertanyaan evaluasi sebagai berikut:

- a. Kebutuhan apa saja yang belum terpenuhi oleh program, misalnya jenis makanan dan siswa yang belum menerima?
- b. Tujuan pengembangan apakah yang belum tercapai oleh program, misalnya peningkatan kesehatan dan prestasi siswa karena adanya makanan tambahan?
- c. Tujuan pengembangan apakah yang dapat membantu mengembangkan masyarakat, misalnya kesadaran orang tua untuk memberikan makanan bergizi kepada anak-anaknya?
- d. Tujuan-tujuan manakah yang paling mudah dicapai, misalnya pemerataan makanan, ketepatan penyediaan makanan?

### **b. *Input Evaluation* (Evaluasi Masukan)**

Evaluasi masukan membantu mengatur keputusan, menentukan sumber-sumber yang ada, alternatif apa yang diambil, apa rencana dan strategi untuk mencapai tujuan, dan bagaimana prosedur kerja untuk mencapainya. Pertanyaan yang berkenaan dengan masukan mengarah pada pemecahan masalah yang mendorong diselenggarakannya program yang bersangkutan. Dalam hal ini komponen evaluasi masukan meliputi: (1) sumber daya manusia, (2) sarana dan peralatan pendukung, (3) dana atau anggaran, dan (4) berbagai prosedur dan aturan yang diperlukan.

Melanjutkan contoh sebelumnya yang telah dikemukakan yaitu program PMTAS maka dalam hal ini pertanyaan-pertanyaan yang dapat diajukan pada tahap evaluasi masukan ini adalah:

- a. Apakah makanan yang diberikan kepada siswa berdampak jelas pada perkembangan siswa?
- b. Berapa orang siswa yang menerima dengan senang hati atas makanan tambahan itu?
- c. Bagaimana reaksi siswa terhadap pelajaran setelah menerima makanan tambahan?
- d. Seberapa tinggi kenaikan nilai siswa setelah menerima makanan tambahan?

### **c. *Process Evaluation* (Evaluasi Proses)**

Evaluasi proses menekankan pada tiga tujuan: *(1) do detect or predict in procedural design or its implementation during implementation stage,*

(2) *to provide information for programmed decision*, dan (3) *to maintain a record of the procedure as it occurs*.

Penjelasan di atas bermakna bahwa evaluasi proses digunakan untuk mendeteksi atau memprediksi rancangan prosedur atau rancangan implementasi selama tahap implementasi, menyediakan informasi untuk keputusan program dan sebagai rekaman atau arsip prosedur yang telah terjadi. Evaluasi proses meliputi koleksi data penilaian yang telah ditentukan dan diterapkan dalam praktik pelaksanaan program.

Pada dasarnya evaluasi proses untuk mengetahui sampai sejauh mana rencana telah diterapkan dan komponen apa yang perlu diperbaiki. Evaluasi proses dalam model CIPP menunjuk pada “apa” (*what*) kegiatan yang dilakukan dalam program, “siapa” (*who*) orang yang ditunjuk sebagai penanggung jawab program, “kapan” (*when*) kegiatan akan selesai. Dalam model CIPP, evaluasi proses diarahkan pada seberapa jauh kegiatan yang dilaksanakan di dalam program sudah terlaksana sesuai dengan rencana (Arikunto dan Jabar, 2009:47).

Pertanyaan-pertanyaan yang diajukan pada untuk evaluasi proses menurut Stufflebeam sebagai berikut:

- a. Apakah pelaksanaan program sesuai dengan jadwal?
- b. Apakah staf yang terlibat didalam pelaksanaan program akan sanggup menangani kegiatan selama program berlangsung dan kemungkinan jika dilanjutkan?
- c. Apakah sarana dan prasarana yang disediakan dimanfaatkan secara maksimal?
- d. Hambatan-hambatan apa saja yang dijumpai selama pelaksanaan program dan kemungkinan jika program dilanjutkan?

#### **d. Product Evaluation (Evaluasi Produk/Hasil)**

Evaluasi produk/hasil adalah: *to allow to project director (or teacher) to make decision of program*. Evaluasi produk diharapkan dapat membantu pimpinan proyek atau guru untuk membuat keputusan yang berkenaan dengan kelanjutan, akhir, maupun modifikasi program. Menurut Tayibnapis (2000:14) evaluasi produk untuk membantu membuat keputusan selanjutnya, baik mengenai hasil yang telah dicapai maupun apa yang dilakukan setelah program itu berjalan.



Stufflebeam dalam naskah yang dipresentasikan pada Annual Conference of the Oregon Program Evaluation Network (OPEN) Portland tahun 2003, memperluas makna evaluasi *product* menjadi *impact evaluation* (evaluasi pengaruh), *effectiveness evaluation* (evaluasi keefektifan), *sustainability evaluation* (evaluasi keberlanjutan), dan *transportability evaluation* (evaluasi transformasi) (Stufflebeam, 2003:59-62)

Berdasarkan penjelasan di atas maka dapat ditarik kesimpulan bahwa, evaluasi produk merupakan penilaian yang dilakukan guna untuk melihat ketercapaian/ keberhasilan suatu program dalam mencapai tujuan yang telah ditentukan sebelumnya. Pada tahap evaluasi inilah seorang evaluator dapat menentukan atau memberikan rekomendasi kepada evaluasi apakah suatu program dapat dilanjutkan, dikembangkan/ modifikasi, atau bahkan dihentikan.

Pada tahap evaluasi ini diajukan pertanyaan evaluasi sebagai berikut:

- a. Apakah tujuan-tujuan yang ditetapkan sudah tercapai?
- b. Pernyataan-pernyataan apakah yang mungkin dirumuskan berkaitan antara rincian proses dengan pencapaian tujuan?
- c. Dalam hal apakah berbagai kebutuhan siswa sudah dapat dipenuhi selama proses pemberian makanan tambahan (misalnya variasi makanan, banyaknya ukuran makanan, dan ketepatan waktu pemberian)?
- d. Apakah dampak yang diperoleh siswa dalam waktu yang relatif panjang dengan adanya program makanan tambahan ini?

Model evaluasi CIPP lebih komprehensif diantara model evaluasi lainnya, karena objek evaluasi tidak hanya pada hasil semata tetapi juga mencakup konteks, masukan, proses, dan hasil. Selain kelebihan tersebut, di satu sisi model evaluasi ini juga memiliki keterbatasan, antara lain penerapan model ini dalam bidang program pembelajaran di kelas mempunyai tingkat keterlaksanaan yang kurang tinggi jika tidak adanya modifikasi.

Dalam konteks pendidikan, Sudjana dan Ibrahim (2004:246) menerjemahkan masing-masing dimensi CIPP tersebut dengan makna:

- a. *Context*, merupakan situasi atau latar belakang yang mempengaruhi jenis-jenis tujuan dan strategi pendidikan yang akan dikembangkan dalam sistem yang bersangkutan, situasi ini merupakan faktor eksternal, seperti misalnya masalah pendidikan yang dirasakan, keadaan ekonomi negara, dan pandangan hidup masyarakat.

- b. *Input*, menyangkut sarana, modal, bahan, dan rencana strategi yang ditetapkan untuk mencapai tujuan pendidikan, komponen input meliputi siswa, guru, desain, saran, dan fasilitas.
- c. *Process*, merupakan pelaksanaan strategi dan penggunaan sarana, modal, dan bahan di dalam kegiatan nyata di lapangan, komponen proses meliputi kegiatan pembelajaran, pembimbingan, dan pelatihan.
- d. *Product*, merupakan hasil yang dicapai baik selama maupun pada akhir pengembangan sistem pendidikan yang bersangkutan, komponen produk meliputi pengetahuan, kemampuan, dan sikap (siswa dan lulusan).

## 2. Model Kirkpatrick

Model evaluasi yang dikembangkan oleh Kirkpatrick dikenal dengan istilah “*Kirkpatrick four levels evaluation model*”. Model Kirkpatrick ini mengevaluasi program pelatihan. Evaluasi terhadap efektivitas program pelatihan menurut Kirkpatrick mencakup empat level evaluasi yaitu:

- a. Level 1: *reaction*.
- b. Level 2: *learning*.
- c. Level 3: *behaviour*.
- d. Level 4: *result*.

Berikut penjelasan dari masing-masing level sebagaimana yang dimaksudkan oleh Kirkpatrick.

### a. Level 1: *reaction* atau *evaluating reaction*.

*Evaluating reaction* atau mengevaluasi terhadap reaksi peserta pelatihan adalah aktivitas mengukur kepuasan peserta (*customer satisfaction*) terhadap program pelatihan yang dilaksanakan. Suatu program pelatihan dianggap efektif apabila proses pelatihan dirasakan menyenangkan dan memuaskan bagi peserta pelatihan sehingga peserta pelatihan tertarik dan termotivasi untuk belajar dan berlatih.

Dengan kata lain peserta pelatihan akan termotivasi apabila proses pelatihan berjalan dan memuaskan bagi peserta pelatihan yang pada akhirnya akan memunculkan reaksi dari peserta pelatihan yang menyenangkan. Sebaliknya apabila peserta pelatihan tidak merasa puas terhadap proses pelatihan yang diikutinya maka peserta pelatihan tidak akan termotivasi untuk mengikuti pelatihan lebih lanjut.

Dengan demikian dapatlah dimaknai bagi keberhasilan proses kegiatan pelatihan tidak terlepas dari minat, perhatian, dan motivasi peserta pelatihan dalam mengikuti jalannya kegiatan pelatihan. Peserta pelatihan akan belajar lebih baik manakala memberikan reaksi positif terhadap lingkungan belajar yang diperolehnya dalam kegiatan pelatihan.

Kepuasan peserta pelatihan terhadap kegiatan pelatihan dapat dikaji dari beberapa aspek yaitu:

- a. Materi yang diberikan.
- b. Fasilitas yang tersedia.
- c. Strategi penyampaian materi yang digunakan instruktur.
- d. Media pembelajaran yang tersedia.
- e. Jadwal kegiatan pelatihan.
- f. Menu dan penyajian konsumsi yang diberikan kepada peserta pelatihan.

Mengukur reaksi peserta pelatihan dapat dilakukan dengan *reaction sheet* dalam bentuk angket sehingga lebih mudah dan lebih efektif dalam menjangkau reaksi peserta pelatihan terhadap program pelatihan.

#### **b. Level 2: *learning atau evaluating leaning.***

Menurut Kirkpatrick (1998:20) *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.* Dalam hal ini terdapat tiga hal yang dapat instruktur ajarkan dalam program yaitu pengetahuan, sikap maupun keterampilan. Peserta dikatakan telah belajar apabila pada dirinya telah mengalami perubahan sikap, perbaikan pengetahuan, maupun peningkatan keterampilan.

Peserta pelatihan dikatakan telah belajar apabila pada dirinya telah mengalami perubahan sikap, perbaikan pengetahuan maupun peningkatan ketrampilan. Oleh karena itu, untuk mengukur efektifitas program pelatihan maka ketiga aspek tersebut perlu untuk diukur. Tanpa adanya perubahan sikap, peningkatan pengetahuan maupun perbaikan keterampilan pada peserta pelatihan maka program pelatihan dapat dikatakan gagal.

Penilaian *evaluating learning* ini ada yang menyebutnya dengan penilaian hasil (*output*) belajar. Oleh karena itu, dalam pengukuran hasil belajar (*learning measurement*) dari peserta pelatihan berarti penentuan satu atau lebih hal-hal berikut:

- a. Pengetahuan apa yang telah dipelajari.
- b. Sikap apa yang telah berubah.
- c. Keterampilan apa yang telah dikembangkan atau diperbaiki dari peserta didik.

**c. Level 3: *behaviour* atau *evaluating behaviour*.**

Evaluasi pada level 3 atau evaluasi tingkah laku berbeda dengan evaluasi terhadap sikap pada level 2. Penilaian sikap pada evaluasi 2 difokuskan pada perubahan sikap yang terjadi pada diri peserta pelatihan saat kegiatan pelatihan berlangsung dilakukan sehingga lebih bersifat internal, sedangkan penilaian tingkah laku pada level 3 ini difokuskan pada perubahan tingkah laku setelah peserta pelatihan kembali ke tempat kerjanya.

Dalam hal ini apakah perubahan sikap yang telah terjadi setelah peserta pelatihan mengikuti pelatihan juga akan diimplementasikan setelah peserta pelatihan kembali ke tempat kerja, sehingga penilaian tingkah laku ini lebih bersifat eksternal. Perubahan perilaku apa yang terjadi di tempat kerja setelah peserta pelatihan mengikuti program pelatihan.

Dengan kata lain yang perlu dinilai adalah apakah peserta pelatihan merasa senang setelah mengikuti pelatihan dan kembali ke tempat kerja. Bagaimana peserta dapat mentransfer pengetahuan, sikap dan keterampilan yang diperoleh selama pelatihan untuk diimplementasikan ditempat kerjanya.

Oleh karena yang dinilai adalah perubahan perilaku peserta pelatihan setelah kembali ke tempat kerja maka evaluasi pada level 3 ini dapat disebut dengan evaluasi terhadap *outcome* dari kegiatan pelatihan.

**d. Level 4: *result* atau *evaluating result*.**

*Evaluating result* atau evaluasi hasil dalam level 4 ini difokuskan pada hasil akhir (*final result*) yang terjadi karena peserta pelatihan telah mengikuti suatu program pelatihan. Termasuk dalam kategori evaluasi hasil akhir dari suatu program pelatihan ini diantaranya adalah:

- a. Kenaikan produksi.
- b. Peningkatan kualitas.
- c. Penurunan biaya.
- d. Penurunan kuantitas terjadinya kecelakaan kerja.
- e. Penurunan *turnover*.
- f. Kenaikan laba.

Beberapa program pelatihan mempunyai tujuan meningkatkan moral kerja maupun membangun *teamwork* yang lebih baik. Dengan kata lain evaluasi level 4 ini adalah evaluasi terhadap *impact* program pelatihan. Tidak semua *impact* dari sebuah program pelatihan dapat diukur dan juga membutuhkan waktu yang cukup lama. Oleh karena itu evaluasi level 4 ini lebih sulit dibandingkan dengan evaluasi pada level-level sebelumnya.

Selanjutnya Kirkpatrick menjelaskan bahwa untuk memperoleh gambaran yang komprehensif terkait dengan evaluasi terhadap program pelatihan maka sekurang-kurangnya terdapat tiga komponen yang dijadikan objek evaluasi yaitu: (1) desain program pelatihan, (2) implementasi program pelatihan, dan (3) hasil yang akan dicapai.

1. Desain program pelatihan.

Desain program pelatihan di evaluasi dari aspek tujuan yang ingin dicapai ataupun kompetensi peserta pelatihan yang akan dikembangkan, strategi pembelajaran yang akan diterapkan pada kegiatan pelatihan, dan isi/materi program pelatihan.

a. Kompetensi yang akan dikembangkan.

Salah satu aspek dari program pelatihan yang dijadikan objek evaluasi adalah kompetensi peserta pelatihan yang akan dikembangkan, khususnya kompetensi dasar dari materi pelatihan yang bersangkutan. Ada beberapa kriteria yang dapat digunakan untuk menilai kompetensi dasar yang akan dikembangkan yaitu antara lain:

- 1) Menunjang pencapaian kompetensi standar kompetensi maupun kompetensi lulusan.
- 2) Jelas rumusan yang digunakan (*observable*) yaitu mampu menggambarkan dengan jelas perubahan tingkah laku yang diharapkan peserta pelatihan.
- 3) Mempunyai kesesuaian dengan tingkat perkembangan peserta pelatihan.

b. Strategi pembelajaran.

Ada beberapa kriteria yang dapat digunakan untuk menilai strategi pembelajaran yang direncanakan pada proses pelaksanaan pelatihan yaitu antara lain:

- 1) Kesesuaian dengan kompetensi yang akan dikembangkan pada diri peserta pelatihan.

- 2) Kesesuaian dengan kondisi belajar yang diinginkan pada pelaksanaan pelatihan.
- 3) Kejelasan rumusan, terutama mencakup aktivitas instruktur maupun peserta pelatihan dalam proses pelatihan.

c. Isi/materi program pelatihan.

Isi atau materi program pelatihan yang dimaksud adalah pengalaman belajar yang akan disiapkan oleh instruktur maupun yang harus diikuti peserta program pelatihan.

Ada beberapa kriteria yang dapat digunakan untuk menilai isi program pelatihan yaitu antara lain:

- 1) Relevansi dengan kompetensi yang akan dikembangkan pada peserta pelatihan.
- 2) Relevansi dengan pengalaman peserta pelatihan dan lingkungan.
- 3) Kesesuaian dengan tingkat perkembangan peserta pelatihan.
- 4) Kesesuaian dengan alokasi waktu yang tersedia.
- 5) Keautentikan pengalaman dengan lingkungan kerja atau institusi peserta pelatihan.

2. Implementasi program pelatihan.

Selain desain program pelatihan, proses implementasi program atau proses pelaksanaannya perlu dijadikan objek evaluasi, khususnya proses belajar dan pembelajaran yang berlangsung di dalam kegiatan pelaksanaan pelatihan.

Beberapa kriteria yang dapat digunakan untuk mengevaluasi proses belajar dan pembelajaran pada kegiatan pelatihan yaitu:

- 1) Konsistensi dengan kegiatan yang terdapat dalam program pelatihan dengan berbagai hal seperti konsistensi materi pelatihan dengan tujuan pelatihan.
- 2) Keterlaksanaan oleh instruktur.
- 3) Keterlaksanaan dari segi peserta pelatihan.
- 4) Perhatian yang diperlihatkan pada peserta pelatihan terhadap pembelajaran yang sedang berlangsung.
- 5) Keaktifan para peserta pelatihan dalam proses belajar.
- 6) Kesempatan yang diberikan untuk menerapkan hasil pembelajaran dalam situasi yang nyata.

- 7) Pola interaksi antara instruktur dan peserta pelatihan.
- 8) Kesempatan untuk mendapatkan umpan balik secara kontiniu dari peserta pelatihan.

3. Hasil yang akan dicapai.

Selain desain program dan implementasi pelatihan, maka komponen ketiga yang perlu dievaluasi adalah hasil yang akan dicapai oleh kegiatan pelatihan. Hasil yang dicapai ini dapat mengacu pada pencapaian tujuan jangka pendek (*output*) maupun mengacu pada pencapaian tujuan jangka panjang (*outcome*).

*Outcome* program pelatihan ini akan dinilai seberapa jauh peserta pelatihan mampu mengimplementasikan kompetensi yang dipelajari sewaktu mengikuti pelatihan ke dalam dunia nyata dalam memecahkan persoalan di tempat kerjanya.

Untuk mengevaluasi keberhasilan program pelatihan tidak cukup hanya dengan mengadakan penilaian terhadap hasil belajar peserta pelatihan saja sebagai produk dari sebuah proses pelatihan. Kualitas suatu produk pelatihan tidak terlepas dari kualitas proses pembelajaran pada pelatihan itu sendiri.

Evaluasi terhadap program pelatihan yang disusun dan dilaksanakan instruktur sebaiknya menjangkau penilaian terhadap program pelatihan yang menyeluruh meliputi:

1. Desain pelatihan yang meliputi kompetensi yang dikembangkan untuk peserta training, strategi pembelajaran yang dipilih dalam pelaksanaan pelatihan dan isi/materi program pelatihan
2. Implementasi program pembelajaran atau kualitas pembelajaran pada pelaksanaan pelatihan
3. Hasil program pelatihan.

Dalam mengadakan penilaian terhadap hasil program pelatihan tidak cukup terbatas pada hasil jangka pendek atau *output* tetapi sebaiknya juga menjangkau *outcome* dari program pelatihan tersebut.

### 3. Model Scriven

Michael Scriven lahir di 28 Maret 1928, di Beaulieu, Hampshire, Inggris 1928. Gelar pertamanya dalam bidang matematika dan gelar doktor dalam

filsafat. Scriven telah membuat kontribusi yang signifikan di bidang filsafat, psikologi, berpikir kritis, dan yang paling terutama, evaluasi (dia telah menciptakan sebuah penemuan untuk evaluasi program).

Scriven merancang dua model evaluasi yaitu: (1) *Goal-Free Evaluation Approach*, dan (2) *Formative and Summative model*.

### **1) Goal-Free Evaluation Approach**

Scriven mengemukakan bahwa dalam melaksanakan evaluasi program, evaluator tidak perlu memperhatikan apa yang menjadi tujuan program. Yang perlu diperhatikan dalam program tersebut adalah bagaimana kerjanya (kinerja) suatu program, dengan jalan mengidentifikasi penampilan-penampilan yang terjadi (pengaruh) baik hal-hal yang positif (yaitu hal yang diharapkan) maupun hal-hal yang negatif (yang tidak diharapkan).

Evaluasi model *goal free evaluation*, fokus pada adanya perubahan perilaku yang terjadi sebagai dampak dari program yang diimplementasikan, melihat dampak sampingan baik yang diharapkan maupun yang tidak diharapkan, dan membandingkan dengan sebelum program dilakukan. Evaluasi juga membandingkan antara hasil yang dicapai dengan besarnya biaya yang dikeluarkan untuk program tersebut atau melakukan *cost benefit analysis*.

Model *goal free evaluation* merupakan titik evaluasi program, di mana objek yang dievaluasi tidak perlu terkait dengan tujuan dari objek atau subjek tersebut, tetapi langsung kepada implikasi keberadaan program apakah bermanfaat atau tidak objek tersebut atas dasar penilaian kebutuhan yang ada.

Tujuan program tidak perlu diperhatikan karena kemungkinan evaluator terlalu rinci mengamati tiap-tiap tujuan khusus. Jika masing-masing tujuan khusus tercapai, artinya terpenuhi dalam penampilan tetapi evaluator lupa memperhatikan sejauh mana masing-masing penampilan tersebut mendukung penampilan terakhir yang diharapkan oleh tujuan umum maka akibatnya jumlah penampilan khusus ini tidak banyak bermanfaat. Dapat disimpulkan bahwa, dalam model ini bukan berarti lepas dari tujuan tetapi hanya lepas dari tujuan khusus. Model ini hanya mempertimbangkan tujuan umum yang akan dicapai oleh program, bukan secara rinci perkomponen yang ada.

Scriven menekankan bahwa evaluasi itu adalah interpretasi *judgement* ataupun *explanation* dan evaluator yang merupakan pengambil keputusan dan sekaligus penyedia informasi. Ciri-ciri evaluasi bebas tujuan yaitu:



- a. Evaluatur sengaja menghindar untuk mengetahui tujuan program.
- b. Tujuan yang telah dirumuskan terlebih dahulu tidak dibenarkan menyempitkan fokus evaluasi.
- c. Evaluasi bebas tujuan berfokus pada hasil yang sebenarnya, bukan pada hasil yang direncanakan.
- d. Hubungan evaluatur dan manajer atau dengan karyawan proyek dibuat seminimal mungkin.
- e. Evaluasi menambah kemungkinan ditemukannya dampak yang tidak diramalkan (Tayibnapis, 2000:35).

Mungkin akan lebih baik apabila evaluasi yang berorientasi pada tujuan dan evaluasi bebas tujuan dipadukan, karena akan saling mengisi dan melengkapi. Evaluatur internal biasanya melakukan evaluasi yang berorientasi pada tujuan, karena ia sulit menghindar atau mau tidak mau ia akan mengetahui tujuan program, akan tidak pantas apabila ia tidak acuh. Manajer program jelas ingin mengetahui sampai seberapa jauh program telah dicapai, dan evaluatur internal akan dan harus menyediakan informasi untuk manajernya.

Di samping itu, perlu diketahui bagaimana orang luar menilai program bukan hanya untuk mengetahui apa yang seharusnya dilakukan, apa yang dilakukan di semua bagian, pada semua yang telah dihasilkan, secara sengaja atau tidak sengaja. Yang belakangan ini merupakan tugas evaluatur bebas tujuan yang tidak mengetahui tujuan program. Jadi, evaluasi yang berorientasi pada tujuan dan evaluasi bebas tujuan dapat bekerja sama dengan baik.

Scriven menjelaskan fokus pada program atau tujuan kegiatan ini dapat menjadi tempat awal yang penting untuk evaluatur bekerja dalam domain evaluasi. Tujuan program tertentu tidak harus diambil sebagai yang diberikan, tetapi diperiksa dan dievaluasi juga.

*Model Goal Free Evaluation* berfokus pada hasil yang sebenarnya dari suatu program atau kegiatan, bukan hanya tujuan-tujuan yang teridentifikasi. Jenis model ini memungkinkan evaluatur untuk mengidentifikasi dan mencatat hasil yang tidak mungkin telah diidentifikasi oleh perancang program. Melalui proses teknik baik terang-terangan dan terselubung, metode ini berusaha untuk mengumpulkan data dalam rangka untuk membentuk deskripsi program, mengidentifikasi proses akurat, dan menentukan pentingnya mereka ke program.

Model *Goal Free Evaluation* berfokus pada hasil tanpa goal (tujuan), sementara model lain berfokus pada proses pengambilan keputusan dan menyediakan administrator kunci dengan analisis mendalam untuk membuat keputusan yang adil dan tidak bias. Fungsi evaluasi bebas tujuan adalah untuk mengurangi bias dan menambah objektivitas. Dalam evaluasi yang berorientasi pada tujuan, seorang evaluator secara subjektif persepsinya akan membatasi sesuai dengan tujuan.

Padahal tujuan pada umumnya hanya formalitas dan jarang menunjukkan tujuan yang sebenarnya dari suatu proyek. Lagipula, banyak hasil program penting yang tidak sesuai dengan tujuan program. Evaluasi bebas tujuan berfokus pada hasil yang sebenarnya bukan pada hasil yang direncanakan. Dalam evaluasi bebas tujuan ini, memungkinkan evaluator untuk menambah temuan hasil atau dampak yang tidak direncanakan.

Model evaluasi *Goal Free Evaluation* ini mempunyai kekurangan dan kelebihan. Kelebihan dari model bebas tujuan di antaranya adalah:

- a. Evaluator tidak perlu memperhatikan secara rinci setiap komponen, tetapi hanya menekankan pada bagaimana mengurangi prasangka (bias).
- b. Model ini menganggap pengguna sebagai audiens utama. Melalui model ini, Scriven ingin evaluator mengukur kesan yang didapat dari sesuatu program dibandingkan dengan kebutuhan pengguna dan tidak membandingkannya dengan pihak penganjur.
- c. Pengaruh konsep pada masyarakat, bahwa tanpa mengetahui tujuan dari kegiatan yang telah dilakukan, seorang penilai bisa melakukan evaluasi.
- d. Kelebihan lain, dengan munculnya model bebas tujuan yang diajukan oleh Scriven adalah mendorong pertimbangan setiap kemungkinan pengaruh tidak saja yang direncanakan, tetapi juga dapat diperhatikan sampingan lain yang muncul dari produk.

Selanjutnya terkait dengan kelemahan model ini adalah:

- a. Model bebas tujuan ini pada umumnya bebas menjawab pertanyaan penting, seperti apa pengaruh yang telah diperhitungkan dalam suatu peristiwa dan bagaimana mengidentifikasi pengaruh tersebut.
- b. Walaupun ide Scriven bebas tujuan bagus untuk membantu kegiatan yang paralel dengan evaluasi atas dasar kejujuran, pada tingkatan

praktis model ini tidak terlalu berhasil dalam menggambarkan bagaimana evaluasi sebaiknya benar-benar dilaksanakan.

- c. Tidak merekomendasikan bagaimana menghasilkan penilaian kebutuhan walau pada akhirnya mengarah pada penilaian kebutuhan.
- d. Diperlukan evaluator yang benar-benar kompeten untuk dapat melaksanakan evaluasi model ini.
- e. Langkah-langkah sistematis yang harus dilakukan dalam evaluasi hanya menekankan pada objek sasaran saja.

## **2) *Formative-Summative Model***

Scriven adalah ahli yang pertama sekali membedakan evaluasi formatif dan evaluasi sumatif, kemudian sejak itu istilah ini menjadi populer dan dapat dikatakan diterima secara universal.

Evaluasi formatif didefinisikan sebagai proses menyediakan dan menggunakan informasi untuk dijadikan dasar pengambilan keputusan dalam meningkatkan kualitas produk atau program yang dirancang. Evaluasi formatif bertujuan untuk menentukan apa yang harus ditingkatkan atau direvisi agar produk atau program tersebut lebih sistematis, efektif dan efisien.

Evaluasi formatif dilaksanakan selama program berjalan untuk memberikan informasi yang berguna kepada pemimpin program untuk perbaikan program. Misalnya selama pengembangan program paket kurikulum, evaluasi formatif akan melibatkan pemeriksaan konten oleh ahli, melakukan pilot tes terhadap sejumlah siswa, tes lapangan terhadap siswa yang lebih banyak dan dengan guru di beberapa sekolah dan lain sebagainya.

Setiap langkah evaluasi pada tahapan evaluasi formatif akan menghasilkan umpan balik yang segera kepada perancang program yang kemudian menggunakan informasi tersebut untuk merevisi program apabila diperlukan. Kegagalan melakukan evaluasi formatif merupakan suatu kerugian karena data evaluasi formatif diperoleh lebih dulu, hal ini dapat menolong penyusunan jadwal kembali, pengaturan pembiayaan, dan sebagainya sehingga dapat diarahkan ke arah yang lebih produktif. Evaluasi yang dilaksanakan pada saat perancangan program akan berakhir akan memungkinkan terlambat dan tidak dapat menolong.

Langkah-langkah pada evaluasi formatif adalah sebagai berikut:

a. *One-to-one evaluation.*

*One-to-one evaluation* yang dilakukan *expert*. Prosedur yang ditempuh adalah: (1) mendatangi *expert* (ahli materi, ahli desain, ahli bahasa dan ahli lainnya), (2) pengembang menjelaskan proses yang dilaksanakan, dan (3) meminta *judgement* dari *expert*. Instrumen yang digunakan adalah wawancara dan lembar penilaian dalam bentuk kuesioner.

Berdasarkan catatan masukan dan saran yang disampaikan oleh *expert* melalui *one-to-one evaluation* tersebut maka dilakukan perbaikan terhadap produk/program yang dikembangkan dan kemudian hasil revisi tersebut dikonfirmasi ulang kepada *expert*.

*One-to-one evaluation* yang dilakukan dengan tiga subjek pengguna produk atau program. Prosedur yang ditempuh adalah:

1. Menjelaskan maksud evaluasi yaitu mendapatkan catatan masukan terhadap produk atau program yang dirancang.
2. Memotivasi subjek pengguna mengikuti kegiatan evaluasi dengan sebaik-baiknya dalam waktu yang telah ditentukan yaitu dengan mengikuti secara seksama.
3. Memberikan tes yang bertujuan untuk melihat bagian-bagian yang digunakan masih perlu diperbaiki atau tidak.
4. Memotivasi subjek pengguna untuk memberikan komentar dengan leluasa menyimpulkan implikasinya terhadap perbaikan secara komprehensif.

Berdasarkan catatan masukan dan saran yang disampaikan tiga subjek pengguna melalui *one-to-one evaluation* maka dilakukan perbaikan, selanjutnya hasil revisi dikonfirmasi ulang kepada subjek pengguna.

b. *Small group evaluation.*

*Small group evaluation* yang dilakukan kepada sepuluh subjek pengguna. Prosedur yang ditempuh adalah:

1. Menyampaikan tujuan diadakannya evaluasi yaitu mendapatkan umpan balik dalam rangka merevisi produk atau program.
2. Menjelaskan kegiatan yang akan dilakukan dan mendorong subjek pengguna untuk memberikan catatan masukan dan saran.

3. Memberikan produk atau program yang akan dievaluasi kepada subjek pengguna.
4. Mencatat seluruh masukan dan saran.
5. Melakukan wawancara untuk mendapatkan informasi yang lebih mendalam tentang:
  - a. Kepraktisan memahami materi yang terdapat dalam produk/program.
  - b. Kegiatan yang dikembangkan tersebut menarik dan sistematis.
  - c. Bagian-bagian yang sulit dipahami dan penyebabnya.
  - d. Kerelevanan butir tes dengan materi yang disajikan.

Berdasarkan catatan masukan dan saran yang disampaikan melalui *small group evaluation* dilakukan perbaikan, selanjutnya hasil revisi dikonformasi ulang kepada subjek pengguna.

c. *Field trial evaluation*.

*Field trial evaluation* atau ujicoba lapangan dilakukan dengan prosedur sebagai berikut:

1. Menentukan subjek pengguna yang menjadi sasaran ujicoba lapangan.'
2. Menyiapkan fasilitas, alat-alat dan lingkungan sesuai dengan strategi dan bentuk kegiatan yang telah ditentukan.
3. Melaksanakan kegiatan sesuai dengan materi/bahan dan bentuk kegiatan.
4. Mencatat informasi tentang kualitas proses dan bahan/materi yang dilakukan dengan memberikan kuesioner dan pedoman wawancara. Di samping itu juga dilakukan observasi untuk kegiatan dan keadaan lingkungan kegiatan.
5. Melakukan tes awal dan tes akhir untuk mengetahui efektivitas kegiatan.

Berdasarkan hasil data kuesioner, wawancara, dan observasi, serta tes yang diperoleh dari aktivitas ujicoba lapangan dilakukan revisi akhir dari produk atau program yang didesain.

Scriven sebagaimana dikutip Stufflebeam dan Shinkfield (2007) menjelaskan evaluasi sumatif adalah proses menilai suatu objek, dalam ini apabila ternyata produk atau program yang dirancang ternyata sama

efektifnya dengan yang lama, maka produk atau program yang dirancang dapat digunakan atau dilanjutkan.

Evaluasi sumatif dilakukan pada akhir program untuk memberi informasi kepada pengguna/konsumen yang potensial tentang manfaat atau kegunaan program. Misalnya, sesudah paket kurikulum dikembangkan, evaluasi sumatif mungkin dilaksanakan untuk menentukan efektifitas paket tersebut pada tingkat nasional atas sampel sekolah khusus, guru dan siswa pada tingkat perkembangan tertentu. Penemuan hasil pada evaluasi sumatif ini akan diberikan kepada konsumen/ pengguna.

Objek atau subjek dan pemakaian evaluasi antara evaluasi formatif dan evaluasi sumatif berbeda. Pada evaluasi formatif, audiensinya adalah personalia program, dalam contoh di atas, adalah mereka yang bertanggung jawab atas pengembangan kurikulum. Pada evaluasi sumatif, audiensinya termasuk konsumen yang potensial seperti siswa, guru, dan lain-lain yang terlibat dalam program. Evaluasi formatif harus mengarah kepada keputusan tentang perkembangan program tersebut termasuk perbaikan atau revisi. Sedangkan evaluasi sumatif mengarah ke arah keputusan tentang kelanjutan program, berhenti atau program diteruskan, pengadopsian, dan sebagainya.

Dengan demikian jelaslah bahwa evaluasi formatif dan evaluasi sumatif sangatlah penting karena keputusan diperlukan selama proses, tingkat pengembangan proyek, untuk memperbaiki, dan memperkuat lagi sesudah stabil, untuk menilai manfaat atau menentukan masa depan program.

#### **4. *Contentance Evaluation Model (Stake Model)***

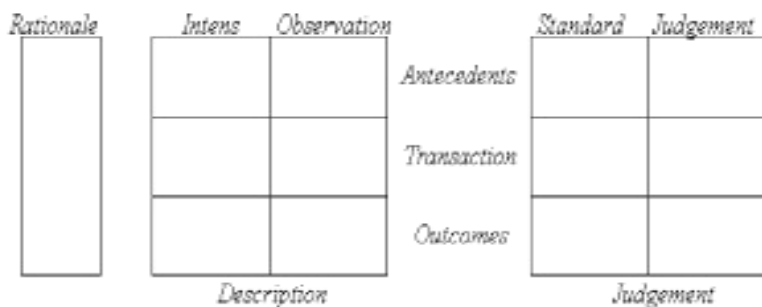
Model evaluasi program yang diperkenalkan oleh Stake dikenal dengan model *Contentance* (keseluruhan). Model ini juga disebut model evaluasi pertimbangan. Maksudnya evaluator mempertimbangkan program dengan membandingkan kondisi hasil evaluasi program dengan yang terjadi di program lain, dengan objek sasaran yang sama dan membandingkan kondisi hasil pelaksanaan program dengan standar yang ditentukan oleh program tersebut.

Tujuan dari model *Contentance Stake* adalah melengkapi kerangka untuk pengembangan suatu rencana penilaian kurikulum. Perhatian utama Stake adalah hubungan antara tujuan penilaian dengan keputusan berikutnya berdasarkan sifat data yang dikumpulkan. Hal tersebut, karena Stake melihat adanya ketidaksesuaian antara harapan penilai dan guru.

Dalam model ini Stake menekankan peran evaluator dalam mengembangkan tujuan kurikulum menjadi tujuan khusus yang terukur. Model *countenance* terdiri atas dua matriks yaitu *description* (gambaran) dan *judgement* (pertimbangan). Matriks pertimbangan baru dapat dikerjakan oleh evaluator setelah matriks deskripsi diselesaikan. Matriks Deskripsi terdiri atas kategori rencana (*intent*) dan observasi. Matriks Pertimbangan terdiri atas kategori standar dan pertimbangan. Pada setiap kategori terdapat tiga fokus yaitu:

- Antecedents* yaitu sebuah kondisi yang ada sebelum instruksi yang mungkin berhubungan dengan hasil, contohnya: latar belakang guru, kurikulum yang sesuai, ketersediaan sumber daya.
- Transaction* yaitu pertemuan dinamis yang merupakan proses instruksi (kegiatan, proses, dan lainnya), contohnya: interaksi guru dan siswa.
- Outcomes* yaitu efek dari pengalaman pembelajaran (pengamatan dan hasil tenaga kerja), contohnya *performance* guru, peningkatan kinerja.

Model evaluasi Stake dapat membawa dampak yang cukup besar dalam penilaian, dan merupakan konsep yang cukup kuat untuk perkembangan yang lebih jauh dalam bidang evaluasi. Dalam model ini, evaluasi dilakukan dengan membandingkan antara satu program dengan program lain yang dianggap standar. Stake berpendapat menilai suatu program pendidikan harus melakukan perbandingan yang relatif antara program satu dan program yang lain, atau perbandingan yang absolut yaitu membandingkan suatu program dengan standar tertentu (Tayibnapis, 2000:19). Model evaluasi Stake digambarkan sebagai berikut:



Gambar VI. 1 Model Evaluasi Stake

Dari gambar di atas terlihat bahwa terdapat dua matriks yaitu: (1) matriks deskripsi, dan (2) matriks pertimbangan. Matriks deskripsi terdiri atas dua kategori:

- a. Kategori pertama adalah sesuatu yang direncanakan pengembang program. Ilustrasi yang diberikan adalah konteks program penyusunan kurikulum 2013 yang dikembangkan oleh instansi terkait dalam hal ini Kementerian Pendidikan. Sedangkan Rencana Program Pengajaran (RPP) yang dikembangkan guru. Guru sebagai pengembang program merencanakan keadaan/ persyaratan yang diinginkannya untuk suatu kegiatan kelas tertentu. Misalnya yang berhubungan dengan minat, kemampuan, pengalaman, dan lain sebagainya dari peserta didik.
- b. Kategori kedua dinamakan observasi, berhubungan dengan apa yang sesungguhnya sebagai implementasi yang diinginkan pada kategori yang pertama. Kategori ini juga sebagaimana yang pertama terdiri atas *antecedents*, transaksi, dan hasil. Evaluator harus melakukan observasi (pengumpulan data) mengenai *antecedents*, transaksi, dan hasil yang ada di suatu satuan pendidikan.

Selanjutnya adalah matriks pertimbangan terdiri atas kategori standar dan pertimbangan, dan fokus *antecedents*, transaksi, dan *outcomes* (hasil yang diperoleh). Standar adalah kriteria yang harus dipenuhi oleh suatu kurikulum atau program yang dijadikan objek evaluasi. Standar dapat dikembangkan dari karakteristik yang dimiliki kurikulum, tetapi dapat juga dari yang lain (*pre-ordinate*, *mutually adaptive*, proses). Kategori kedua adalah kategori pertimbangan. Kategori ini menghendaki evaluator melakukan pertimbangan dari apa yang telah dilakukan dari kategori yang pertama dan kedua matriks Deskripsi sampai kategori pertama matriks pertimbangan. Suatu evaluasi harus sampai kepada pemberian pertimbangan. Keseluruhan matriks yang mendukung model Stake ini terdiri dari 12 kotak.

Untuk melakukan evaluasi menggunakan model evaluasi Stake ini dilakukan melalui langkah-langkah berikut:

- a. Pengumpulan data.  
Evaluator mengumpulkan data mengenai apa yang diinginkan pengembang program baik yang berhubungan dengan kondisi awal, transaksi, dan hasil. Data dapat dikumpulkan melalui studi dokumen dapat pula melalui wawancara. Sebelum melakukan pengumpulan data, maka para evaluator harus bertemu terlebih dahulu untuk membuat kerangka



acuan yang berhubungan dengan *antecedents*, transaksi dan hasil. Hal tersebut dilakukan tidak hanya untuk memperjelas tujuan evaluasi tetapi juga untuk melihat apakah konsisten terhadap *transactions* yang dimaksud dengan *antecedent* dan *outcome*.

b. Analisis Data.

Analisis data yang dilakukan meliputi analisis logis dan empirik. Analisis logis diperlukan dalam memberikan pertimbangan mengenai keterkaitan antara prasyarat awal, transaksi, dan hasil dari kotak-kotak tujuan. Evaluator harus dapat menentukan apakah prasyarat awal yang telah dikemukakan pengembang program akan tercapai dengan rencana transaksi yang dikemukakan. Atau sebetulnya ada model transaksi lain yang lebih efektif. Demikian pula mengenai hubungan antara transaksi dengan hasil yang diharapkan. Selanjutnya analisis empirik, pada dasar bekerjanya sama dengan analisis logis tapi data yang digunakan adalah data empirik.

c. Analisis *congruence* (kesesuaian).

Analisis *congruence* (kesesuaian) merupakan analisis, di mana evaluator membandingkan antara apa yang dikemukakan dalam tujuan (inten) dengan apa yang terjadi dalam kegiatan (observasi). Dalam hal ini evaluator menganalisis apakah yang telah direncanakan dalam tujuan telah sesuai dengan pelaksanaannya di lapangan atau terjadi penyimpangan. Apabila analisis *congruence* telah selesai, maka evaluator menyerahkannya kepada tim yang terdiri dari para ahli dan orang yang terlibat dalam program. Tim ini yang akan meneliti kesahihan hasil analisis evaluator dan memberikan persepsinya mengenai faktor penting *congruence*.

d. Pertimbangan hasil. Tugas evaluator berikutnya adalah memberikan pertimbangan mengenai program yang sedang dikaji. Untuk itu, evaluator memerlukan standar.

Penekanan yang umum atau hal yang penting dalam model Stake ini adalah evaluator yang membuat penilaian tentang program yang di evaluasi. Stake mengatakan bahwa *description* di satu pihak berbeda dengan *judgement* atau menilai. Dalam model ini, *antecedents* (masukan), *transaction* (proses) dan *outcomes* (hasil) data dibandingkan tidak hanya menentukan apakah ada perbedaan tujuan dengan keadaan yang sebenarnya, tetapi juga dibandingkan dengan standar yang absolut untuk menilai manfaat program. Stake menyatakan

bahwa tak ada penelitian dapat diandalkan apabila tidak dinilai (Tayibnapis, 2000:22).

Kelebihan evaluasi model Stake adalah:

- a. Dalam evaluasi memasukkan data tentang latar belakang program, proses dan hasil yang merupakan perluasan ruang lingkup evaluasi.
- b. Evaluator memegang kendali dalam evaluasi dan juga memutuskan cara yang paling tepat untuk hadir dan menggambarkan hasil.
- c. Memiliki potensi besar untuk memperoleh wawasan baru dan teori-teori tentang lapangan dan program yang akan di evaluasi.

Selanjutnya terkait dengan kelemahan evaluasi model Stake ini adalah:

- a. Pendekatan yang dilakukan terlalu subjektif.
- b. Terjadinya kemungkinan dalam meminimalkan pentingnya instrumen pengumpulan data dan evaluasi kuantitatif.
- c. Kemungkinan biaya yang terlalu besar.

## 5. Model Alkin

Dalam merumuskan model evaluasi program yang disusunnya, Alkin membuat batasan konstruk evaluasi sebagai suatu proses penentuan area yang akan di evaluasi, pemilihan informasi yang cocok untuk dievaluasi, pengumpulan dan analisis informasi serta penyusunan laporan atau ringkasan data yang berguna bagi pengambil keputusan dalam memilih alternatif yang berguna yang tepat dari berbagai alternatif yang ada.

Model Alkin dikembangkan berdasarkan 4 (empat) asumsi. Apabila keempat asumsi ini sudah dipenuhi maka model ini dapat digunakan. Adapun keempat asumsi itu yaitu:

- a. Variabel perantara adalah satu-satunya variabel yang dimanipulasi.
- b. Sistem eksternal tidak langsung dipengaruhi oleh keluaran sistem (persekolahan).
- c. Para pengambil keputusan sekolah tidak memiliki kontrol mengenai pengaruh yang diberikan sistem eksternal terhadap sekolah.
- d. Faktor masukan mempengaruhi aktivitas faktor perantara dan pada gilirannya faktor perantara berpengaruh terhadap hasil.

Deskripsi dari model evaluasi dari Alkin terdiri dari 5 (lima) macam evaluasi yakni:

- a. Sistem *assessment*, yaitu memberikan informasi tentang keadaan atau posisi sistem. Contohnya dalam hal penerapan metode pembelajaran. Hasil evaluasi dengan menggunakan model ini antara lain dapat menunjukkan peningkatan hasil belajar siswa.
- b. Program *planning*, membantu pemilihan program tertentu yang mungkin akan berhasil memenuhi kebutuhan program. Dalam program *planning* dapat dilakukan melalui evaluasi internal dan evaluasi eksternal. Evaluasi internal dilakukan dengan cara menilai ketepatan, kesesuaian dan kebermaknaan sub-sub program yang dirumuskan dalam kaitannya dengan tujuan program yang dinilai, baik dari segi konstruksi, kepraktisan dan biaya. Sedangkan evaluasi eksternal adalah evaluasi yang dilakukan sesudah suatu program diimplementasikan.

Salah satu cara yang dapat digunakan adalah *Delphi Techniques* atau teknik lain yang menggunakan pendekatan sistem analisis. Untuk contoh penerapan metode pembelajaran, metode pembelajaran disesuaikan dengan karakteristik materi dan karakteristik siswa. Setelah terpilih, metode pembelajaran tersebut direalisasikan dalam proses pembelajaran.

- c. Program *implementation* yaitu menyiapkan informasi apakah program sudah diperkenalkan kepada kelompok tertentu yang tepat seperti direncanakan. Dalam contoh penerapan metode pembelajaran, model ini dimaksudkan untuk mengevaluasi, misalnya apakah metode yang digunakan telah sesuai dengan karakteristik materi dan karakteristik siswa.
- d. Program *improvement*, yaitu memberikan informasi tentang bagaimana program berfungsi, bagaimana program bekerja, atau berjalan? Apakah dalam menuju pencapaian tujuan ada hal-hal atau masalah-masalah baru yang muncul tak terduga? Dengan kata lain, evaluator mengidentifikasi masalah-masalah yang muncul, mengumpulkan dan menganalisis data serta menyerahkan pada pengambil keputusan untuk melakukan perbaikan pelaksanaan program dengan segera.

Dalam contoh penerapan metode pembelajaran, model ini dimaksudkan untuk menilai proses pembelajaran, apakah berjalan dengan baik dan sesuai dengan rencana, bagaimana penanggulangan masalah jika terjadi kendala selama terjadi proses pembelajaran.

- e. Program *certification*, yaitu memberikan informasi tentang nilai atau guna program. Dalam contoh penerapan metode pembelajaran, model ini dimaksudkan untuk mengevaluasi apakah metode yang diterapkan memberikan dampak positif pada siswa, yakni siswa semakin termotivasi untuk belajar sehingga dapat meningkatkan hasil belajar siswa.

Model Alkin ini sedikit unik karena selalu memasukkan unsur pendekatan ekonomi mikro dalam pekerjaan evaluasi. Adapun pendekatan yang digunakan disebut Alkin dengan pendekatan sistem. Disebut pendekatan sistem karena model ini mengutamakan sistem yang berjalan seperti halnya pendidikan yang diartikan sebagai sebuah sistem.

Dua hal yang harus diperhatikan oleh evaluator dalam model ini adalah pengukuran dan variabel kontrol. Alkin membagi model ini dibagi atas tiga komponen yaitu masukan, proses yang dinamakan dengan istilah perantara, dan keluaran (hasil). Alkin juga mengenal sistem internal yang merupakan interaksi antar komponen yang langsung berhubungan dengan pendidikan dan sistem eksternal yang mempunyai pengaruh dan dipengaruhi oleh pendidikan.

Kelebihan model ini adalah keterikatannya dengan sistem. Dengan model ini, kegiatan sekolah dapat diikuti dengan seksama mulai dari variabel-variabel yang ada dalam komponen masukan, proses, dan keluaran. Komponen masukan yang dimaksud adalah semua informasi yang berhubungan dengan karakteristik siswa, kemampuan intelektual, hasil belajar sebelumnya, kepribadian, kebiasaan, latar belakang keluarga, latar belakang lingkungan dan sebagainya.

Adapun yang dimaksud dengan proses di sini meliputi faktor perantara yang merupakan kelompok variabel yang secara langsung mempengaruhi keluaran. Adapun yang masuk dalam variabel perantara ini diantaranya adalah ratio jumlah guru dengan siswa, jumlah siswa dalam kelas, pengaturan administrasi, penyediaan buku bacaan, prosedur pengajaran dan sebagainya.

Adapun keluaran siswa adalah setiap perubahan yang terjadi pada diri siswa sebagai akibat dari pengalaman belajar yang diperolehnya. Perubahan ini harus diikuti sejak siswa masuk sistem hingga keluar sistem. Perubahan harus diukur meliputi setiap aspek perubahan yang mungkin terjadi termasuk di dalamnya kemampuan siswa dalam melanjutkan pelajaran di tingkat pendidikan yang lebih tinggi, pada waktu memasuki lapangan kerja, dalam melakukan pekerjaan bahkan termasuk aktivitas dalam kehidupan di masyarakat.

Kelemahan dari model Alkin adalah keterbatasannya dalam fokus kajian yaitu yang hanya fokus pada kegiatan persekolahan. Sehingga model ini hanya dapat digunakan untuk mengevaluasi kurikulum yang sudah siap dilaksanakan di sekolah.

## 6. CSE-UCLA Evaluation Model

CSE-UCLA adalah akronim dari *Center for the Study of Evaluation-University of California in Los Angeles*. Pada awalnya, karakteristik dari model CSE-UCLA adalah adanya 5 (lima) tahap yang dilakukan dalam evaluasi yaitu: perencanaan, pengembangan, implementasi, hasil dan dampak. Seiring dengan perkembangannya, Fernandes sebagaimana dikutip Arikunto dan Jabar (2009:44) memaparkan bahwa langkah-langkah dari model CSE-UCLA menjadi empat tahap yaitu:

a. *Need assessment.*

Pada tahap pertama ini yaitu analisis kebutuhan, evaluator memusatkan perhatian pada penentuan masalah pertanyaan yang dapat diajukan yaitu:

- 1) Hal-hal apakah yang perlu dipertimbangkan sehubungan dengan keberadaan program?
- 2) Kebutuhan apakah yang terpenuhi sehubungan dengan adanya pelaksanaan program ini?
- 3) Tujuan jangka panjang apakah yang dapat dicapai melalui program ini?

b. *Program planning.*

Pada tahap kedua ini yaitu perencanaan program, evaluator mengumpulkan data yang terkait langsung dengan program dan mengarahkan pada pemenuhan kebutuhan yang telah diidentifikasi pada tahap pertama. Dalam tahap perencanaan ini program yang di evaluasi dengan cermat untuk mengetahui apakah rencana program yang telah disusun berdasarkan analisis kebutuhan. Evaluasi tahap ini tidak lepas dari tujuan yang telah dirumuskan sebelumnya.

c. *Formative evaluation.*

Dalam tahap ketiga ini yaitu evaluasi formatif, evaluator memusatkan perhatian pada keterlaksanaan program. Dengan demikian, evaluator

diharapkan betul-betul terlibat dalam program karena harus mengumpulkan data dan berbagai informasi dari pengembang program.

d. *Summative evaluation*.

Dalam tahap keempat yaitu evaluasi sumatif, evaluator diharapkan dapat mengumpulkan semua data tentang hasil dan dampak dari program. Melalui evaluasi sumatif ini diharapkan dapat diketahui apakah tujuan yang dirumuskan untuk program sudah tercapai dan jika belum dicari bagian mana yang dan apa faktor-faktor penyebabnya.

## 7. *Discrepancy Model*

*Discrepancy model* atau model kesenjangan digagas oleh Malcolm Probus yaitu evaluasi yang dilakukan dengan maksud untuk mengetahui tingkat kesesuaian antara standar yang telah ditentukan dalam program dengan penampilan aktual dari program tersebut. Standar adalah kriteria yang telah dikembangkan dan ditetapkan dengan hasil yang efektif, sedangkan penampilan adalah sumber, prosedur, manajemen, dan hasil nyata yang tampak ketika program dilaksanakan.

Langkah-langkah dalam model kesenjangan ini adalah:

a. Penyusunan desain.

Dalam tahap ini dilakukan kegiatan sebagai berikut:

- 1) Merumuskan tujuan program.
- 2) Menyiapkan klien, staf dan kelengkapan lain.
- 3) Merumuskan standar dalam bentuk rumusan yang menunjuk pada sesuatu yang dapat diukur. Biasanya di dalam langkah ini evaluator berkonsultasi dengan pengembangan program.

b. Pemasangan instalasi (*installation*).

Tahap ini melihat apakah kelengkapan yang tersedia sudah sesuai dengan yang diperlukan atau belum. Dalam tahap ini dilakukan kegiatan sebagai berikut:

- 1) Meninjau kembali penetapan standar.
- 2) Meninjau program yang sedang berjalan.
- 3) Meneliti kesenjangan antara yang direncanakan dengan yang sudah dicapai.

c. Proses (*process*).

Dalam tahap ini adalah mengadakan penilaian tujuan-tujuan manakah yang sudah dicapai. Dalam hal ini adalah mengumpulkan data dari pelaksanaan program.

d. Pengukuran tujuan (*product*)

Tahap ini adalah mengadakan analisis data dan menetapkan tingkat *output* yang diperoleh. Pertanyaan yang diajukan dalam tahap ini adalah: apakah program sudah mencapai tujuan terminalnya?

e. Pembandingan (*programme comparison*).

Tahap ini adalah membandingkan hasil yang telah dicapai dengan tujuan yang telah ditetapkan. Dalam tahap ini evaluator menuliskan semua penemuan tentang kesenjangan untuk disajikan kepada pengambil keputusan, agar dapat memutuskan kelanjutan dari program tersebut.

## 8. Model Brinkerhoff

Desain evaluasi program pada umumnya terdiri atas elemen-elemen yang sama, ada banyak cara untuk menggabungkan elemen tersebut, masing-masing ahli atau evaluator mempunyai konsep yang berbeda dalam hal ini Brinkerhoff et-al mengemukakan tiga golongan evaluasi yang disusun berdasarkan penggabungan elemen-elemen yang sama seperti model lainnya, namun dalam komposisi dan versi sendiri.

Model Brinkerhoff et-al terdiri dari: (1) *fixed vs emergent evaluation design*, (2) *formative vs summative evaluation*, dan (3) *experimental and quasi experimental design vs natural/unobtrusive inquiry* (Tayibnapis, 2000:16).

### a. *Fixed vs emergent evaluation design.*

Desain evaluasi yang tetap (*fixed*) ditentukan dan direncanakan secara sistematis sebelum implementasi dikerjakan. Desain dikembangkan berdasarkan tujuan program disertai seperangkat pertanyaan yang akan dijawab oleh informasi yang akan diperoleh dari sumber-sumber tertentu.

Rencana analisis dibuat sebelumnya di mana si pemakai akan menerima informasi seperti yang telah ditentukan dalam tujuan. Desain biasanya dibicarakan dan dirundingkan dengan pemakai utama atau pemesan, apabila ada perubahan biasanya hanya lebih memperlancar pencapaian tujuan dan rencana utama.

Walaupun desain *fixed* ini lebih terstruktur daripada desain *emergent*, desain *fixed* juga dapat disesuaikan dengan kebutuhan yang mungkin berubah. Kebanyakan evaluasi formal yang dibuat secara individu berdasarkan desain *fixed*, karena tujuan program telah ditentukan dengan jelas sebelumnya, dibiayai dan melalui usulan atau proposal evaluasi.

Desain *fixed* relatif memakai biaya yang cukup besar, kegiatan-kegiatan berkisar membuat pertanyaan-pertanyaan menyiapkan dan membuat instrumen, menganalisis hasil evaluasi dan melaporkan secara formal hasil evaluasi kepada pemakai. Komunikasi antara evaluator dan klien dilakukan secara teratur, formal atau tertulis.

Evaluator berpedoman pada tujuan program untuk merumuskan masalah atau pertanyaan-pertanyaan desain dan menstimulasi klien yang relevan untuk mengembangkan dan membetulkan pertanyaan tersebut. Strategi pengumpulan informasi khususnya menggunakan cara-cara formal juga metode penelitian, kriteria penelitian dan pengumpulan data biasanya dilakukan secara kuantitatif.

Selanjutnya desain evaluasi *emergent*, dibuat untuk beradaptasi dengan pengaruh dan situasi yang sedang berlangsung dan berkembang seperti menampung pendapat-pendapat klien, masalah-masalah, kegiatan program. Evaluasi ini menghabiskan banyak waktu dari permulaan sampai akhir mencari tujuan dan isu, karena semuanya pada dasarnya tidak dikhususkan dan ditentukan sebelumnya.

Biaya relatif lebih besar, sumber-sumber dalam desain ditentukan untuk mengamati program dan memfokuskan pertanyaan lebih lanjut. Evaluator tidak mendorong klien memikirkan tentang program atau isu evaluasi. Klien menentukan isu-isu penting dan informasi yang diperlukan desain. Komunikasi antar evaluator dengan klien terus berkesinambungan selama proses evaluasi.

Metode yang digunakan dalam desain ini adalah observasi, studi kasus dan laporan tim penyokong. Pengukuran yang tidak selalu berpedoman pada tujuan biasanya dilakukan dan evaluator sering kali mengorbankan ketepatan pengukuran untuk lebih berguna. Informasi yang bersifat kualitatif biasanya dikumpulkan. Desain terus berkembang, berubah dan bereaksi sesuai dengan situasi dan kondisi yang dapat dikatakan tak pernah berhenti.



**b. *Formative vs summative evaluation.***

Evaluasi formatif digunakan untuk memperoleh informasi yang dapat membantu memperbaiki program. Dibuat untuk karyawan, dan juga mengorbankan kepentingan orang luar untuk lebih bermanfaat bagi program. Ada yang mengatakan bahwa evaluasi yang paling melindungi yaitu evaluasi formatif.

Fokus evaluasi formatif berkisar paada kebutuhan yang dirumuskan oleh karyawan atau orang-orang yang terlibat dalam program. Evaluator sering merupakan bagian daripada program dan bekerja sama dengan orang-orang program. Strategi pengumpulan informasi mungkin juga dipakai, tetapi penekanan pada usaha memberikan informasi yang berguna secepatnya bagi perbaikan program. Desain evaluasi dibuat bersama orang-orang program dan direvisi untuk mencapai kebutuhan mereka.

Selanjutnya evaluasi sumatif dibuat untuk menilai kegunaan suatu objek yang sering diminati atau dibiayai oleh pemakai/pemesan atau oleh sponsor. Evaluasi sumatif digunakan untuk menilai apakah suatu program akan diteruskan atau dihentikan saja. Evaluator harus dapat dipercaya oleh sejumlah klien yang akan dipengaruhi oleh keputusan tersebut. Untuk usaha-usaha yang dibiayai perorangan, evaluasi sumatif ini lebih populer daripada evaluasi formatif.

Pada evaluasi sumatif, evaluasi berfokus pada variabel-variabel yang dianggap penting oleh sponsor atau pembuat keputusan (pengambil kebijakan). Evaluator luar atau tim *review* sering dipakai, karena evaluator internal dapat mempunyai minat yang berbeda. Strategi pengumpulan informasi akan memaksimalkan validitas eksternal dan internal yang mungkin dikumpulkan dalam waktu yang cukup lama.

**c. *Experimental and quasi experimental design vs natural/unobtrusive inquiry.***

Desain eksperimen dan quasi eksperimen menggunakan metodologi penelitian yang subjek penelitiannya di *random* (acak), perlakuan diberikan dan pengukuran dampak dilakukan. Tujuan dari penelitian yaitu untuk menilai manfaat suatu objek, suatu program atau strategi yang dicobakan. Apabila subjek atau program dipilih secara acak, maka generalisasi dapat populasi.

Dalam beberapa hal, intervensi tidak mungkin dilakukan atau tidak dikehendaki. Apabila proses sudah terjadi, evaluator harus melihat dokumen-

dokumen sejarah, mempelajari nilai tes, atau menganalisis penelitian yang dilakukan. Bila dianggap penting untuk mengevaluasi suatu lingkungan atau program agar dapat diperbaiki, evaluator mungkin memilih mengamatinya, bicara dengan orang-orang yang terlibat, dan selalu merendah (*low profile*) sehingga program yang dievaluasi tidak terancam dan mengubah diri karena kehadiran evaluator. Banyak metodologi termasuk observasi, survey, analisis meta evaluasi, studi kasus dan wawancara dapat dilakukan seperti itu untuk mengurangi dampak evaluasi pada orang dan proyek dan memaksimalkan laporan yang sebenarnya.

Desain penelitian ini menggunakan waktu dan biaya yang relatif banyak yang digunakan untuk mempersiapkan instrumen untuk menilai perlakuan, data kuantitatif biasanya dikumpulkan, dan kriteria statistik juga digunakan. Kriteria statistik berfokus pada hasil program dan dibuat sebelumnya. Interaksi dengan klien dilakukan untuk membuat rencana, mengumpulkan informasi dan melaporkan kembali.

Strategi pengumpulan data terutama menggunakan instrumen formal seperti tes, survey, kuesioner dan *skala rating* serta memakai metode penelitian yang standar. Kriteria penelitian seperti *internal* dan *external validity* dianggap penting. Data yang dikumpulkan kebanyakan kuantitatif. Desain penelitian biasanya dibuat bersama pemesan atau pemakai. Bila ada perubahan hanya untuk memperlancar pencapaian tujuan sesuai rencana.

Selanjutnya penelitian *natural inquiry*, evaluator menghabiskan banyak waktu untuk mengamati dan berbicara dengan klien yang relevan. Strategi yang *multiple* dan sumber-sumber dipakai untuk mempertinggi reliabilitas pengumpulan data. Evaluator merundingkan isu dengan klien, hal ini dilakukan sesuai dengan cara evaluator. Interaksi dengan klien berkesinambungan dan informal. Observasi, studi kasus, laporan tim penyokong, merupakan ciri-ciri desain *natural inquiry*.

## BAB IV

# PERENCANAAN EVALUASI PROGRAM

### A. Analisis Kebutuhan

Aktivitas ilmiah apakah itu melakukan penelitian (kuantitatif, kualitatif, R&D dan lainnya), mendesain program, mendesain pembelajaran dan sebagainya tak terkecuali melakukan evaluasi terhadap suatu program maka tahapan awal yang menjadi dasar atau argumentasi rasional untuk melakukan aktivitas ilmiah tersebut adalah dilakukannya kajian analisis kebutuhan (*need assessment*). Dengan kata lain analisis kebutuhan dilakukan untuk menjawab urgensinya suatu aktivitas ilmiah itu perlu dan penting untuk dilakukan.

Oleh karena itu, jika ada pihak-pihak yang mempertanyakan suatu tindakan yang dilakukan, maka pihak yang menjadi pelaku tindakan tersebut dapat menunjukkan fakta-fakta merupakan data yang diperoleh untuk melakukan tindakan tersebut melalui aktivitas analisis kebutuhan. Jadi tidak karena selera ataupun karena kepentingan yang tidak ilmiah maka melakukan suatu aktivitas.

#### 1. Pengertian.

Anderson sebagaimana dikutip Arikunto dan Jabar (2009:71) memaparkan bahwa analisis kebutuhan diartikan sebagai suatu proses yang dilakukan oleh seseorang untuk mengidentifikasi kebutuhan sekaligus menentukan prioritas diantaranya. Dalam konteks pendidikan dan pembelajaran kebutuhan diartikan sebagai suatu kondisi yang memperlihatkan adanya kesenjangan antara keadaan nyata (yang ada) dengan kondisi yang diharapkan. Kebutuhan tersebut dapat terjadi pada diri individu, kelompok maupun lembaga.

Kaufman dan English (1979:8) memaknai analisis kebutuhan sebagai suatu proses formal untuk menentukan jarak atau kesenjangan antara keluaran dan dampak yang nyata dengan keluaran dan dampak yang diinginkan, kemudian menempatkan deretan kesenjangan tersebut dalam skala prioritas lalu memilih hal yang paling penting untuk diselesaikan masalahnya. Dalam hal ini kebutuhan diartikan sebagai jarak antara keluaran nyata dengan keluaran yang diinginkan untuk memperoleh keluaran dan dampak yang ditentukan.

Hardjodipuro (1982:22) menyatakan analisis kebutuhan adalah proses mengidentifikasi perbedaan-perbedaan (*gaps, discrepancies*) untuk dipecahkan yang dinyatakan berdasarkan prioritas dan menunjukkan persyaratan keseluruhannya bagi suatu sistem pendidikan. Selanjutnya Witkin sebagaimana dikutip Sagala (2010:114) menjelaskan analisis kebutuhan adalah mengenali gap dalam hasil-hasil yang menunjukkan apakah sesuatu terlalu banyak atau sedikit.

Menurut Briggs (1977:xxiv) analisis kebutuhan adalah proses sistematis untuk menentukan tujuan, mengidentifikasi kesenjangan antara tujuan dan keadaan sekarang dan menetapkan prioritas yang digunakan untuk mengambil tindakan.

Berdasarkan pemaparan di atas maka dapatlah dimaknai bahwa analisis kebutuhan sesungguhnya adalah proses menemukan, memilah dan menyisihkan kemudian mengambil skala prioritas terhadap kebutuhan yang menjadi kesenjangan antara kenyataan yang ada dengan kenyataan yang diharapkan. Proses menemukan, memilah, menyisihkan kemudian mengambil skala prioritas tersebut tentunya dilakukan secara ilmiah dengan menggunakan perspektif keilmuan yang merujuk kepada teori-teori maupun standar atau kriteria yang terukur.

## **2. Peranan Analisis Kebutuhan.**

Analisis kebutuhan adalah alat yang konstruktif dan positif untuk melakukan perubahan, yang dimaksud dengan perubahan di sini bukanlah perubahan yang radikal dan tidak berdasar, tetapi perubahan yang didasarkan atas logika yang bersifat rasional, perubahan fungsional yang dapat memenuhi kebutuhan. Perubahan itu menunjukkan upaya formal yang sistematis menentukan dan mendekatkan jarak kesenjangan antara seperti apa yang ada dengan bagaimana seharusnya (Arikunto dan Jabar, 2009:72).

### **3. Langkah pelaksanaan analisis kebutuhan.**

Anderson sebagaimana dikutip Arikunto dan Jabar (2009:76-77) menjelaskan langkah pelaksanaan analisis kebutuhan dapat dilakukan dengan dua teknik. Kedua teknik tersebut adalah analisis kebutuhan secara objektif dan analisis kebutuhan secara subjektif. Berikut penjelasannya.

#### **a. Analisis kebutuhan secara objektif.**

Analisis kebutuhan secara objektif dilakukan dengan tahapan-tahapan sebagai berikut:

- 1) Mengidentifikasi lingkup tujuan-tujuan penting dalam program yang akan dievaluasi.
- 2) Menentukan indikator dan cara pengukuran tujuan-tujuan.
- 3) Menyusun kriteria (standar) untuk tiap-tiap indikator dengan acuan pedoman atau acuan apa saja yang ada dalam sistem dan bidang yang dievaluasi.
- 4) Menyusun alat pengukuran untuk tiap-tiap indikator.
- 5) Membandingkan kondisi yang diperoleh dengan kriteria, jika data yang diperoleh lebih rendah dari tingkat standar maka maknanya berarti ada kebutuhan.

#### **b. Analisis kebutuhan secara subjektif.**

Analisis kebutuhan secara subjektif dilakukan dengan tahapan-tahapan sebagai berikut:

- 1) Mengidentifikasi tujuan-tujuan penting dalam program yang akan dievaluasi.
- 2) Menentukan pilihan kriteria atau menyusun kriteria yang sesuai dengan setiap tujuan masing-masing bidang dan indikator. Dalam langkah ini evaluator perlu mengumpulkan banyak bukti formal yang akan digunakan untuk dasar pertimbangan kebutuhan.
- 3) Menyusun skala bertingkat yang digunakan untuk mempertimbangkan tingkat penampilan indikator. Dalam pembuatan skala sedapat mungkin diurutkan sesuai dengan tingkat penerimaan, dan skala tersebut sebaiknya berbentuk interval.
- 4) Jika sudah selesai membuat skala, kumpulkan semua calon evaluator untuk bersama-sama menentukan urutan kebutuhan dan skala prioritas kebutuhan. Jika kebetulan terdapat dua kebutuhan yang sejajar, diperlukan

lagi kesepakatan untuk menentukan masa kebutuhan yang lebih mendesak untuk diprioritaskan dalam penyelesaiannya. Yang terpenting untuk diingat adalah bahwa menentukan urutan kebutuhan jangan sampai ada subjektivitas yang masuk sehingga menyebabkan hasilnya menyimpang dari kenyataan yang ada.

## **B. *Scheduling* (Penjadwalan).**

Kegiatan evaluasi diharapkan selesai dalam waktu yang diberikan. Jika evaluasi tidak selesai dalam waktu yang dialokasikan, tambahan biaya tidak disediakan, kecuali jika dibuat ketentuan khusus untuk itu. Untuk menepati jadwal itulah, maka evaluator harus memutuskan kapan setiap tahapan kegiatan evaluasi akan dilakukan, urutan kegiatan dan lamanya waktu yang diperlukan untuk setiap kegiatan.

Penyusunan jadwal kegiatan evaluasi diperlukan, terutama untuk kepentingan, kelancaran kegiatan evaluasi itu sendiri, mengatur batas waktu penyelesaian setiap kegiatan, dan menentukan jumlah waktu yang diberikan untuk setiap kegiatan.

Kegiatan yang dilakukan selama evaluasi diarahkan untuk mengecek tujuan program serta menjawab pertanyaan-pertanyaan seperti bagaimana pengembangan rancangan strategi evaluasi dan prosedur *sampling*nya, bagaimana perencanaan dan pelaksanaan pengumpulan data/informasi dan analisisnya, dan bagaimana pelaporan hasil evaluasinya.

Batas waktu atau *deadline* untuk menyelesaikan setiap kegiatan mengacu pada waktu di mana kegiatan diharapkan selesai. Jumlah waktu yang diberikan untuk setiap kegiatan dibuat dalam satuan waktu jam, hari atau minggu yang akan dihabiskan untuk kegiatan, dan itu biasanya dihitung berdasarkan pertimbangan kemampuan dan waktu yang tersedia bagi evaluator.

Kegiatan evaluasi, batas dan alokasi waktu biasanya ditetapkan setelah melalui kompromi antara evaluator dan sponsor. Jadwal waktu biasanya disusun evaluator, meskipun sebagian atau seluruhnya dapat ditentukan sebelumnya oleh sponsor. Misalnya, sponsor dapat meminta adanya dua *progress report* dan satu *final report* dalam suatu batas waktu tertentu. Kegiatan evaluasi, batas waktu dan alokasi waktu dapat dikombinasikan dalam beberapa cara untuk menyusun jadwal evaluasi.

Penentuan jadwal evaluasi mulai dari tahap persiapan pelaksanaan sampai pelaporan atau tindak lanjutnya perlu dibicarakan secara jelas

antara sponsor (termasuk *stakeholder*) dengan evaluator. Termasuk di dalam pengelolaan waktu ini adalah masalah kejelasan tentang; kegiatan spesifik yang dikerjakan, kapan dan berapa lama kegiatan spesifik tersebut dilaksanakan, siapa yang akan bertanggungjawab menyelesaikannya.

Pengelolaan waktu dalam evaluasi haruslah menerapkan prinsip efektivitas ketimbang prinsip efisiensi, artinya tercapaian sasaran lebih diutamakan daripada penghematan waktu. Analisis mengenai tugas-tugas atau tahapan kegiatan dan perhitungan-perhitungan secara akurat mengenai lamanya waktu yang diperlukan untuk suatu kegiatan harus dilakukan secara cermat.

Bahkan untuk kepentingan tertentu, perhitungan tentang waktu juga diperlukan untuk menghitung jam kerja evaluator. Rincian jumlah jam kerja masing-masing evaluator pada setiap tahap kegiatan evaluasi, maupun secara keseluruhan kegiatan evaluasi, terkadang diperlukan untuk menentukan beban kerja. Selain itu, rincian jam kerja evaluator dapat juga digunakan sebagai standar perhitungan imbalan gaji/honor.

### **C. Penugasan dan Monitoring.**

Merencanakan dan mengelola evaluasi berarti juga melakukan aktivitas mengawasi atau melihat bagaimana staf secara efisien melakukan kegiatan evaluasi. Berbagai informasi harus dikumpulkan, misalnya informasi tentang jumlah waktu yang dibutuhkan untuk setiap kegiatan, bagaimana setiap kegiatan diselesaikan, dan masalah-masalah apa yang timbul. Informasi ini dapat dikumpulkan dengan berbagai cara, mulai dari sistem pelaporan yang terstruktur sampai dengan pertemuan informal dengan staf.

Siapa yang akan bertanggungjawab dalam evaluasi harus jelas. Suatu daftar tentang siapa menjadi penanggungjawab kegiatan apa haruslah dibuat. Terlebih apabila evaluator melibatkan suatu tim yang cukup besar, maka mekanisme kerja tenaga yang ada harus direncanakan dan dikelola dengan baik.

Daftar pembagian tanggungjawab dengan rincian tugas hendaknya meliputi pekerjaan-pekerjaan; mengkonsep evaluasi, mendesain evaluasi, menyusun instrumen, mengumpulkan data, menganalisis data, menentukan metode pengkodean, penyimpanan dan akses data, menghubungi dan negosiasi dengan audien dan responden, menyiapkan kontrak, menulis laporan, menafsirkan dan menyusun rekomendasi serta mengelola dan berinteraksi dengan segenap personel yang berperan serta.

Tenaga yang berperan serta dalam evaluasi program, pada dasarnya dapat dikelompokkan menjadi dua yaitu tenaga evaluator dan tenaga penunjang atau pembantu. Tenaga evaluator adalah tenaga inti yang bertanggung jawab penuh atas keseluruhan pekerjaan evaluasi, sedangkan tenaga penunjang atau pembantu adalah orang-orang yang menyelesaikan sebagian pekerjaan evaluasi atas dasar permintaan dan arahan dari evaluator.

Apabila tenaga evaluator merupakan tim, maka perlu ada kesepakatan tentang siapa yang menjadi ketua tim dan siapa menjadi anggota. Lebih jauh siapa yang menjadi evaluator pertama, kedua, ketiga dan seterusnya. Hal ini penting terutama dalam kaitannya dengan hak-hak, wewenang dan yang terpenting adalah tanggungjawab dalam melaksanakan evaluasi.

#### **D. *Budgeting* (Pembiayaan).**

Masalah pembiayaan evaluasi merupakan hal yang perlu dikelola secara efektif dan efisien. Dalam perencanaan evaluasi hendaknya dicantumkan dengan jelas segala jenis biaya beserta jumlah biaya yang diperlu. Biaya evaluasi biasanya meliputi komponen-komponen; gaji, konsultan, perjalanan dan *lumpsum*, pencetakan dan pengiriman, rapat dan pertemuan, pengolahan data, alat, bahan dan ongkos sewa, biaya komunikasi, pajak dan lain-lain.

Sebagai bahan pertimbangan untuk menentukan besar kecilnya biaya evaluasi, maka hal-hal yang harus diperhatikan dan diindahkan adalah semua ketentuan yang berlaku tentang satuan biaya pada instansi yang bersangkutan. Selain itu besarnya biaya evaluasi dipengaruhi oleh hal-hal yang sebagai berikut; jenis informasi/data yang dikumpulkan, jumlah informasi yang diperlukan, lokasi sumber informasi, jadwal waktu, jumlah orang yang berperan serta dalam pengumpulan data, analisis data, dan pelaporan.

Dalam hal ini evaluator harus dapat mengelola biaya yang dialokasikan atau yang telah disepakati dalam kontrak. Meskipun dalam kontrak dimungkinkan adanya perubahan melalui *adendum* atau revisi, hal ini ditempuh jika benar-benar telah terjadi sesuatu yang luar biasa seperti kejadian *force majeure*.

Suatu evaluasi biasanya diberikan sejumlah biaya dan jumlah yang telah ditentukan, tidak lebih dan tidak kurang. Untuk merencanakan biaya, evaluator harus mempertimbangkan antara lain keperluan yang dibutuhkan dan ketersediaan sumber, rincian kegiatan, alokasi waktu, dan jumlah staf



yang ditugasi. Seluruhnya diatur sesuai dengan biaya dan jadwal. Biaya evaluasi dibedakan menjadi dua yaitu biaya langsung dan biaya tak langsung.

### **1. Biaya langsung.**

Biaya langsung adalah biaya-biaya yang diperlukan untuk mendanai seluruh kegiatan evaluasi. Pos-pos pengeluaran untuk biaya langsung ini terdiri atas; pengeluaran untuk staf dan nonstaf. Pengeluaran untuk staf meliputi; gaji/honor, biaya lainnya (asuransi, kesehatan, keamanan), biaya konsultan. Sedangkan pengeluaran untuk nonstaf meliputi; sewa, ATK, peralatan, penggunaan komputer, telepon, surat-menyurat, percetakan dan penggandaan, serta perjalanan (*lumpsum*, *tranpor*, *penginapan*).

### **2. Biaya tak langsung.**

Biaya tak langsung adalah biaya yang diperlukan untuk mendukung kegiatan evaluasi, tetapi tidak termasuk dalam pos biaya langsung. Biaya tak langsung biasanya dihitung dalam persentase dari total biaya langsung atau persentase dari gaji/honor, misalnya 20% dari biaya langsung. Dalam hal ini segala sumber yang tersedia dan mungkin dapat digunakan dalam proses evaluasi hendaknya ditentukan, dimanfaatkan berdasarkan kesepakatan yang jelas. Apa saja dokumen-dokumen, buku-buku, sarana prasarana dan lain-lain sumber yang akan dimanfaatkan dalam evaluasi harus direncanakan, dimanfaatkan dan dikelola dengan prinsip efektivitas dan efisiensi.

## **E. Proposal Evaluasi Program**

### **1. Pengertian**

Proposal adalah sebuah rencana kerja yang menggambarkan semua kegiatan yang akan dilakukan dalam pelaksanaan evaluasi program (Arikunto dan Jabar, 2009:79).

### **2. Penyusun Proposal**

Pertanyaan yang sering kali timbul dalam melakukan evaluasi program adalah siapakah yang menyusun proposal evaluasi program? Menjawab pertanyaan ini memerlukan berbagai pertimbangan, diantaranya adalah pertimbangan konsekuensi pembiayaan yang ditimbulkannya. Jika pertimbangan

adalah masalah pembiayaan yang kurang memadai maka penyusunan proposal evaluasi program dan dilakukan dengan memanfaatkan evaluator dari dalam lembaga/institusi tersebut, namun jika masalah pembiayaan tidaklah menjadi ukuran maka penyusunan proposal evaluasi program dapat dilakukan evaluator eksternal atau mengkombinasi antara evaluator internal dan evaluator eksternal.

Berdasarkan pemaparan di atas maka dapatlah dilihat bahwa sesungguhnya terdapat tiga kemungkinan alternatif untuk pihak yang dapat menyusun proposal evaluasi program, yaitu: (1) evaluator internal, (2) evaluator eksternal, dan (3) kombinasi dari evaluator internal dan evaluator eksternal.

### **3. Format Rancangan Proposal**

Format proposal evaluasi program pada umumnya bervariasi, namun demikian tidak jauh berbeda substansinya. Substansi atau sistematika proposal bergantung pada ketentuan yang diberlakukan di institusi tertentu, maupun tujuan dilakukannya evaluasi program.

Namun demikian setidaknya terdapat 6 (enam) komponen yang terdapat format rancangan proposal evaluasi program menurut Arikunto dan Jabar (2009:61-62) yaitu: (1) judul kegiatan, (2) rasional dilaksanakan evaluasi, (3) tujuan, (4) pertanyaan evaluasi, (5) metodologi yang digunakan, dan (6) prosedur dan langkah-langkah kegiatan. Berikut penjelasan dari masing-masing bagian tersebut:

#### **a. Judul kegiatan.**

Pada bagian ini menyebutkan isi pokok kegiatan evaluasi yang mencantumkan nama kegiatan, program apa yang dievaluasi dan dapat juga mencantumkan model yang digunakan serta menyebutkan unit dan lokasi program.

Contoh:

Evaluasi program koperasi di UIN Sumatera Utara.

#### **b. Rasional dilaksanakannya evaluasi.**

Pada bagian ini menjelaskan adanya kebijakan tentang program yang menjadi objek sasaran, perkiraan adanya hambatan tentang pelaksanaan atau alasan mengapa perlu dilaksanakan evaluasi. Di samping itu untuk meyakinkan pembaca bahwa urgensi dilakukannya evaluasi program adalah memaparkan atau menunjukkan adanya kesenjangan.

Kesenjangan yang dimaksudkan di sini adalah penjelasan tentang kondisi yang diharapkan dengan kondisi yang terjadi atau yang ada saat ini. Apabila evaluator penyusun proposal dapat menunjukkan bahwa kondisi yang terjadi saat ini masih jauh dari kondisi yang diharapkan maka kesenjangan dimaksud semakin jelas, sehingga terdapat alasan yang kuat dan dapat diterima untuk dilakukannya evaluasi program tersebut.

Contoh:

Memaparkan nama dan jenis koperasi UIN Sumatera Utara tersebut apakah koperasi simpan pinjam, konsumsi atau koperasi umum. Pemaparan dikuatkan dengan memberikan penjelasan tentang koperasi dengan akte pendirian. Kemudian dipaparkan data mengenai koperasi UIN Sumatera Utara belum pernah di evaluasi atau bisa juga bukti adanya kesenjangan antara kenyataan dengan harapan sehingga perlu dilakukan evaluasi. Tunjukkan kemanfaatan evaluasi yang akan dilaksanakan sekurang-kurangnya bagi pelanggan. Pemaparan dilengkapi dengan penjelasan mengenai program koperasi UIN Sumatera Utara itu sendiri.

### **c. Tujuan.**

Taylor dkk (1996) mengidentifikasi beberapa dimensi umum yang biasanya digali dalam tujuan evaluasi suatu program yaitu:

1) Dampak/pengaruh program.

Dalam dimensi ini, evaluator akan mengkaji seberapa jauh program yang akan, sedang atau telah dijalankan memiliki konsekuensi terhadap konteks, partisipan dan subjek, sistem atau lainnya.

2) Implementasi program.

Evaluator melakukan kajian terhadap seberapa jauh pelaksanaan program ini akan dan sedang dijalankan.

3) Konteks program.

Evaluator mengamati dan mengkaji kondisi konteks (lingkungan) dari program yang akan, sedang dan telah dijalankan, seberapa jauh keterkaitannya, dan apa sajakah konteksnya.

4) Kebutuhan program.

Evaluator mengkaji tentang faktor-faktor penentu keberhasilan program dan keberlanjutannya di masa yang akan datang.

Merujuk kepada penjelasan di atas maka dapatlah dimaknai bahwa ruang lingkup yang harus tercantum dalam rumusan tujuan evaluasi adalah dampak, pengaruh program, implementasi program, konteks program dan kebutuhan program yang dirumuskan tujuan umum dan tujuan khusus dari evaluasi program. Dalam tujuan khusus disebutkan secara rinci yang harus dicapai dari evaluasi yang dilakukan. Banyaknya butir tujuan tidak dibatasi, tetapi tergantung kepada ruang lingkup kajian evaluasi yang dilakukan.

Pemilihan tujuan manakah yang akan mendasari kegiatan evaluasi sebaiknya dilakukan bersama-sama dengan sponsor. Berdasarkan pilihan tujuan yang telah ditentukan selanjutnya menetapkan jenis evaluasi yang akan dilaksanakan yaitu apakah evaluasi formatif ataukah evaluasi sumatif.

Pilihan ini juga telah membatasi cakupan kegiatan evaluasi serta jenjangnya. Beberapa kriteria yang digunakan dalam merumuskan tujuan evaluasi adalah: (1) kejelasan, (2) keterukuran, (3) kegunaan dan kebermanfaatn, dan (4) relevansi dan kesesuaian atau *compatibility*.

Contoh:

Tujuan umum dari evaluasi program yang dilakukan adalah untuk mengumpulkan informasi yang terkait dengan keterlaksanaan program koperasi UIN Sumatera Utara.

Tujuan khusus, bertitik tolak dari tujuan umum maka tujuan khusus dari kegiatan evaluasi ini adalah untuk mengetahui:

- 1) Jumlah dan perilaku pelanggan (banyak pengunjung yang datang sehari-hari, variasi pengunjung yaitu dosen, pegawai, mahasiswa, masyarakat umum, saat-saat ramai pengunjung, saat-saat sepi pengunjung, banyaknya barang yang dibeli dan sebagainya).
- 2) Tingkat kepuasan pelanggan terhadap layanan koperasi, di mulai dari keadaan pelanggan itu sendiri sampai pendapat pelanggan terhadap kualitas koperasi.
- 3) Kualitas barang-barang yang dijual (jenis, banyaknya tiap jenis, penataan, kondisi barang dan harga).
- 4) Kondisi perlengkapan koperasi (jenis perlengkapan, kondisi atau kualitas penataan dan perawatan).
- 5) Kualitas layanan (jumlah dan kualitas yang melayani, jam layanan dan cara melayani).

**d. Pertanyaan evaluasi.**

Setelah tujuan evaluasi dirumuskan, maka evaluator kemudian mengoperasionalkan tujuan evaluasi tersebut ke dalam pertanyaan evaluasi yang akan dijawab dalam kegiatan evaluasi.

Dalam hal ini Arikunto dan Jabar (2009:84-85) memaparkan model pertanyaan yang biasanya muncul dalam evaluasi program yaitu:

1) Pertanyaan tentang dampak/pengaruh.

- Apakah perilaku/aktivitas/orang-orang berubah akibat dari program yang dijalankan?
- Siapakah yang diuntungkan dan bagaimana?
- Apakah semua partisipan program puas dengan apa yang mereka dapat dari program tersebut?
- Apakah capaian program yang didapat sebanding dengan sumber daya yang diinvestasikan?
- Apa yang bisa orang pelajari, dapatkan, dan capai dari hasil program tersebut?
- Apa dampak program ditinjau dari segi sosial, ekonomis, dan lingkungan (baik positif maupun negatif) terhadap orang, masyarakat dan lingkungan?
- Apa kekuatan dan kelemahan dari program?
- Kegiatan apa dari program yang paling banyak atau sedikit berkontribusi terhadap pencapaian tujuan program?
- Jika ada, apa pengaruh tak langsung, baik positif atau negatif dari program?
- Seberapa baik program mampu merespon kebutuhan?
- Seberapa efisienkah sumber daya digunakan dalam pencapaian tujuan program?

2) Pertanyaan tentang implementasi program.

- Terdiri dari aktivitas atau *event* apakah program yang akan/ sedang/telah berjalan itu?
- Metode apa yang digunakan dalam menjalankan program?
- Siapa yang sebenarnya menjalankan program dan seberapa baik mereka melakukannya?

- Siapa yang berpartisipasi dan dalam aktivitas apa? Apa semua pihak yang terlibat memiliki akses yang adil terhadap program?
  - Sumber daya dan input apakah yang diinvestasikan dalam program?
  - Seberapa banyak pihak yang terlibat, siapa saja, dan apa perannya?
  - Apakah sumber daya keuangan dan manusia tersedia dengan cukup?
- 3) Pertanyaan tentang konteks program.
- Seberapa baik program sesuai dengan keadaan setempat? Misalnya Gekonomi sasaran target?
  - Seberapa besar kondisi sosial, ekonomi, politik yang ada berkontribusi atau mempengaruhi keberhasilan program?
  - Bagaimana keadaan wilayah/tempat program itu dijalankan, adakah *setting* yang bisa diubah?
  - Adakah pihak lain yang melakukan hal yang sama seperti apa yang ingin dicapai oleh program yang sedang dijalankan itu? Adakah duplikasi?
  - Siapa pendukung dan penghalang kesuksesan pencapaian program?
- 4) Pertanyaan tentang kebutuhan program.
- Kebutuhan-kebutuhan apa saja yang bisa diidentifikasi melalui program?
  - Bagaimanakah karakteristik dari populasi target program?
  - Aset apakah yang ada di konteks program dan kelompok target yang bisa dikembangkan?
  - Apa yang selama ini telah dijalankan terkait dengan pelaksanaan program?
  - Perubahan apa yang dianggap sasaran target yang memungkinkan atau sangat perlu?
  - Apakah program yang dijalankan sudah tepat?

Purwanto dan Suparman (1999:77-78) memaparkan beberapa metode dalam merumuskan pertanyaan evaluasi, sebagai berikut:

1) Menganalisis objek.

Pengertian objek di sini dapat meminjamnya dari model CIPP yaitu terdiri atas komponen-komponen konteks, input, proses dan produk.

Di antara objek-objek tersebut dianalisis sampai dapat ditentukan objek mana yang paling menarik dan perlu dijadikan fokus evaluasi. Penentuan objek ini dapat pula dilakukan setelah dilakukan analisis terhadap kekuatan dan kelemahannya masing-masing. Dengan mengidentifikasi hal-hal penting dari objek yang dievaluasi, evaluator dapat merumuskan pertanyaan evaluasi.

Contoh: evaluator menetapkan objek evaluasi yakni proses pelaksanaan pelatihan. Setelah dianalisis, maka hal paling menarik dalam proses tersebut adalah masalah metode pelatihan dan bahan yang dipergunakan.

Oleh karena itu selanjutnya pertanyaan evaluasi dirumuskan berkaitan dengan hal-hal penting dalam proses pelatihan tersebut yaitu pertanyaan yang berkaitan dengan metode pelatihan yang digunakan dan bahan pelatihan yang dipakai.

2) Menggunakan kerangka teoretis.

Pertanyaan evaluasi dirumuskan berdasarkan model tertentu, teori-teori atau asumsi-asumsi tertentu yang dikemukakan oleh ahli, baik dalam publikasi ilmiah (jurnal), buku, makalah ilmiah maupun karya ilmiah lainnya.

Contoh: pertanyaan evaluasi tentang program pelatihan didasarkan kepada model CIPP. Pertanyaan lainnya adalah evaluasi tentang program pelatihan yang dilaksanakan dengan sistem jarak jauh dirumuskan berdasarkan suatu teori tentang efektivitas pendidikan jarak jauh yang dikemukakan ahli.

3) Memanfaatkan keahlian dan pengalaman dari luar.

Evaluator dapat merumuskan pertanyaan evaluasi dengan bantuan dari orang lain yang ahli di bidang yang dievaluasi, atau yang memiliki pengalaman dalam mengevaluasi program sejenis. Ahli atau orang berpengalaman dari luar sistem yang memiliki minat dan perhatian terhadap program dapat diminta partisipasinya dalam kegiatan perumusan pertanyaan evaluasi.

Contoh: evaluator mengundang konsultan dan mereka diminta mengkaji pertanyaan evaluasi yang telah disusun, memberikan saran untuk memperbaiki pertanyaan tersebut. Contoh lain: pertanyaan evaluasi disusun atas dasar pendapat kritis dari pengamat (dari luar sistem pelatihan) yang secara intens mengikuti perkembangan program pelatihan.

4) Berinteraksi dengan sponsor atau audien kunci.

Evaluator berdiskusi dengan audien atau orang-orang penting tentang apa yang menjadi perhatian dan kepeduliannya, serta menanyakan kepada mereka pertanyaan apa yang ingin dijawab melalui evaluasi. Dengan berinteraksi dengan orang-orang penting tersebut maka evaluator akan dapat menentukan pertanyaan evaluasi yang benar-benar memenuhi kebutuhan audien.

Contoh: evaluator mewawancarai beberapa orang yang menjadi sponsor dan akan memanfaatkan hasil evaluasi. Orang-orang yang diwawancarai tersebut dipilih berdasarkan tingkat kebutuhan dan kepeduliannya terhadap evaluasi.

5) Mendefinisikan tujuan evaluasi.

Dengan menggunakan penalaran logika, evaluator merumuskan pertanyaan evaluasi berdasarkan kepada pemahaman logis tentang tujuan evaluasi. Pendefinisian ini dapat dilakukan dalam suatu kelompok dengan cara curah pendapat.

Contoh: evaluator mengadakan pertemuan dengan stafnya untuk curah pendapat tentang pertanyaan evaluasi. Tujuan evaluasi yang telah ditetapkan dielaborasi dan ditentukan pertanyaan-pertanyaan pokoknya yang perlu dijawab melalui evaluasi.

6) Membuat pertanyaan tambahan.

Setelah memiliki sejumlah pertanyaan evaluasi, evaluator dapat menambahkan pertanyaan lain sembari melakukan kegiatan evaluasi. Beberapa pertanyaan evaluasi dirumuskan sambil jalan dan diperoleh dari lapangan.

Contoh: evaluator menggali pertanyaan evaluasi dari orang-orang yang berperan serta dalam kegiatan evaluasi. Pertanyaan tambahan ini bisa berasal dari peserta, penyelenggara, instruktur atau staf evaluator di lapangan.

Contoh pertanyaan evaluasi:

- 1) Bagaimanakah variasi jumlah dan perilaku pelanggan koperasi UIN Sumatera Utara?
- 2) Bagaimanakah tingkat kepuasan pelanggan terhadap layanan koperasi UIN Sumatera Utara?



- 3) Bagaimanakah kualitas barang-barang yang dijual di koperasi UIN Sumatera Utara?
- 4) Bagaimanakah kondisi perlengkapan koperasi UIN Sumatera Utara?
- 5) Bagaimanakah kualitas layanan koperasi UIN Sumatera Utara?

#### **e. Metodologi yang digunakan.**

Bagian ini memaparkan tentang objek sasaran evaluasi yang dihasilkan dari identifikasi komponen program dan indikator, sumber data, metode yang digunakan, instrumen yang digunakan sebagai pelengkap metode pengumpulan data dan menentukan teknik analisis data. Pemaparan terkait dengan metodologi dalam evaluasi program termasuk bagian terpenting dalam perencanaan evaluasi program.

Objek sasaran evaluasi dapat meliputi berbagai hal dan berbagai jenis. Menurut isinya sasaran objek tersebut dapat dikelompokkan menjadi informasi deskriptif dan informasi tentang pendapat. Informasi deskriptif meliputi data dan informasi tentang berbagai hal yang berkaitan dengan karakteristik, keadaan, dokumentasi yang berguna bagi penetapan tujuan, penentuan strategi, penentuan implementasi dan tindak lanjut.

Informasi tentang pendapat berupa opini, keyakinan, norma yang berkaitan dengan tujuan, strategi pelaksanaan dan kelanjutan program. Dipaparkan juga deskripsi dan pendapat apa saja yang dikumpulkan dalam evaluasi yang akan dilaksanakan. Jika memang ada informasi tentang pendapat, nyatakan siapa sumbernya.

Sumber data perlu dinyatakan apakah *person*, *paper* atau *place*. Perlu juga dijelaskan seberapa banyak informasi akan dikumpulkan dan apakah digunakan teknik *sampling*. Jika digunakan *sampling*, maka dijelaskan prosedur yang ditempuh.

Kriteria yang dapat digunakan dalam menentukan prioritas sumber data antara lain: kredibilitas, kepraktisan, ketepatan waktu, ketepatan isi, kemudahan di analisis, objektivitas, kejelasan, cakupan isi, ketersediaan, kegunaan, keseimbangan dan efektivitas biaya. Tentu saja kriteria tersebut dapat dijadikan sesuai dengan situasi dan kondisi (Purwanto dan Suparman, 1999:95).

Paparan mengenai teknik pengumpulan data merupakan bagian terpenting yang menguraikan dua hal yaitu teknik yang ditempuh evaluator

dalam memperoleh data dan prosedur pengumpulan data. Uraian tentang prosedur yang ditempuh untuk memperoleh data merupakan upaya untuk memberikan jawaban atas pertanyaan evaluasi.

Di samping itu perlu dijelaskan pula kepada siapa pertanyaan-pertanyaan evaluasi tersebut diarahkan, bagaimana jadwal pengumpulan data dan siapa respondennya. Nyatakan teknik pengumpulan data yang digunakan misalnya wawancara, observasi, kuesioner, tes atau lainnya.

Pada bagian metodologi ini juga dipaparkan instrumen pengumpulan data yang digunakan untuk pengumpulan data yaitu terkait dengan bentuk atau jenis instrumennya. Demikian juga halnya dengan pemaparan prosedur yang ditempuh evaluator dalam memperoleh instrumen yang akan digunakan untuk pengumpulan data.

Apabila instrumen yang digunakan dalam pengumpulan data dirancang sendiri, maka dipaparkan hal-hal terkait langkah-langkah menyusun instrumen tersebut seperti membuat spesifikasi atau kisi-kisi, menulis butir, mengkaji dan merevisi dan pemaparan terkait dengan syarat validitas dan reliabilitasnya terpenuhi, apakah dilakukan ujicoba dan bagaimana cara memvalidasinya.

Di samping merancang sendiri instrumen yang digunakan untuk mengumpulkan data maka evaluator dapat memperoleh instrumen yang sudah baku melalui rekan sejawat, lembaga penerbit tes, katalog, perpustakaan dan pusat-pusat sumber belajar lainnya.

Pada bagian metodologi ini juga dipaparkan teknik analisis data yang digunakan. Penentuan teknik analisis data ditentukan oleh tingkat pengukuran atau *level of measurement* dan jenis datanya. Teknik analisis yang sering digunakan dalam evaluasi program adalah statistik deskriptif seperti distribusi frekuensi, mean, median, modus, simpangan baku, persentil rank dan skor standar.

#### **f. Prosedur kerja dan langkah-langkah kegiatan.**

Bagian ini memaparkan hal-hal yang terkait dengan proses yang akan dilalui oleh evaluator, berupa prosedur kerja dan langkah-langkah kerja. Di samping itu diikuti dengan estimasi waktu pelaksanaannya. Hal ini dilakukan agar pentahapan kerja dan langkah-langkahnya diketahui dengan jelas oleh evaluator dan sponsor (pemberi tugas).

Contoh:

**Tabel 4. 1 Rencana Evaluasi Program**

No	Jenis Kegiatan	Bulan															
		Januari				Februari				Maret				April			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	Penyusun Proposal																
2	Penyiapan instrumen																
3	Pengumpulan data																
4	Analisis data																
5	Penulisan dan Pengadaan laporan																

Guba dan Lincoln sebagaimana dikutip Sukmadinata (2006:131) memaparkan 12 (dua belas) langkah perencanaan dalam desain evaluasi program yaitu:

1. Pembuatan kontrak.  
Membuat kontrak dengan sponsor atau klien yang membutuhkan evaluasi.
2. Pengorganisasian.  
Memilih dan melatih tim evaluator, menyusun rancangan awal, menyusun kebutuhan logistik, mengidentifikasi faktor-faktor sosial-politis setempat yang mungkin berpengaruh.
3. Mengidentifikasi pengguna dan pihak terkait.  
Mengidentifikasi perantara, pengguna, pihak yang diuntungkan dan dirugikan, memilih strategi yang akan digunakan, memperhitungkan kegagalan dan sanksi bila gagal, menyusun persetujuan formal.
4. Pengembangan kerjasama.  
Merancang siklus hermaneutik, menyusun siklus, membangun kerjasama, mencek kredibilitas pelaksana evaluasi.
5. Memperluas kerjasama dengan pengguna dan sponsor berdasarkan informasi baru.  
Penyempurnaan siklus, menggunakan informasi dokumenter, melaksanakan wawancara dan observasi, kajian literatur, penyusunan etika evaluator.
6. Menyaring keluhan-keluhan, kepedulian dan isu-isu.  
Mengidentifikasi keluhan-keluhan, kepedulian, isu-isu, pemecahan

melalui konsensus, pembuatan catatan-catatan samping sebagai komponen laporan.

7. Memberikan prioritas pada butir-butir yang belum terpecahkan.  
Proses penentuan prioritas secara partisipatif menyusun butir-butir prioritas, mengecek kemampuan mengatasi yang menjadi prioritas.
8. Mengumpulkan informasi dan melengkapinya.  
Mengumpulkan informasi, melatih penggunaannya melalui menggunakan siklus hermaneutik lebih lanjut, mengumpulkan informasi yang ada, menggunakan instrumen yang ada dan yang baru, melakukan studi kasus.
9. Menyiapkan agenda untuk negosiasi.  
Merumuskan dan menjelaskan butir-butir yang belum terpecahkan, menjelaskan kegiatan yang dipilih, menjelaskan, memperkuat butir yang dipilih, membuang yang tidak cocok, menyiapkan pelatihan lengkap, mencek agenda.
10. Melakukan negosiasi.  
Memilih siklus yang tepat, melaksanakan siklus, membuat penyusunan bersama, mencek kemampuan, menentukan tindakan.
11. Menyusun laporan.  
Laporan kasus-kasus dan laporan lengkap.
12. Pengulangan.  
Pengulangan seluruh proses.

Purwanto dan Suparman (1999:91-98) memaparkan 11 (sebelas) komponen-komponen yang terdapat dalam desain evaluasi adalah: (1) menentukan objek atau sasaran evaluasi, (2) merumuskan tujuan evaluasi, (3) menentukan penerima laporan evaluasi, (4) merumuskan pertanyaan evaluasi, (5) menentukan informasi yang dikumpulkan, (6) menentukan teknik pengumpulan informasi, (7) menentukan instrumen, (8) menentukan teknik analisis data, (9) cara interpretasi dan penarikan kesimpulan, (10) menentukan jenis/bentuk laporan, dan (11) menentukan teknik pelaporan. Berikut penjelasannya sebagai berikut:

### **1) Menentukan objek atau sasaran evaluasi.**

Apa yang akan dievaluasi? Apabila menjadi pertanyaan ini berarti telah menentukan objek evaluasi. Nyatakanlah apa yang akan dievaluasi

tersebut dan rumuskan secara pasti. Untuk menentukan objek ini perlu ada kesepakatan dengan penanggungjawab program yang menjadi sponsor kegiatan evaluasi.

Setelah objek evaluasi dirumuskan, selanjutnya objek tersebut dideskripsikan detilnya. Dalam mendeskripsikan objek ini hendaknya dijelaskan secara singkat hal-hal yang berkenaan dengan pertanyaan, apa, mengapa, siapa, kapan, dan di mana?

Contoh rumusan objek evaluasi beserta deskripsi ringkasnya:

Objek evaluasi ini adalah kegiatan pendidikan dan latihan profesi guru (PLPG) yang dilaksanakan oleh Fakultas Ilmu Tarbiyah dan Keguruan (FITK) Universitas Islam Negeri Sumatera Utara. Evaluasi program ini diarahkan kepada peninjauan efektivitas pelaksanaan PLPG.

## **2) Merumuskan tujuan evaluasi.**

Apa yang akan dicapai oleh kegiatan evaluasi dan mengapa evaluasi dilakukan? Apabila telah dijawab pertanyaan ini, berarti telah merumuskan tujuan evaluasi. Lalu dalam merumuskan tujuan evaluasi tersebut apa kriterianya? Kriteria tujuan evaluasi adalah harus jelas dapat dipahami, dapat dicapai, berguna dan relevan atau cocok, dan menguntungkan.

Contoh:

- Memperoleh informasi tentang input, proses, dan produk pelatihan, yang akan berguna bagi perancangan kembali dan penyempurnaan sistem pelatihan.
- Memonitor dan mengawasi jalannya program, mendeskripsikan masalah dan merevisi kebutuhan, dan menentukan tujuan mana yang telah berhasil dicapai.

Tujuan evaluasi tersebut perlu dijabarkan menjadi informasi yang lebih terbatas (terdefinisi) yang biasa disebut variabel. Jika variabel tersebut dianggap belum cukup rinci dan jelas, maka dapat dirinci menjadi indikator-indikator.

## **3) Menentukan penerima laporan evaluasi.**

Bagian ini memberikan jawaban atas pertanyaan; untuk siapa evaluasi dilakukan dan siapa yang akan memanfaatkan hasilnya? Penerima laporan

yang harus diperhatikan adalah orang-orang yang berperan sebagai sponsor, pembuat keputusan, peserta, orang-orang yang berperan serta dalam perencanaan dan pelaksanaan program, orang-orang yang tertarik dan berminat serta berkepentingan dengan program, orang-orang yang karena kedudukannya mempunyai hak atas informasi hasil evaluasi.

#### **4) Merumuskan pertanyaan evaluasi.**

Bagian ini hendaknya memuat seluruh pertanyaan evaluasi yang diajukan. Pertanyaan-pertanyaan tentang keadaan dan nilai serta objek yang akan dievaluasi harus tercantum secara lengkap. Evaluator harus mengidentifikasi pertanyaan penting yang berhubungan dengan tujuan evaluasi.

Contoh: sebuah evaluasi terhadap program pelatihan jarak jauh bagi pegawai baru di sebuah bank. Tujuan evaluasinya adalah untuk menilai efektivitas program pelatihan tersebut. Pertanyaan evaluasinya dirumuskan sebagai berikut:

- Apakah satuan biaya (*unit cost*) per peserta program pelatihan jarak jauh tersebut lebih rendah atau lebih tinggi daripada sistem pelatihan sejenis sebelumnya?
- Apakah prestasi/kinerja peserta pelatihan jarak jauh tersebut lebih baik jika dibandingkan dengan peserta pelatihan dengan sistem sebelumnya?
- Apakah pelatihan jarak jauh tersebut memerlukan dukungan administratif?
- Bagaimana penilaian peserta terhadap program pelatihan jarak jauh tersebut?
- Apakah para peserta menyukai program pelatihan jarak jauh tersebut?
- Apakah kompetensi yang diberikan dalam pelatihan jarak jauh tersebut tepat/valid?
- Apakah beberapa bagian dari pelatihan tersebut lebih efektif dari yang lainnya?

Pertanyaan-pertanyaan tersebut masih merupakan pertanyaan umum yang harus dielaborasi atau dirinci lebih detil pada saat evaluator menyusun instrumen. Cara yang dapat ditempuh untuk merumuskan pertanyaan evaluasi adalah melalui:

- Menganalisis objek.

- Mempergunakan kerangka teoretik.
- Memanfaatkan tenaga ahli yang berpengalaman dari luar.
- Berinteraksi dengan orang penting (*key person*).
- Mendefinisikan tujuan evaluasi secara logis.

### **5) Menentukan informasi yang dikumpulkan.**

Informasi tentang objek atau konteks terpilih, dikembangkan dari pertanyaan evaluasi yang telah dirumuskan. Dalam bagian ini tugas evaluator adalah menguraikan tentang pertanyaan terinci dari masing-masing pertanyaan evaluasi yang diajukan. Sebuah pertanyaan evaluasi dapat dijabarkan menjadi beberapa pertanyaan yang lebih spesifik.

Jenis informasi yang dikumpulkan dapat meliputi berbagai hal dan berbagai jenis. Menurut isinya informasi tersebut dapat dikelompokkan menjadi informasi deskriptif dan informasi tentang pendapat.

Informasi deskriptif meliputi data dan informasi tentang berbagai hal yang berkaitan dengan karakteristik, keadaan, dokumentasi yang berguna bagi penetapan tujuan, penentuan strategi, penentuan implementasi dan tindak lanjut.

Informasi tentang pendapat berupa opini, keyakinan, norma yang berkaitan dengan tujuan, strategi pelaksanaan dan kelanjutan. Kemudian dijelaskan atau diuraikan deskripsi dan pendapat apa saja yang dikumpulkan dalam evaluasi yang akan dilaksanakan. Jika memang ada informasi tentang pendapat, nyatakan siapa saja sumbernya.

Sumber informasi perlu dinyatakan apakah ia orang, kinerja, dokumen atau konteks. Perlu pula dijelaskan seberapa banyak informasi akan dikumpulkan dan apakah digunakan teknik *sampling*. Jika digunakan *sampling*, jelaskan prosedur yang ditempuh.

Apabila ternyata begitu banyak informasi yang bisa dikumpulkan dan evaluator menemui kesulitan dalam menentukan informasi manakah yang akan diprioritaskan, maka tips, berikut ini mungkin bermanfaat. Pelajarilah konteks evaluasi dengan membaca literatur terkait. Wawancarailah sponsor, kemudian tentukan pertanyaan apa dan variabel mana yang dinilai penting oleh sponsor. Susunlah urutan atau ranking pertanyaan evaluasi. Tentukanlah sumbernya dan buatlah daftarnya.

Kriteria yang digunakan dalam menentukan prioritas informasi antara lain adalah kredibilitas, kepraktisan, ketepatan waktu, ketepatan isi, kemudahan

dianalisis, objektivitas, kejelasan, cakupan isi, ketersediaan, kegunaan, keseimbangan, dan efektivitas biaya. Tentu saja kriteria tersebut dapat dipilih sesuai dengan situasi dan kondisi.

## **6) Menentukan teknik pengumpulan informasi.**

Bagian ini termasuk bagian penting dalam dalam perencanaan evaluasi yang menguraikan dua hal yaitu teknik yang ditempuh evaluator dalam memperoleh informasi dan prosedur pengumpulan data atau informasi. Uraian tentang prosedur yang ditempuh untuk memperoleh informasi merupakan upaya untuk memberikan jawaban atas pertanyaan evaluasi. Perlu dijelaskan pula kepada siapa pertanyaan-pertanyaan evaluasi tersebut diarahkan, bagaimana jadwal pengumpulan informasinya, siapa respondennya dan sampel (jika ada).

Kemudian dinyatakan pendekatan yang dipilih dalam pengumpulan data, apakah pendekatan kuantitatif atau kualitatif. Teknik pengumpulan data/informasi yang digunakan misalnya dengan wawancara, observasi, kuesioner *self report*, tes atau teknik lainnya.

## **7) Menentukan instrumen.**

Dalam desain atau perencanaan evaluasi program, perlu dinyatakan hal-hal yang berkaitan dengan bentuk atau jenis instrumen yang akan digunakan untuk pengumpulan data. Nyatakan pula prosedur yang ditempuh oleh evaluator dalam memperoleh instrumen yang akan dipergunakan untuk mengumpulkan data.

Jelaskanlah jenis instrumen yang digunakan, karena instrumen mencerminkan tentang bagaimana informasi/data diperoleh dan dicatat. Ada banyak jenis instrumen, misalnya protokol/pedoman wawancara, ceklis, kuesioner, tes dan sebagainya. Setelah itu dijelaskan pula bagaimana instrumen tersebut diperoleh. Biasanya ada dua cara untuk memperoleh instrumen yaitu dengan menyusun sendiri dan bisa pula dengan mengadaptasi, atau memakai instrumen yang sudah baku dan banyak digunakan.

Bila menyusun sendiri instrumen yang digunakan, maka dijelaskan bagaimana syarat-syarat validitas dan reliabilitasnya dipenuhi. Apakah dilakukan ujicoba, dan bagaimana cara memvalidasinya. Langkah menyusun instrumen secara garis besar adalah membuat spesifikasi (kisi-kisi), menulis butir, mengkaji dan merevisi, mengujicoba dan revisi lagi.



Apabila memilih cara mengadaptasi atau memakai instrumen yang sudah baku, maka masalah validitasi dan reliabilitasnya, bukanlah masalah besar. Evaluator dapat memperoleh instrumen dari kolega, teman sejawat, lembaga penerbit tes, katalog, proyek lain, atau program sejenis, perpustakaan dan pusat-pusat sumber belajar.

### **8) Menentukan teknik analisis data.**

Pada perencanaan evaluasi, perlu dijelaskan cara yang akan digunakan untuk memahami informasi yang dikumpulkan. Ada empat tahap utama dalam menganalisis data yaitu mengkaji pertanyaan-pertanyaan (atau tujuannya), menyiapkan analisis deskriptif setiap data dan distribusi frekuensi (untuk data kuantitatif), menyajikan hasilnya, menyiapkan ringkasan isu dasar, kecenderungan dan keterkaitan serta bukti-bukti, dan menilai ketersediaan bukti-bukti untuk menegaskan isu dan pertanyaan evaluasi.

Hal-hal yang perlu diperhatikan dalam menganalisis data adalah jangan terlalu menyederhanakan permasalahan. Sebaliknya harus peka, perhatikanlah keanekaragaman pengaruh dan kondisi, gunakan teknik analisis yang bervariasi. Yakinkan bahwa asumsi yang dituntut oleh teknik yang akan digunakan terpenuhi oleh data yang diperoleh. Gunakan teknik yang cocok dengan tujuan, pilihlah teknik yang praktis, terjangkau dan tidak terlalu rumit.

Penentuan teknik analisis data ditentukan oleh tingkat pengukuran atau *level of measurement* dan jenis data/informasinya. Teknik analisis yang sering digunakan dalam evaluasi program adalah statistik deskriptif, seperti distribusi frekuensi, mean, median, mode, simpangan baku, percentile rank, skor standar, analisis korelasi, analisis isi dan lain-lain.

### **9) Cara interpretasi dan penarikan kesimpulan.**

Suatu desain evaluasi hendaknya memuat dan menjelaskan bagaimana cara mendapatkan nilai-nilai dan arti informasi yang dikumpulkan dan kriteria apa yang dipergunakan untuk menilai. Menginterpretasikan atau menafsirkan adalah memberi makna dari hasil analisis, memutuskan signifikansi dan implikasi data yang disajikan. Menafsirkan meliputi pula tindakan membuat klaim, misalnya tentang manfaat pelatihan, keberhasilan pelatihan dan sebagainya.

Penafsiran data juga melibatkan penilaian atau perbandingan tentang apa yang ditunjukkan oleh data dengan nilai-nilai, standar, atau harapan yang relevan.

### **10) Menentukan jenis/bentuk laporan.**

Desain hendaknya juga menjelaskan tentang bentuk laporan evaluasi yang akan disusun untuk penerima laporan evaluasi. Secara garis besar perlu dijelaskan format laporan yang akan digunakan (sistematikanya) dan isi laporan meliputi apa saja. Laporan yang harus dibuat ada berapa macam, apakah dalam bentuk laporan lengkap ataukah cukup rangkumannya saja (*executive report*), apakah akan dibuat laporan berkala atau laporan sementara, dan laporan intern. Akankah juga dibuat pengumuman atau *press release*, atau dibuat abstrak untuk penerbitan ilmiah seperti jurnal.

### **11) Menentukan teknik pelaporan.**

Desain sebaiknya juga menjelaskan tentang bagaimana laporan evaluasi akan disampaikan kepada penerima laporan evaluasi. Secara garis besar perlu dijelaskan bagaimana dan kapan laporan tersebut akan disampaikan kepada audien, apakah tertulis, lisan dalam bentuk presentasi. Jika tertulis apakah dalam bentuk laporan lengkap ataukah cukup rangkumannya saja (*executive report*). Jika ada presentase, siapa saja penerima laporan evaluasi atau audiennya.

Strahan, Cooper dan Wood sebagaimana dikutip Sukmadinata (2006:133-136) memaparkan perencanaan evaluasi program meliputi langkah-langkah sebagai berikut: (1) klarifikasi alasan melakukan evaluasi, (2) memilih model evaluasi, (3) mengidentifikasi pihak-pihak yang terkait, (4) penentuan komponen yang akan dievaluasi, (5) mengidentifikasi pertanyaan-pertanyaan evaluasi, (6) menyusun desain evaluasi dan jadwal kegiatan, (7) pengumpulan dan analisis data, dan (8) pelaporan hasil evaluasi. Berikut penjelasannya:

#### **a) Klarifikasi alasan melakukan evaluasi.**

Menjelaskan alasan-alasan mengapa evaluasi diadakan. Banyak alasan yang menjadi latar belakang mengadakan evaluasi. Alasan tersebut bisa bersumber dari peneliti sendiri, karena peneliti mempunyai minat yang cukup besar terhadap sesuatu program, peneliti melihat keunggulan atau keberhasilan, atau sebaliknya peneliti melihat adanya kelambanan, kejanggalan, dampak negatif, bahkan kegagalan. Alasan mengadakan penelitian juga

dapat bersumber dari pihak luar, karena adanya tawaran dari lembaga atau pimpinan pemegang otoritas, karena adanya keluhan dari masyarakat khususnya masyarakat pengguna.

#### **b) Memilih model evaluasi.**

Alasan melakukan evaluasi program berhubungan erat dengan model evaluasi yang akan digunakan. Alasan karena adanya keunggulan, keberhasilan dan dampak positif dari suatu program, akan menggunakan model atau pendekatan yang berbeda dengan alasan karena adanya kelambanan, kegagalan atau dampak negatif. Pemilihan model atau pendekatan penelitian didasarkan atas:

1. Tujuan evaluasi dan pertanyaan penelitian.
2. Metode pengumpulan data.
3. Hubungan antara evaluator dengan administrator, melihat evaluasi, individu-individu dalam program dan organisasi yang akan dievaluasi.

#### **c) Mengidentifikasi pihak-pihak yang terkait.**

Identifikasi pihak-pihak terkait atau *stakeholder* sangat penting untuk kelancaran pelaksanaan evaluasi. Siapa yang akan dilibatkan dalam perencanaan, dalam pelaksanaan, siapa yang akan menjadi partner, narasumber, sumber data, partisipan, dan lain-lain. Pelaksanaan evaluasi membutuhkan dukungan, bantuan, kerjasama dengan berbagai pihak. Hubungan yang kurang harmonis dengan pihak-pihak tertentu dapat menghambat kelancaran evaluasi, bahkan bisa menggagalkan.

#### **d) Penentuan komponen yang akan dievaluasi.**

Langkah selanjutnya yang cukup penting dalam evaluasi program adalah penentuan komponen yang akan dievaluasi. Sebelum ditentukan komponen yang akan dievaluasi terlebih dahulu perlu diidentifikasi komponen-komponen yang ada dalam suatu program, mana komponen utama dan mana komponen penunjang. Pemilihan komponen yang akan dievaluasi didasarkan atas pertimbangan kesesuaian dengan tujuan evaluasi, manfaat hasil, keluaran dan kompleksitas komponen, keluasan target populasi, waktu serta biaya yang tersedia.

Komponen utama dari suatu program pendidikan meliputi:

1. Tujuan program, merupakan sasaran-sasaran atau hasil-hasil yang ingin dicapai oleh suatu program. Tujuan program harus dirumuskan secara jelas, rinci dan teratur.
2. Sumber program yaitu segala kekuatan yang mendukung pelaksanaan dan keberhasilan program. Sumber program mencakup sumber daya manusia, sarana dan fasilitas, dan biaya. Sumber daya pendukung keberhasilan program pendidikan disebut juga sumber daya pendidikan.
3. Prosedur pelaksanaan program adalah langkah-langkah pelaksanaan program yang di dalamnya tergambar metode, teknik, strategi yang digunakan bagi keberhasilan program.
4. Manajemen program, adalah sistem yang digunakan dalam merencanakan, mengorganisasikan, melaksanakan, memonitor dan menyempurnakan pelaksanaan program pendidikan. Pelaksanaan suatu program pendidikan melibatkan banyak pihak dan tenaga pelaksana. Koordinasi semua pelaksana program dan kerjasama dengan berbagai pihak membutuhkan sistem manajemen yang efisien dan efektif.

#### **e) Mengidentifikasi pertanyaan-pertanyaan evaluasi.**

Rincian dari fokus atau aspek-aspek yang di evaluasi dirumuskan dalam bentuk pertanyaan, hipotesis atau tujuan. Perumusan pertanyaan tujuan evaluasi program yaitu tahapan divergen dan konvergen.

Tahap divergen, pertanyaan penelitian dirumuskan secara komprehensif. Sebanyak mungkin pertanyaan, isu, informasi, kepedulian, dan masalah berkenaan dengan program yang akan dievaluasi diajukan. Ke dalam pertanyaan-pertanyaan atau informasi-informasi tersebut termasuk kriteria ketercapaiannya. Tahapan kedua adalah tahap konvergen, dalam tahapan ini pertanyaan-pertanyaan, isu-isu atau informasi-informasi yang diajukan pada tahap pertama diseleksi mana yang layak dan penting diajukan dan mana yang tidak.

Beberapa pertanyaan penting yang dapat diajukan dalam evaluasi program yaitu:

1. Tujuan atau sasaran-sasaran apa yang ingin dicapai oleh program pendidikan?
2. Kegiatan-kegiatan utama apa yang dilakukan untuk mencapai sasaran atau target tersebut?

3. Strategi, metode, teknik apa yang digunakan dalam kegiatan-kegiatan tersebut?
4. Bagaimana kondisi sumber daya pendidikan pendukung pelaksanaan program?
5. Bagaimana manajemen pelaksanaan program dan sumber daya pendukungnya?
6. Bagaimana tingkat ketercapaian tujuan atau sasaran program dengan kegiatan dan strategi yang telah dilakukan?

**f) Menyusun desain evaluasi dan jadwal kegiatan.**

Desain evaluasi program pendidikan tidak jauh berbeda dengan desain penelitian, berisi langkah-langkah kegiatan yang akan dilakukan, sasaran evaluasi (aspek atau komponen serta sampel evaluasi), teknik pengukuran atau pengumpulan data yang digunakan, serta para evaluator baik evaluator internal (orang yang terlibat dalam program) maupun evaluator eksternal (peneliti, ahli dari luar). Pelaksanaan kegiatan evaluasi disusun dalam jadwal yang rinci dan kronologis.

**g) Pengumpulan dan analisis data.**

Sebelum pengumpulan data dilakukan kegiatan penting yang harus dilakukan adalah penyusunan instrumen evaluasi. Instrumen evaluasi dapat berbentuk tes dan non tes. Instrumen tes bersifat mengukur, menghasilkan data hasil pengukuran berbentuk angka yang dapat dianalisis secara statistik. Instrumen tes membutuhkan validasi instrumen yaitu suatu proses untuk menguji validitas dan reliabilitas instrumen. Instrumen non tes membutuhkan validasi instrumen walaupun tidak menggunakan analisis statistik seperti pada instrumen tes.

Pengumpulan data dilakukan sesuai dengan jadwal yang telah disusun. Pengumpulan data yang bersifat kuantitatif menggunakan instrumen-instrumen baku (baik instrumen tes maupun non tes), sedang data yang bersifat kualitatif menggunakan multimetode seperti wawancara, observasi, dokumen dan sebagainya.

Data yang diperoleh di analisis secara kuantitatif maupun kualitatif. Analisis kuantitatif menggunakan statistik deskriptif maupun statistik inferensial, analisis kualitatif menggunakan analisis naratif-kualitatif. Hasil analisis kuantitatif berbentuk tabel, profil, bagan, peta (analisis deskriptif),

atau berbentuk skor rata-rata, koefisien korelasi, regresi, perbedaan, analisis jalur, dan sebagainya (analisis inferensial). Hasil analisis kualitatif berupa deskripsi naratif-kualitatif tentang hal-hal yang esensial.

#### **h) Pelaporan hasil evaluasi.**

Berdasarkan temuan-temuan yang diperoleh dari hasil-hasil analisis, maka disusunlah laporan hasil evaluasi. Isi dari laporan penelitian, evaluasi hampir sama dengan laporan penelitian biasa memuat rancangan penelitian, metodologi, temuan-temuan serta kesimpulan dan rekomendasi. Kesimpulan hendaknya berisi jawaban terhadap pertanyaan atau pembuktian hipotesis yang diajukan. Kesimpulan dari pertanyaan kualitatif berisi deskripsi tentang hal-hal yang esensial dari pertanyaan. Rekomendasi berisi masukan-masukan dari temuan-temuan evaluasi bagi penyempurnaan, perbaikan program. Rekomendasi hendaknya memperhatikan segi kelayakan praktis, dirumuskan secara operasional atau rinci.

# BAB V

## INSTRUMEN

### EVALUASI PROGRAM

#### A. Sumber Data

Berdasarkan sumber asal data diperoleh maka sumber data dalam evaluasi program dapat dibedakan atas dua jenis yaitu:

##### 1. Data Internal

Data internal yaitu data yang berasal dari dalam lingkungan sendiri. Seperti diketahui setiap sekolah melakukan aktivitas pencatatan atas segala aktivitas yang dilakukannya baik di bidang personalia, kesiswaan, keuangan, sarana dan prasarana.

Sekiranya Kepala Sekolah menginginkan untuk mengetahui perkembangan siswa dari tahun ke tahun, maka ia dapat melihat dari catatan kesiswaannya. Buku catatan itulah yang merupakan sumber data internal, karena ia berada pada sekolah itu sendiri.

##### 2. Data Eksternal

Data eksternal adalah data yang berasal dari luar lingkungan sendiri. Demi untuk kelancaran pengelolaan sekolah maka setiap Kepala Sekolah memerlukan informasi yang berasal dari luar lingkungan sekolah. Misalnya informasi tentang peraturan atau edaran terkait dengan pengelolaan sekolah yang dikeluarkan oleh Pemerintah dalam hal ini Kementerian Pendidikan.

Informasi-informasi tersebut tidak dipunyai oleh sekolah yang bersangkutan dan harus dicari di luar sekolah. Informasi-informasi demikian itu, dapat diperoleh baik secara langsung maupun secara tidak langsung. Data yang demikian ini disebut data eksternal.

Data eksternal yang diperoleh langsung dari sumbernya disebut data primer. Misalnya sekolah ingin mengetahui tentang peraturan pengelolaan sekolah, maka ia dapat memperoleh langsung dari sumbernya dalam hal ini Kementerian Pendidikan.

Data eksternal yang diperoleh secara tidak langsung dari sumbernya disebut data sekunder. Misalnya data tentang perkembangan hasil ujian nasional diperoleh dari membaca koran, majalah atau melihat siaran televisi.

Data eksternal ada yang diterbitkan dan ada pula yang tidak diterbitkan. Data yang diterbitkan berupa tulisan-tulisan di koran, majalah dan sebagainya. Sedangkan yang tidak diterbitkan berupa dokumen-dokumen, arsip-arsip, laporan dan sebagainya.

Menurut Arikunto dan Jabar (2009:88) sumber data dalam evaluasi program dikenal dengan istilah 3P, rinciannya adalah:

### 1. **Person**

*Person* atau orang dalam hal ini responden yang terlibat secara langsung dengan program yang dievaluasi maupun secara tak langsung berhubungan dengan program. Pengungkapan data dari sumber data *person* ini dapat dilakukan dengan melakukan wawancara maupun menggunakan angket/kuesioner.

### 2. **Paper**

*Paper* atau kertas, dalam hal ini bukan hanya dibatasi pada dokumen dalam bentuk kertas saja tetapi lebih dari itu adalah segala bentuk simbol berupa grafis, tulisan, gambar, tabel, denah, motif dan sebagainya. *Paper* dimaksudkan juga bukan ditulis pada media kertas saja, tetapi dapat juga ditulis di media batu, kayu, plastik dan sebagainya bahkan ditulis dalam media yang saat ini modern seperti di *compact disk*, *hard disk*, *flash disk* maupun sarana digital lainnya termasuk *e-mail*. Untuk sumber data *paper* ini maka metode yang tepat digunakan dalam mengungkapkan atau mengumpulkan data adalah melalui dokumentasi.

### 3. **Place**

*Place* atau tempat/lokasi. Dalam hal ini yang dimaksudkan dengan tempat bukan hanya terbatas pada ruangan, tetapi hal lain yang lebih dalam berada di suatu tempat (ruang). Istilah *place* adalah untuk mempermudah pengumpul data menelusuri lebih jauh apa yang menjadi objek pengamatan, tempat, benda diam, bergerak atau kegiatan. Jadi



jika di dalam rencana tertera tempat sebagai sumber data, harus langsung dilanjutkan dengan penelusuran spesifikasi tempat tersebut.

Metode yang digunakan untuk mengungkapkan data dari sumber data *place* adalah menggunakan observasi atau pengamatan.

## **B. Instrumen Pengumpulan Data**

Instrumen yang biasa dipakai dalam prosedur pengumpulan data dan informasi dalam evaluasi program menurut Tayibnapis (2000:102-103) adalah:

### **1. Surveys.**

Survey dilakukan dengan *open ended instruments* (instrumen terbuka) dan *forced choice instruments* (instrumen pilihan).

### **2. Interviews.**

Interviews dilakukan dengan: (1) *closed formats* yaitu wawancara dengan format tertutup, pertanyaan dan jawaban dibacakan kepada responden, (2) *semi open* (semi terbuka) pertanyaan ditentukan, dan pewawancara membuat interpretasi jawaban ke dalam formulir, dan (3) *open format* (format terbuka), petunjuk umum diberikan kepada pewawancara, jawaban didesain atau dicatat atau dapat juga direkam dengan *tape*.

### **3. Observations.**

Observation dilakukan dengan: (1) *open format*, observer membuat catatan atau reaksi umum, perilaku, dan sebagainya tentang subjek yang dievaluasi, (2) *logs*, semacam buku harian di mana observer mencatat reaksi dan perilakunya sendiri, (3) *sign system*, di mana setiap perilaku khusus dihitung, dilakukan untuk merekam perilaku tertentu dalam tempo waktu yang telah ditentukan, dan (4) *category system*, di mana perilaku diamati, digolongkan ke dalam kategori tertentu untuk membuat rekaman tentang perilaku yang telah ditentukan dalam waktu yang telah ditentukan.

### **4. Tests.**

Tes dilakukan dengan: (1) *multiple choice tests* (tes pilihan ganda), (2) *true-false* (benar-salah), (3) *matching* (tes menjodohkan), *shorts answers*, *fill in blanks* (jawaban pendek, mengisi), dan (5) *essay tests* (tes uraian).

## 5. *Inventories.*

Inventories dilakukan dengan: (1) *open ended* yaitu responden membuat catatan tentang objek tertentu dan item yang mereka temukan, dan (2) *checklist formats*, yaitu responden mengecek atau menghitung dan memberi nomor di sebelah item yang terdaftar dalam instrumen.

## 6. *Site visits, expert reviews, panel hearing.*

Dalam hal ini peneliti atau evaluator itu sendiri sebagai instrumen. Instrumen bentuk ini dapat berupa para ahli, wakil konsumen, anggota/karyawan, publik, orang tua murid dan lain-lain yang terlibat dalam suatu program.

Berikut dipaparkan beberapa instrumen yang digunakan dalam melakukan evaluasi program sebagai berikut:

### 1. **Kuesioner.**

Kuesioner atau angket adalah seperangkat pertanyaan tertulis yang diberikan kepada responden untuk mengungkapkan pendapat, keadaan, kesan yang ada pada diri responden sendiri maupun di luar dirinya (Arikunto, 1988:77).

Penyusunan sebuah kuesioner hendaknya memperhatikan beberapa hal yang menjadi fokus perhatian penyusun kuesioner, sehingga kuesioner yang disajikan kepada responden memiliki tingkat keterbacaan tinggi. Hal-hal yang menjadi perhatian dalam menyusun kuesioner antara lain:

#### a. Membuat kata pengantar.

Kata pengantar ini menjadi urgen karena melalaui pemaparan pada kata pengantar dapat menjelaskan kepada responden tentang maksud dari kuesioner tersebut. Melalui kata pengantar yang dikemas dengan bahasa yang santun dan arif akan menumbuhkan *trust* dari responden sehingga terjalin komunikasi yang baik dan pada akhirnya responden memberikan jawaban yang sesuai dengan keadaan sebenarnya. Jadi kata pengantar merupakan media yang dapat menghubungkan antara peneliti dengan responden.

#### b. Menyertakan petunjuk pengisian angket yang menjelaskan tentang cara menjawab pertanyaan/pernyataan yang terdapat dalam instrumen.

Petunjuk disusun dengan bahasa yang mudah dimengerti, singkat dan jelas serta dilengkapi dengan contoh tentang bagaimana cara mengisi atau menjawab kuesioner tersebut.

- c. Item pertanyaan dalam kuesioner disusun sedemikian rupa sehingga dapat dipahami setiap responden.
- d. Hindari pertanyaan yang dapat menimbulkan kecurigaan, menimbulkan potensi permusuhan atau perselisihan.
- e. Beri penekanan secara khusus pada kalimat atau kata yang difokuskan melalui penggunaan garis bawah atau penebalan.

Secara umum kelebihan penggunaan kuesioner adalah:

- a. Dalam waktu singkat dan serentak dapat diperoleh data yang relatif banyak.
- b. Menghemat waktu, tenaga dan biaya, jika dibandingkan metode wawancara.
- c. Dalam mengisi angket responden dapat memilih waktu senggangnya sehingga tidak terlalu terganggu bila dibandingkan dengan wawancara.
- d. Secara psikologis responden tidak merasa dipaksa dan dapat menjawab lebih terbuka.

Sedangkan kelemahan penggunaan kuesioner adalah:

- a. Dengan adanya bentuk (susunan) pertanyaan yang sama untuk responden yang sangat heterogen, maka penafsiran akan berbeda-beda sesuai dengan latar belakang sosial pendidikan.
- b. Penggunaan kuesioner hanya dapat diterapkan bagi responden yang bisa baca tulis.
- c. Apabila responden tidak dapat memahami pertanyaan atau tidak dapat menjawab, maka akan terjadi kemacetan dan mungkin responden tidak akan menjawab seluruh kuesioner.
- d. Jawaban terhadap kuesioner dapat diisi orang lain. Jika peneliti tidak dapat mengontrol waktu responden mengisi kuesioner.
- e. Sangat sulit untuk menyusun pertanyaan-pertanyaan secara tepat dengan menggunakan bahasa yang jelas sesuai dengan karakteristik responden.

Bentuk kuesioner antara lain: kuesioner terbuka/tak terstruktur, kuesioner tertutup/terstruktur, kuesioner kombinasi terstruktur dan tak terstruktur, dan kuesioner semi terbuka. Berikut pemaparannya:

### a. Kuesioner terbuka/tak terstruktur.

Kuesioner terbuka/tak terstruktur meminta informasi atau pendapat dengan kata-kata responden sendiri. Pertanyaan semacam ini berguna bagi tahap-tahap eksplorasi, tetapi dapat menghasilkan jawaban-jawaban yang sulit untuk disatukan. Contoh kuesioner terbuka/tak terstruktur:

- Bagaimana pendapat anda tentang pembelajaran yang dilakukan guru dengan menerapkan strategi kooperatif?
  - a. \_\_\_\_\_
  - b. \_\_\_\_\_
  - c. \_\_\_\_\_
  - d. \_\_\_\_\_

Dari pertanyaan di atas responden diharapkan mengisi sesuai dengan pendapatnya masing-masing.

### b. Kuesioner tertutup/terstruktur.

Kuesioner tertutup/terstruktur meminta responden untuk memilih kalimat atau deskripsi yang paling dekat dengan pendapat, perasaan, penilaian, atau posisi responden. Responden diminta membubuhkan tanda silang (X) pada alternatif jawaban yang sesuai menurut pendapat responden.

Pilihan jawaban dapat berupa selalu, jarang, kadang-kadang, tidak pernah, sangat setuju, setuju, kurang setuju, tidak setuju, sangat berminat, berminat, kurang berminat, tidak berminat, sangat baik, baik, kurang baik, tidak baik.

Contoh kuesioner tertutup/terstruktur:

- Bapak/Ibu guru memeriksa pekerjaan rumah dan memberikan penilaian.
  - a. Selalu
  - b. Jarang
  - c. Kadang-kadang
  - d. Tidak pernah
- Bapak/Ibu guru memberikan perhatian secara khusus kepada peserta didik yang mengalami kesulitan belajar.
  - a. Sangat Setuju
  - b. Setuju
  - c. Kurang setuju
  - d. Tidak setuju

**c. Kuesioner kombinasi terstruktur dan tak terstruktur.**

Kuesioner kombinasi terstruktur dan tak terstruktur meminta responden untuk memberi alternatif jawaban yang harus dipilih, di lain pihak responden juga diberi kebebasan untuk memberikan jawaban sesuai dengan pendapatnya sendiri. Contoh kuesionernya sebagai berikut:

1. Bapak/ibu guru mengikuti pendidikan dan pelatihan keprofesian?
  - a. Pernah (teruskan ke nomor 2)
  - b. Tidak pernah (lanjutkan ke nomor 3)
2. Pendidikan dan pelatihan keprofesian yang Bapak/Ibu Guru ikuti?
  - a. \_\_\_\_\_
  - b. \_\_\_\_\_
  - c. \_\_\_\_\_
  - d. \_\_\_\_\_

**d. Kuesioner semi terbuka.**

Kuesioner semi terbuka memberi kebebasan kemungkinan menjawab selain dari alternatif jawaban yang sudah tersedia. Contoh kuesionernya adalah:

- Berapa lama Bapak/Ibu mengajar di sekolah ini?
  - a. 1 tahun
  - b. 2 tahun
  - c. 3 tahun
  - d. .... tahun (tuliskan sesuai dengan keadaan Bapak/Ibu)

**2. Wawancara.**

Wawancara merupakan instrumen pengumpulan data yang mengkehendaki komunikasi langsung antara peneliti dengan subjek penelitian. Dalam wawancara biasanya terjadi tanya jawab yang berorientasi pada pencapaian tujuan penelitian. Wawancara sangat tepat diterapkan untuk mengungkap persoalan-persoalan yang sedang diijak dari pada persoalan-persoalan yang sudah dibatasi dari awal.

Beberapa hal yang perlu diperhatikan agar wawancara berlangsung efektif dan efisien adalah:

1. Bersikaplah sebagai pewawancara yang simpatik yang perhatian dan pendengar yang baik, tidak berperan terlalu aktif untuk menunjukkan bahwa anda menghargai pendapat responden.
2. Bersikaplah netral dalam relevansinya dengan materi wawancara. Janganlah menyatakan pendapat sendiri tentang hal itu, atau mengomentari pendapat responden. Upayakan jangan menunjukkan sikap terheran-heran atau tidak menyetujui terhadap apa yang dinyatakan atau ditunjukkan responden.
3. Bersikaplah tenang, tidak terburu-buru atau ragu-ragu dan responden akan menunjukkan sikap yang sama.
4. Mungkin responden yang diwawancarai merasa takut kalau-kalau mereka menunjukkan sikap atau gagasan yang salah menurut pewawancara. Yakinkanlah responden bahwa pendapatnya penting dan bahwa wawancara ini bukan tes atau ujian.
5. Secara khusus perhatikan bahasa yang digunakan untuk wawancara, ajukan frasa yang sama pada setiap pertanyaan, selalu ingat akan garis tujuan wawancara, ulangi pertanyaan apabila responden menjawab terlalu umum atau kabur sifatnya.

Esterberg sebagaimana dikutip Sugiyono (2010:319) menjelaskan 3 macam jenis wawancara yaitu: (1) wawancara terstruktur, (2) wawancara semiterstruktur, dan (3) wawancara tidak terstruktur.

#### **a. Wawancara terstruktur.**

Wawancara terstruktur digunakan apabila peneliti telah mengetahui dengan pasti tentang informasi apa yang akan diperoleh. Oleh karena itu dalam melakukan wawancara peneliti telah menyiapkan instrumen penelitian berupa pertanyaan-pertanyaan tertulis yang alternatif jawabannya telah disiapkan dan pada wawancara terstruktur ini setiap responden diberi pertanyaan yang sama.

#### **b. Wawancara semiterstruktur.**

Wawancara semi terstruktur dilakukan lebih bebas dibandingkan dengan wawancara terstruktur. Tujuan wawancara semi terstruktur adalah menemukan permasalahan secara lebih terbuka, di mana pihak yang diajak wawancara diminta pendapat dan ide-idenya. Dalam melakukan wawancara peneliti mendengarkan secara teliti dan mencatat apa yang dikemukakan oleh responden.

**c. Wawancara tidak terstruktur.**

Wawancara tidak terstruktur adalah wawancara bebas yang dilakukan peneliti dengan tidak menggunakan pedoman wawancara yang telah tersusun secara sistematis dan lengkap untuk pengumpulan datanya. Pedoman wawancara yang digunakan hanya berupa garis-garis besar permasalahan yang akan ditanyakan kepada responden.

**3. Observasi.**

Observasi yaitu pengamatan secara langsung terhadap proses yang berlangsung di *setting* program yang di evaluasi. Observasi dapat dilakukan terhadap klien terkait proses, aktivitas dan interaksinya. Observasi dapat dilakukan menggunakan daftar cek (*checklist*) ataupun catatan terbuka (tulisan bebas). Pedoman observasi menggunakan daftar cek lebih mudah digunakan karena berisi daftar kriteria tertentu, sehingga *observer* (pengamat) hanya memberikan tanda cek pada kriteria yang sesuai dengan pengamatan.

Instrumen observasi yang digunakan dalam evaluasi memiliki kelebihan dan kelemahan. Kelebihan observasi adalah:

- a. Tidak perlu biaya banyak, mudah dilakukan dan dapat digunakan untuk penelitian terhadap berbagai macam gejala.
- b. Tidak banyak mengganggu subjek penelitian.
- c. Dapat secara simultan melakukan pencatatan kepada *observee*.
- d. Banyak gejala yang hanya diteliti dengan observasi sehingga hasilnya yang akurat sulit dibantah.
- e. Banyak objek yang hanya bersedia diambil datanya hanya dengan observasi, karena sulit untuk diwawancarai ataupun mengisi kuesioner.
- f. Kejadian yang serempak dapat diamati dan dicatat secara serempak pula dengan memperbanyak observer.
- g. Banyak kejadian yang dipandang kecil yang tidak dapat ditangkap oleh alat pengumpulan data yang lain, ternyata sangat menentukan hasil penelitian justru diungkap oleh observasi.

Sedangkan kelemahan observasi adalah:

- a. Observasi bergantung pada kemampuan pengamatan dan mengingat. Kemampuan ini dipengaruhi oleh beberapa aspek yaitu:

- 1) Daya adaptasi yaitu kemampuan menyesuaikan diri dengan objek yang akan diamati.
  - 2) Kebiasaan-kebiasaan, yaitu kebiasaan atau pengalaman dalam kehidupan yang berperan dalam pengamatan, tetapi pola ini seringkali tidak mampu menangkap fakta-fakta sebagaimana adanya.
  - 3) Keinginan yaitu memperoleh hasil tertentu dalam penelitiannya, sehingga pengamatannya lebih terarah pada fakta yang sesuai dengan keinginannya.
  - 4) Prasangka terhadap objek yang diamati sehingga pengamatan tidak dapat dilakukan secara objektif dan bahkan terjerumus pada penafsiran secara baik.
  - 5) Proyeksi yaitu kecenderungan melemparkan kejadian di dalam diri observer kepada objek yang berada di luar, sehingga pengamatan tidak dapat dilakukan secara baik.
  - 6) Ingatan yaitu ingatan observer cenderung tidak tahan lama dan tidak luas sehingga fakta-fakta yang dilupakan menjadi tidak tercatat dan fakta-fakta yang dilupakan diganti menurut interpretasi observer.
  - 7) Keadaan fisik dan psikis terutama perasaan yang dalam kondisi fisik letih, sakit, mengantuk, marah dan lain-lain sehingga sulit untuk melakukan pengamatan secara cermat.
- b. Kelemahan dalam pencatatan meliputi:
- 1) Pengaruh kesan umum (*hallo effects*).  
Observasi terpengaruh oleh kesan umum dari objek yang diamati sehingga ia mencatat tidak tepat.
  - 2) Pengaruh keinginan (*generosity effects*).  
Observer ingin membuat baik dalam bentuk kecenderungan memberikan penilaian yang menguntungkan walaupun gejala yang dialami sebenarnya tidaklah demikian.
  - 3) Pengaruh pengamatan sebelumnya (*carry out effects*).  
Kesesatan ini terjadi karena observasi tidak dapat memisahkan kesan terdahulu pada saat mengamati gejala berikutnya.
  - 4) Banyaknya kejadian atau keadaan objek yang sulit diobservasi, terutama yang menyangkut kehidupan pribadi yang sangat rahasia. Di samping itu seringkali terjadi munculnya suatu gejala yang akan diamati pada saat diamati.



- 5) Banyak gejala yang hanya diamati dalam kondisi lingkungan tertentu, sehingga kalau terjadi gangguan yang tiba-tiba, observasi tidak dapat dilaksanakan.
- 6) Observasi memerlukan waktu yang lama, sehingga kadang merasa membosankan karena tingkah laku/gejala yang diharapkan diamati tidak segera muncul.
- 7) Terjadinya subjektivitas dari observer sangat tinggi.

Secara umum terdapat empat jenis metode observasi yaitu: (1) observasi terbuka, (2) observasi terfokus, (3) observasi terstruktur, dan (4) observasi sistematis. Evaluator program dapat memilih salah satu diantaranya atau menggabungkan dua atau lebih metode observasi tersebut dalam pelaksanaan evaluasi program.

**a. Observasi terbuka.**

Observasi terbuka dimulai dari pemikiran netral dan tidak diadakan pengarahannya terlebih dahulu sebelumnya, sehingga *observer* dapat berimprovisasi untuk merekam hal-hal penting dalam proses pembelajaran dalam rangka penerapan tindakan perbaikan. Tujuannya agar *observer* dapat merekonstruksi proses penerapan tindakan perbaikan dalam kerangka diskusi balikan.

Secara umum, format yang digunakan untuk observasi terbuka adalah lembaran kosong yang ditulis pengamat dalam menggambarkan proses pembelajaran. Observasi terbuka dapat juga dilakukan lebih terarah dengan memberikan lembar observasi yang memuat hal-hal yang diamati. Contoh lembar format observasi terbuka sebagai berikut:

**Tabel V. 1. Contoh Lembar Format  
Lembar Observasi Terbuka**

“Program pembelajaran aspek yang diamati keterampilan guru”

No	Aspek Yang Diamati	Catatan Pengamat
1	Melakukan presentasi	
2	Memberikan penjelasan	
3	Menggunakan strategi bertanya	
4	Memberikan umpan balik	
5	Menguasai bahan ajar	

### b. Observasi terfokus

Observasi terfokus adalah observasi yang dilakukan secara spesifik, yaitu observasi yang diarahkan kepada aspek tertentu dalam tindakan guru atau aktivitas siswa dalam proses pembelajaran. Contoh lembar format observasi terfokus sebagai berikut:

**Tabel V. 2. Contoh Lembar Format  
Lembar Observasi Terfokus**

No	Strategi Bertanya	Option			
		SL	SR	JR	TP
1	Jenis Pertanyaan: 1. Faktual (akademik) 2. Opini (akademik) 3. Non-akademik				
2	Jenis jawaban yang diharapkan: 1. Membutuhkan alasan 2. Membutuhkan fakta 3. Membutuhkan pilihan				
3	Pemilihan responden 1. Memberikan kesempatan pada semua siswa setelah mengajukan pertanyaan 2. Memilih siswa untuk menjawab pertanyaan yang telah diajukan 3. Menyebut nama siswa sebelum bertanya				
4	Penggunaan waktu tunggu 1. Memberikan kesempatan yang cukup kepada semua siswa untuk berpikir 2. Berhenti beberapa detik sebelum menyebut nama siswa 3. Tidak memberikan waktu pada siswa untuk berpikir				
5	Bahasa dan cara dalam bertanya 1. Pertanyaan diberikan sebagai tantangan 2. Pertanyaan diberikan untuk menguji pemahaman 3. Pertanyaan diberikan dengan menakuti siswa				

Keterangan

SL = Selalu, SR = Sering, JR = Jarang, TP = Tidak Pernah

c. **Observasi terstruktur**

Observasi terstruktur adalah observasi yang ditandai dengan perekam data yang sederhana tetapi dengan format lebih rinci, sehingga pengamat tinggal membubuhkan tanda cacah atau tanda-tanda lain pada kolom yang disediakan. Gejala yang diamati itu dapat diidentifikasi peristiwa kejadiannya dengan menggunakan format terstruktur. Contoh lembar format observasi terstruktur sebagai berikut:

**Tabel V. 3. Contoh Lembar Format  
Lembar Observasi Terstruktur**

Petunjuk: Berilah tanda *tally* pada kolom sebelah kanan setiap mengamati kejadian sesuai deskripsi observasi

Aktivitas Guru	Tally Kejadian
Menginformasikan tujuan pembelajaran pada siswa	
Memotivasi siswa agar tertarik belajar	
Memberikan review singkat tentang pelajaran sebelumnya (apersepsi)	
Menyajikan materi baru secara singkat dan bertahap	
Mengaitkan materi ajar dengan realitas kehidupan (kontekstual)	
Berbicara kurang dari 10 menit dalam tiap tahapan dan memberikan kesempatan pada siswa untuk bertanya atau mengerjakan latihan/ praktek/ demonstrasi	
Memberikan penjelasan/instruksi secara jelas, rinci, serta menarik perhatian	
Menggunakan media/alat peraga secara efektif dan efisien	
Memfasilitasi siswa melakukan latihan secara bertahap	
Memandu siswa pada tahap latihan	
Memberikan perhatian khusus pda bagian-bagian utama dalam pelajaran	
Memberikan contoh yang bervariasi	
Memberikan petunjuk atau arahan dalam menyelesaikan persoalan	
Menjaga kontak mata pada siswa selama pembelajaran	
Mengajukan beberapa pertanyaan yang relevan dengan topik yang diajarkan	

Menyebutkan nama siswa ketika bertanya	
Memberikan waktu berpikir bagi siswa sebelum menjawab pertanyaan	
Meminta beberapa siswa mengulang kembali jawaban suatu pertanyaan	
Mendorong interaksi/komunikasi antar siswa	
Memberikan penguatan dan menghargai keberhasilan siswa dalam menjawab atau mengerjakan sesuatu yang diminta guru	
Mengecek pemahaman siswa selama pembelajaran	
Berupaya memperoleh respon/perhatian dari semua siswa dan menunjukkan sikap terbuka atas respon siswa.	
Menumbuhkan partisipasi aktif siswa dalam belajar	
Melibatkan siswa dalam menggunakan media/sumber belajar	
Memberikan selingan (humor, istirahat, jeda) ketika siswa lelah	
Menumbuhkan keceriaan dan antusiasme siswa dalam belajar	
Bergerak mendekati siswa dan tidak hanya berada di depan kelas	
Memonitor pekerjaan siswa	
Melakukan refleksi atau membuat kesimpulan dengan melibatkan siswa	

#### d. Observasi sistematis

Observasi sistematis adalah bentuk observasi yang diarahkan pengkategorian bentuk dan jenis data pengamatan yang disusun secara rinci. Penggunaan observasi sistematis dilakukan akibat banyaknya kode yang harus diberikan dalam format observasi. Contoh lembar format observasi sistematis sebagai berikut:

**Tabel V. 4. Contoh Lembar Format Lembar Observasi Sistematis**

[illegible]

Keterangan:

Indikator 1 = siswa aktif mengerjakan tugas kerja kelompok

Indikator 2 = mau menulis

Indikator 3 = interaksi dalam kelompok

Indikator 4 = mau bertanya dan mau menjawab pertanyaan

Indikator 5 = mau memberi ide

Indikator 6 = kreatif dalam pemanfaatan media

Indikator 7 = bertanggung jawab dalam tugas kelompok

Indikator 8 = berpartisipasi dalam kelompok

Indikator 9 = mengerjakan tugas kelompok dengan sungguh-sungguh

#### **4. Dokumentasi.**

Dokumentasi berfokus pada objek yang diamati dalam bentuk dokumen. Dalam hal ini dokumen dalam evaluasi program bukan hanya berupa dokumen dalam catatan tertulis, baik yang tertera pada surat keterangan, artikel, cerita, buku, arsip administrasi, maupun yang tertera pada manuskrip-manuskrip. Lebih dari itu dokumen juga termasuk benda-benda hasil budaya yang mengandung keterangan misalnya porselin, keramik, alat-alat rumah tangga, candi-candi dan sebagainya.

Data yang terkandung dalam dokumen dapat digali, dicacahkan, dikumpulkan dengan menggunakan daftar centang ataupun pedoman dokumentasi yang telah disusun seperti halnya dengan observasi.

#### **5. Tes.**

Di dalam penelitian evaluasi program dikenal ada dua macam tes yaitu tes yang terstandar dan tes buatan. Tes yang terstandar merupakan tes yang sudah dibakukan dan diperoleh dari pihak yang *expert* sebagai pemegang hak. Sedangkan tes buatan sendiri adalah tes yang disusun sendiri oleh evaluator untuk keperluan khusus untuk penelitian evaluasi program.

Bermacam-macam tes yang dapat digunakan untuk mengumpulkan data penelitian evaluasi program yaitu antara lain:

- a. Tes prestasi belajar.
- b. Tes kemampuan akademik.
- c. Tes bakat.
- d. Tes kepribadian.

- e. Tes inventori minat.
- f. Tes skala sikap.

### C. Penyusunan Instrumen.

Sebelum melakukan penyusunan instrumen evaluasi program maka evaluator program terlebih dahulu menentukan instrumen yang digunakan dalam evaluasi program, dalam hal ini sangat dipengaruhi oleh karakteristik program yang dievaluasi, aspek-aspek dari program yang dievaluasi dan data yang ingin dikumpulkan.

Petunjuk umum tentang penyusunan instrumen dipaparkan oleh Brinkerhoff sebagaimana dikutip Tayibnapis (2000:104-105) sebagai berikut:

1. Apa konten yang diperlukan?  
Hal ini langsung berhubungan dengan variabel yang telah ditentukan sebelumnya. Konten instrumen harus dibatasi sebatas apa yang termasuk dalam variabel.
2. Apa dan bagaimana bahasa yang akan dipakai?  
Hal ini tergantung dari responden yang akan menjawab instrumen, apakah responden termasuk golongan yang berpendidikan rendah atau tinggi? Yang penting harus diingat yaitu hindari pemakaian bahasa asing, istilah-istilah asing yang aneh, jangan sampai responden tidak dapat menjawab pertanyaan karena tidak mengerti bahasanya. Usahakan menggunakan bahasa yang mudah, kalimat singkat dan sederhana.
3. Prosedur analisis apa yang akan dipakai?  
Bila akan memakai mesin *scoring*, atau *coding automatic* atau manual maka instrumen harus disiapkan untuk itu.
4. Apakah ada pertimbangan khusus lainnya?  
Dalam hal ini mungkin termasuk versi khusus untuk responden yang cacat (*handicapped*) yang memerlukan petunjuk khusus dan lain sebagainya. Perlu dibuat rencana (*blue print*) dan kisi-kisi untuk setiap instrumen yang akan dibuat, atau mungkin memerlukan konsultasi khusus dari rekan sejawat atau ahlinya.
5. Tentukan seberapa ketepatan yang diperlukan.  
Dalam hal ini diperhatikan kelengkapan, ketepatan waktu, presentasi dan sebagainya.

6. Kapasitas responden.

Responden dilihat dari kemampuannya, pendidikan dan penataran yang telah dilakukan sehubungan dengan hal yang akan diukur.

7. Kesesuaian dengan rencana analisis.

Mengetahui sebelumnya apa yang akan dilakukan terhadap data sesudah terkumpul akan membantu menentukan ketepatan yang diperlukan. Ketepatan pengukuran dapat diperoleh misalnya dengan membuat instrumen yang lebih rinci, petunjuk jawaban, dan kategori.

Djaali dan Muljono (2004:81-85) mendeskripsikan secara garis besar langkah-langkah dalam penyusunan dan pengembangan instrumen sebagai berikut:

1. Berdasarkan sintesis dari teori-teori yang dikaji tentang suatu konsep dari variabel yang hendak diukur, kemudian dirumuskan konstruk dari variabel tersebut. Konstruk pada dasarnya adalah bangun pengertian dari suatu konsep yang dirumuskan.
2. Berdasarkan konstruk tersebut dikembangkan dimensi dan indikator variabel yang hendak diukur yang sesungguhnya telah tertuang secara eksplisit pada rumusan konstruk variabel pada langkah 1.
3. Membuat kisi-kisi instrumen dalam bentuk tabel spesifikasi yang memuat dimensi, indikator, nomor butir dan jumlah butir untuk setiap dimensi dan indikator.
4. Menetapkan besaran atau parameter yang bergerak dalam suatu rentangan kontinum dari suatu kutub ke kutub lain yang berlawanan, misalnya dari rendah ke tinggi, dari negatif ke positif, dari otoriter ke demokratik, dari dependen ke independen dan sebagainya.
5. Menulis butir-butir instrumen yang dapat berbentuk pernyataan atau pertanyaan. Biasanya butir instrumen yang dibuat terdiri atas dua kelompok pernyataan atau pertanyaan yaitu kelompok butir positif dan kelompok butir negatif. Butir positif adalah pernyataan mengenai ciri atau keadaan yang menjadi indikasi sikap atau persepsi positif atau mendekat ke kutub positif, sedang butir negatif adalah pernyataan mengenai ciri atau keadaan yang mengindikasikan persepsi atau sikap negatif atau mendekat ke kutub negatif.
6. Butir-butir yang telah ditulis merupakan konsep instrumen yang harus melalui proses validasi, baik validasi teoretik maupun validasi empirik.

7. Tahap validasi pertama yang ditempuh adalah validasi teoretik yaitu melalui pemeriksaan pakar atau melalui panel yang pada dasarnya menelaah seberapa jauh dimensi merupakan jabaran yang tepat dari konstruk, seberapa jauh indikator merupakan jabaran yang tepat dari dimensi dan seberapa jauh butir-butir instrumen yang dibuat secara tepat dapat mengukur indikator.
8. Revisi atau perbaikan berdasarkan saran dari pakar atau berdasarkan hasil panel.
9. Setelah konsep instrumen dianggap valid secara teoretik atau secara konseptual, dilakukanlah penggunaan instrumen secara terbatas untuk keperluan ujicoba.
10. Ujicoba instrumen di lapangan merupakan bagian dari proses validasi empirik. Melalui ujicoba tersebut, instrumen diberikan kepada sejumlah responden sebagai sampel ujicoba yang mempunyai karakteristik sama atau ekuivalen dengan karakteristik populasi penelitian. Jawaban atau respon dari sampel ujicoba merupakan data empiris yang akan dianalisis untuk menguji validitas empiris atau validitas kriteria dari instrumen yang dikembangkan.
11. Pengujian validitas empiris dilakukan dengan menggunakan kriteria baik kriteria internal maupun kriteria eksternal. Kriteria internal adalah instrumen itu sendiri sebagai satu kesatuan yang dijadikan kriteria, sedangkan kriteria eksternal adalah instrumen atau hasil ukur tertentu di luar instrumen yang dijadikan sebagai kriteria.
12. Berdasarkan kriteria tersebut diperoleh kesimpulan mengenai valid atau tidaknya sebuah butir atau sebuah perangkat instrumen. Jika menggunakan kriteria internal yaitu skor total instrumen sebagai kriteria, maka keputusan pengujian adalah mengenai valid atau tidaknya butir instrumen dan proses pengujiannya biasanya disebut dengan analisis butir. Dalam kasus lainnya, yaitu jika menggunakan kriteria eksternal yaitu instrumen atau ukuran lain di luar instrumen yang dijadikan kriteria, maka keputusan pengujiannya adalah mengenai valid atau tidaknya perangkat instrumen sebagai suatu kesatuan.
13. Untuk kriteria internal atau validitas internal, berdasarkan hasil analisis butir maka butir-butir yang tidak valid dikeluarkan atau diperbaiki untuk diujicoba ulang, sedangkan butir-butir yang valid dirakit kembali menjadi sebuah perangkat instrumen untuk melihat kembali validitas



kontennya berdasarkan kisi-kisi. Jika secara konten butir-butir yang valid tersebut dianggap valid atau memenuhi syarat, maka perangkat instrumen terakhir ini menjadi instrumen final yang akan digunakan untuk mengukur variabel penelitian.

14. Selanjutnya dihitung koefisien reliabilitas. Koefisien reliabilitas dengan rentangan nilai (0 – 1) adalah besaran yang menunjukkan kualitas atau konsistensi hasil ukur instrumen. Makin tinggi koefisien reliabilitas, maka makin tinggi pula kualitas instrumen tersebut. Mengenai batas nilai koefisien reliabilitas yang dianggap layak tergantung pada presisi yang dikehendaki oleh suatu penelitian. Untuk itu dapat merujuk pendapat-pendapat yang sudah ada, karena secara eksak tidak ada tabel atau distribusi statistika mengenai angka reliabilitas yang dapat dijadikan rujukan.
15. Perakitan butir-butir instrumen yang valid untuk dijadikan instrumen final.

Menurut Tayibnapis (2000:105-106) tahapan yang harus dilakukan seorang evaluator dalam menyusun instrumen yang digunakan dalam pengumpulan data dalam evaluasi program sebagai berikut:

1. Sediakan waktu untuk mengecek apakah telah ada instrumen yang serupa, walaupun nanti perlu diperbaiki dan dimodifikasi sesuai dengan masalah. Informasi ini dapat diperoleh melalui rekan sejawat, *publisher* materi dan tes, katalog, proyek lain atau program serupa, perpustakaan, internet dan sebagainya.
2. Setelah instrumen ditemukan, periksalah instrumen tersebut, apakah kontennya sesuai dengan keperluan? Apakah variabel-variabel sama? Apakah ada konten yang hilang? Apakah ketepatan cukup? Dapatkah instrumen tersebut diperoleh pada waktu yang tepat? Mampukah atau terjangkau?
3. Jika tidak ditemukan instrumen yang dimaksud atau instrumen yang ada tidak memadai, maka untuk keperluan evaluasi program yang dilakukan tentulah membuat sendiri rancangan instrumennya.
4. Buatlah tabel kisi-kisi, konten sesuai dengan variabel yang telah ditentukan ketepatannya. Keputusan apa yang akan bergantung pada hasil pengukurannya? Siapakah respondennya akan menentukan bahasa yang akan dipakai? Analisisnya bagaimana? Bagaimana daya rekam instrumen? dan lain-lain.

5. Membuat konsep (*draft*) instrumen, periksalah konsep tersebut, cek kejelasan dan kemudahan membacanya, dan lakukan revisi konsep.
6. Lakukan ujicoba instrumen untuk memeriksa validitas dan realibilitasnya, apakah sudah memenuhi syarat-syarat atau belum lakukan cek dan recek, sehingga validitas dan reliabilitas yang diharapkan tercapai.

Arikunto (1988:71-74) memaparkan langkah-langkah dalam menyusun instrumen pengumpulan data sebagai berikut: (1) merumuskan tujuan, (2) membuat kisi-kisi instrumen, (3) menuliskan butir-butir instrumen, dan (4) menyunting instrumen.

### **1. Merumuskan tujuan.**

Evaluator haruslah merumuskan tujuan yang akan dicapai dengan instrumen yang akan disusun. Tahapan ini diperlukan untuk menyusun instrumen karena tanpa tahu apa data yang terkumpul, apa yang harus dilakukan sesudah itu, apa fungsi setiap jawaban dalam setiap butir dan sebagainya.

### **2. Membuat kisi-kisi instrumen.**

Tahapan berikutnya adalah membuat kisi-kisi instrumen yang berisikan perincian tentang variabel atau aspek yang diukur dalam evaluasi program dan jenis instrumen yang akan digunakan untuk mengambil data dari variabel atau aspek yang diukur tersebut.

### **3. Menuliskan butir-butir instrumen.**

Tahapan berikutnya setelah menuliskan kisi-kisi instrumen maka langkah berikutnya adalah menuliskan butir-butir instrumen yang merujuk kepada kisi-kisi yang telah disusun.

### **4. Menyunting instrumen.**

Apabila pembuatan butir-butir instrumen sudah selesai dilakukan, maka evaluator melakukan pekerjaan terakhir dari penyusunan instrumen yaitu mengadakan penyuntingan (*editing*). Hal-hal yang diperhatikan dalam tahapan menyunting ini adalah:

- a. Mengurutkan butir menurut sistematika yang dikehendaki evaluator untuk mempermudah pengolahan data.

- b. Menuliskan petunjuk pengisian, identitas dan sebagainya.
- c. Membuat pengantar permohonan pengisian instrumen yang akan ditunjukkan kepada responden.

## **D. Validitas dan Reliabilitas Instrumen**

### **1. Validitas Instrumen**

Validitas (*validity*) berasal dari kata *valid* artinya sah atau tepat. Validitas atau kesahihan berarti sejauhmana ketepatan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya. Jadi suatu instrumen yang valid berarti instrumen tersebut merupakan alat ukur yang tepat untuk mengukur suatu objek.

Berdasarkan pengertian ini, maka validitas instrumen pada dasarnya berkaitan dengan ketepatan dan kesesuaian antara instrumen sebagai alat ukur dengan objek yang diukur. Mengukur berat badan tentu tidak valid menggunakan meteran. Di kilang padi, ada timbangan yang valid untuk mengukur berat beras, akan tetapi timbangan ini tidak valid untuk mengukur berat emas dengan bentuk cincin. Mengukur keterampilan siswa, misalnya mengukur unjuk kerja siswa, tentu tidak valid menggunakan tes pilihan ganda. Jadi, tes yang digunakan perlu disesuaikan dengan karakteristik hasil belajar yang diukur.

Beberapa bentuk pengujian validitas instrumen yaitu: (1) validitas konstruksi, (2) validitas isi, dan (3) validitas empirik.

#### **a. Validitas konstruksi (*construct validity*).**

Validitas konstruksi adalah validitas yang mempermasalahkan seberapa jauh item-item instrumen mampu mengukur apa yang benar-benar dimaksudkan yang hendak diukur sesuai dengan konstruk atau konsep khusus atau definisi konseptual yang telah ditetapkan. Untuk menentukan validitas konstruk suatu instrumen harus dilakukan proses penelaahan teoritis terhadap suatu konsep dari variabel yang hendak diukur, mulai dari perumusan konstruk, penentuan dimensi dan indikator sampai kepada penjabaran dan penulisan item-item instrumen.

Perumusan konstruk dilakukan berdasarkan sintesis dari teori-teori mengenai konsep variabel yang hendak diukur melalui proses analisis dan komparasi yang logis dan cermat. Proses selanjutnya adalah dilakukan

penelaahan atau justifikasi *expert* yaitu pakar yang menguasai substansi atau konten dari variabel yang hendak diukur.

### **b. Validitas isi (*content validity*).**

Validitas isi suatu instrumen mempermasalahkan seberapa jauh suatu instrumen mengukur tingkat penguasaan terhadap isi suatu materi tertentu yang seharusnya dikuasai sesuai dengan tujuan. Dengan kata lain instrumen yang mempunyai validitas isi yang baik adalah instrumen yang benar-benar mengukur penguasaan materi yang seharusnya dikuasai sesuai dengan konten yang diukur.

Menurut Gregory sebagaimana dikutip Djaali dan Muljono (2004:66) menjelaskan validitas isi sejauhmana pertanyaan, tugas atau butir dalam suatu instrumen maupun mewakili secara keseluruhan dan proporsional keseluruhan perilaku sampel menjadi tujuan penelitian yang akan diukur pencapaiannya. Artinya instrumen mencerminkan keseluruhan konten atau materi yang diujikan atau yang seharusnya dikuasai secara proporsional.

Untuk mengetahui apakah instrumen itu valid atau tidak, harus dilakukan melalui penelaahan kisi-kisi instrumen untuk memastikan bahwa item-item tersebut sudah mewakili atau mencerminkan keseluruhan konten atau materi yang seharusnya dikuasai secara proporsional. Oleh karena itu validitas isi suatu instrumen tidak mempunyai besaran tertentu yang dihitung secara statistik, tetapi dipahami bahwa instrumen itu sudah valid berdasarkan telaah kisi-kisi instrumen.

### **c. Validitas empirik (*empiric validity*).**

Validitas empiris atau validitas kriteria suatu instrumen ditentukan berdasarkan data hasil ukur instrumen baik melalui ujicoba maupun pengukuran yang sesungguhnya. Validitas empiris diartikan sebagai validitas yang ditentukan berdasarkan kriteria baik kriteria internal maupun kriteria eksternal.

Kriteria internal adalah instrumen itu sendiri yang menjadi kriteria, sedangkan kriteria eksternal adalah hasil ukur instrumen lain di luar instrumen itu yang menjadi kriteria. Ukuran lain yang sudah dianggap baku atau dapat dipercaya dapat pula dijadikan sebagai kriteria eksternal.

Pada garis besarnya, cara-cara menentukan validitas instrumen berupa tes dibedakan kepada dua, yaitu validitas rasional/logis dan validitas empiris atau validitas berdasarkan pengalaman. Validitas rasional dapat dicapai

dengan menjawab pertanyaan berikut ini: (1) apakah tes benar-benar mengukur kompetensi atau hasil belajar yang akan diukur ?, dan (2) apakah bentuk tes sesuai digunakan untuk mengukur hasil belajar siswa?

Untuk menentukan validitas instrumen secara empiris, peneliti harus melakukan uji coba (*try out*). Uji coba dilakukan kepada sebagian responden. Kemudian hasil uji coba tersebut diuji validitasnya. Cara yang dapat kita tempuh untuk menguji validitas tes secara empiris yaitu:

#### **a) Validitas eksternal**

Validitas eksternal dilakukan dengan cara mengkorelasikan skor hasil uji coba instrumen yang dibuat dengan instrumen yang sudah baku. Misalnya seorang peneliti membuat tes yang dipergunakan untuk mendapatkan data hasil belajar sebagai data penelitian. Untuk menguji validitas eksternal tes yang dibuat, dapat dibandingkan dengan tes yang sudah baku, misalnya tes Toefl.

Tes kemampuan berbahasa Inggris yang dibuat peneliti dapat diuji validitas eksternal dengan cara: (1) mengujicobakan secara bersamaan tes yang dibuat guru dan tes Toefl yang telah baku, (b) memberi skor-skor tes buatan dan tes Toefl, (c) mencari angka korelasi antara skor-skor tes buatan dengan skor-skor tes Toefl. Teknik korelasi yang digunakan adalah teknik korelasi *Product Moment*, dan (4) menguji signifikansi angka korelasi yang diperoleh pada langkah ketiga. Jika angka korelasi yang diperoleh ternyata signifikan, berarti tes yang dibuat dapat dianggap valid.

#### **b) Validitas Internal**

Validitas Internal dapat dilakukan dengan dua cara yaitu:

##### **1) Analisis Faktor.**

Analisis faktor dilakukan dengan cara mengkorelasikan skor faktor dengan skor total. Teknik korelasi yang digunakan adalah teknik korelasi *Product Moment* jika skor butir kontinum atau menggunakan teknik koefisien korelasi biserial jika skor butir dikotomi. Jika terdapat korelasi positif dan signifikan, berarti item-item pada faktor tersebut dianggap valid.

##### **2) Analisis Butir**

Analisis butir dilakukan dengan cara mengkorelasikan skor-skor item dengan skor total. Korelasi dilakukan dengan teknik korelasi *Product Moment* jika skor butir kontinum atau menggunakan teknik koefisien

korelasi biserial jika skor butir dikotomi. Jika terdapat korelasi positif dan signifikan antara skor item dengan skor total berarti item tersebut dianggap valid.

a. Pengujian Validitas Tes Berbentuk Objektif Test

Tes berbentuk objektif seperti pilihan ganda (*multiple choice*), benar-salah (*true-false*), menjodohkan (*matching*) merupakan tes dengan skor butir berbentuk dikotomi dengan penilaian 0 dan 1. Jika skor butir dikotomi maka untuk menguji validitas butir tes dilakukan dengan menghitung koefisien korelasi antara skor butir dengan skor total instrumen dengan menggunakan rumus:

$$r_{bis(i)} = \frac{\bar{X}_i - \bar{X}_t}{S_t} \sqrt{\frac{p_i}{q_i}}$$

Keterangan :

$r_{bis(i)}$  : = Koefisien korelasi biserial antara skor butir soal nomor i dengan skor total

$\bar{X}_i$  = rerata skor skor total responden yang menjawab benar pada butir nomor i

$\bar{X}_t$  = rerata skor total seluruh responden

$S_t$  = Standar deviasi dari skor total

$p_i$  = proporsi jawaban yang benar untuk butir soal nomor i

(  $p = \frac{\text{banyaknya siswa yang benar}}{\text{jumlah seluruh siswa}}$  )

$q_i$  = proporsi peserta didik yang menjawab salah ( $q_i = 1 - p_i$ ).

Contoh penggunaannya;

Memberikan skor kepada responden dengan ketentuan setiap item tes yang yang dijawab benar diberikan skor 1 dan bila salah diberi skor 0. Datanya tertera pada tabel berikut:

Tabel V.5. Data Hasil Belajar Matematika

No	Nama	Butir soal / item										Skor Total
		1	2	3	4	5	6	7	8	9	10	
1	Ahmad	1	0	1	0	1	1	1	1	1	1	8
2	Bakri	0	0	1	0	1	0	0	1	1	1	5
3	Cici	0	1	0	0	0	1	0	1	0	1	4
4	Dhani	1	1	0	0	1	1	0	0	1	0	5
5	Eko	1	1	1	1	1	1	0	0	0	0	6
6	Fatur	1	0	1	0	1	0	1	0	0	0	4
7	Gugun	1	1	1	1	1	1	1	0	0	0	7
8	Hamid	0	1	0	1	1	1	1	1	1	1	8

Penghitungan validitas butir tes nomor 1:  
Langkah-langkah penyelesaian :  
1. Tabel persiapan menghitung validitas item sebagai berikut ;

Tabel V.6. Persiapan Menghitung Validitas Butir

No	Nama	Skor i	Skor Total
1	Ahmad	1	8
2	Bakri	0	5
3	Cici	0	4
4	Dhani	1	5
5	Eko	1	6
6	Fatur	1	4
7	Gugun	1	7
8	Hamid	0	8
	Jumlah	5	47

2. Menghtung harga  $\overline{X}_i$

$$\begin{aligned}\overline{X}_i &= \frac{8+5+6+4+7}{5} \\ &= \frac{30}{5} \\ &= 6\end{aligned}$$

3. Menghitung harga  $\bar{X}_t$

$$\begin{aligned}\bar{X}_t &= \frac{8 + 5 + 4 + 5 + 6 + 4 + 7 + 8}{8} \\ &= \frac{47}{8} \\ &= 5,87\end{aligned}$$

4. Menghitung harga  $S_t$  (standar deviasi total)

Berdasarkan data pada tabel di atas maka dapat dibuat tabel baru untuk dipergunakan dalam mencari harga standar deviasi.

**Tabel V.7. Tabel Penolong**

X	X <sup>2</sup>
8	64
5	25
4	16
5	25
6	36
4	16
7	49
8	64
$\sum X = 47$	$\sum X^2 = 295$

Dari data di atas maka dapat dihitung:

$$\begin{aligned}S_t &= \sqrt{\left(\frac{\sum X^2}{N}\right) - \left(\frac{\sum X}{N}\right)^2} \\ &= \sqrt{\left(\frac{295}{8}\right) - \left(\frac{47}{8}\right)^2} \\ &= \sqrt{36,87 - 34,51} \\ &= \sqrt{2,36} \\ &= 1,53\end{aligned}$$



5. Menghitung harga  $p_i$  sebagai berikut:

$$\begin{aligned} p_i &= 5/8 \\ &= 0,625 \end{aligned}$$

6. Mengitung harga  $q_i$  sebagai berikut:

$$\begin{aligned} q_i &= 1 - 0,625 \\ &= 0,375 \end{aligned}$$

Sehingga diperoleh:

$$\begin{aligned} r_{bis(i)} &= \frac{6 - 5,87}{1,53} \sqrt{\frac{0,625}{0,375}} \\ &= 0,08 \times 1,29 \\ &= 0,10 \end{aligned}$$

Selanjutnya untuk menerima apakah butir tes yang dicari tersebut valid atau invalid, maka harga yang diperoleh tersebut dibandingkan dengan harga kritik yang terdapat dalam tabel statistik *Product Moment* dengan  $N = 8$  maka d.b =  $N - 1 = 8 - 1 = 7$  pada taraf signifikansi 5% yaitu 0,666. Oleh karena nilai koefisien hitung (0,10) lebih kecil dari harga hitung (0,666) maka butir tes nomor 1 tersebut tidak valid.

#### b. Pengujian Validitas Tes Berbentuk Essay

Tes berbentuk essay seperti uraian (*essay*), isian (*fill in*) merupakan tes dengan skor butir berbentuk kontinum. Jika skor butir kontinum maka untuk menguji validitas butir tes dilakukan dengan menghitung koefisien korelasi *Product Moment* yaitu penghitungan koefisien korelasi antara skor butir dengan skor total instrumen dengan menggunakan rumus:

$$r = \frac{N (\sum XY) - (\sum X) (\sum Y)}{\sqrt{[N (\sum X^2) - (\sum X)^2][N (\sum Y^2) - (\sum Y)^2]}}$$

Keterangan:

$N$  = jumlah responden

$X$  = skor butir

$Y$  = skot total

Contoh penggunaannya ;

Guru memberikan skor kepada anak didiknya dengan ketentuan setiap item tes *essay* yang yang dijawab benar dan sempurna diberi skor 3, benar namun kurang sempurna diberi skor 2 dan salah diberikan skor 1. Datanya tertera pada tabel berikut:

**Tabel V. 8. Hasil Belajar Fisika**

No	Nama	Butir soal / item										Skor Total
		1	2	3	4	5	6	7	8	9	10	
1	Ahmad	3	3	3	2	1	1	1	1	1	1	17
2	Bakri	2	3	1	3	3	1	1	2	3	3	22
3	Cici	3	1	3	3	3	1	2	1	2	1	20
4	Dhani	1	1	2	2	3	2	1	2	1	2	17
5	Eko	1	3	1	1	1	1	3	3	1	2	17
6	Fatur	1	2	1	2	1	3	1	2	2	3	18
7	Gugun	1	1	1	1	1	1	1	2	3	1	13
8	Hamid	3	1	2	1	2	1	2	1	3	1	17

Penghitungan validitas butir tes nomor 1:

Langkah-langkah penyelesaian :

1. Tabel persiapan menghitung validitas item sebagai berikut:

**Tabel V.9. Persiapan Menghitung Validitas Butir Tes**

No	X	Y	X <sup>2</sup>	Y <sup>2</sup>	XY
1	3	17	9	289	51
2	2	22	4	484	44
3	3	20	9	400	60
4	1	17	1	289	17
5	1	17	1	289	17
6	1	18	1	324	18
7	1	13	1	169	13
8	3	17	9	289	51
	$\sum X = 15$	$\sum Y = 141$	$\sum X^2 = 35$	$\sum Y^2 = 2533$	$\sum XY = 271$

## 2. Melakukan penghitungan:

$$\begin{aligned}
 r &= \frac{N (\sum XY) - (\sum X) (\sum Y)}{\sqrt{[N (\sum X^2) - (\sum X)^2][N (\sum Y^2) - (\sum Y)^2]}} \\
 &= \frac{8 (271) - (15) (141)}{\sqrt{[8 (36) - (15)^2][8 (2533) - (141)^2]}} \\
 &= \frac{2168 - 2115}{\sqrt{[288 - 225][20264 - 19881]}} \\
 &= \frac{53}{\sqrt{[63][383]}} \\
 &= \frac{53}{\sqrt{24129}} \\
 &= \frac{53}{155,33} \\
 &= 0,341
 \end{aligned}$$

3. Selanjutnya untuk menerima apakah butir tes yang dicari tersebut valid atau invalid, maka harga yang diperoleh tersebut dibandingkan dengan harga kritik yang terdapat dalam tabel statistik *Product Moment* dengan  $N = 8$  maka  $d.b = N - 1 = 8 - 1 = 7$  pada taraf signifikansi 5% yaitu 0,666. Oleh karena nilai koefisien hitung (0,341) lebih kecil dari harga hitung (0,666) maka butir tes nomor 1 tersebut tidak valid.

## c. Pengujian Validitas instrumen Berbentuk Kuesioner/Angket

Pengujian validitas instrumen berbentuk kuesioner atau angket menggunakan *Product Moment* yaitu penghitungan koefisien korelasi antara skor butir kuesioner dengan skor total instrumen dengan menggunakan rumus:

$$r = \frac{N (\sum XY) - (\sum X) (\sum Y)}{\sqrt{[N (\sum X^2) - (\sum X)^2][N (\sum Y^2) - (\sum Y)^2]}}$$

Keterangan:

N = jumlah responden

X = skor butir

Y = skot total

Contoh:

Peneliti memberikan skor angket/kuesioner yang terdiri dari 4 (empat) option. Untuk pernyataan positif yaitu Sangat Setuju (SS) = 4, Setuju (S)= 3, Kurang Setuju (KS)= 2, Tidak Setuju (TS)= 1. Sedangkan untuk pernyataan negatif yaitu Sangat Setuju (SS)= 1, Setuju (S)= 2, Kurang Setuju (KS)= 3, Tidak Setuju (TS)= 4. Datanya tertera pada tabel berikut:

**Tabel V.10. Skor Angket**

No	Nama	Butir Angket										Skor Total
		1	2	3	4	5	6	7	8	9	10	
1	Ahmad	3	3	3	2	4	4	4	3	2	2	30
2	Bakri	2	3	2	3	3	2	3	2	3	3	26
3	Cici	4	3	3	3	3	2	2	3	2	2	27
4	Dhani	3	2	2	2	3	2	2	2	4	2	24
5	Eko	4	3	4	3	2	3	3	3	3	2	30
6	Fatur	3	2	2	2	3	3	3	2	2	3	25
7	Gugun	4	2	2	2	3	2	2	2	3	2	24
8	Hamid	3	4	2	3	2	4	2	3	3	2	28

Penghitungan validitas butir angket nomor 1:

Langkah-langkah penyelesaian :

1. Tabel persiapan menghitung validitas butir angket sebagai berikut;

**Tabel V.11. Tabel Persiapan Menghitung Validitas Butir Angket**

No	X	Y	X <sup>2</sup>	Y <sup>2</sup>	XY
1	3	30	9	900	90
2	2	26	4	676	52
3	4	27	16	729	108
4	3	24	9	576	72
5	4	30	16	900	120
6	3	25	9	625	75
7	4	24	16	576	96
8	3	28	9	784	84
	$\sum X = 26$	$\sum Y = 214$	$\sum X^2 = 88$	$\sum Y^2 = 5766$	$\sum XY = 697$

## 2. Melakukan penghitungan:

$$\begin{aligned}
 r &= \frac{N (\sum XY) - (\sum X) (\sum Y)}{\sqrt{[N (\sum X^2) - (\sum X)^2][N (\sum Y^2) - (\sum Y)^2]}} \\
 &= \frac{8 (697) - (26) (214)}{\sqrt{[8 (88) - (26)^2][8 (5766) - (214)^2]}} \\
 &= \frac{5576 - 5564}{\sqrt{[704 - 676][46128 - 45796]}} \\
 &= \frac{12}{\sqrt{[28][332]}} \\
 &= \frac{12}{\sqrt{9296}} \\
 &= \frac{12}{9641} \\
 &= 0,124
 \end{aligned}$$

3. Selanjutnya untuk menerima apakah butir angket nomor 1 yang dicari tersebut valid atau invalid, maka harga yang diperoleh tersebut dibandingkan dengan harga kritik yang terdapat dalam tabel statistik *Product Moment* dengan  $N = 8$  maka  $d.b = N - 1 = 8 - 1 = 7$  pada taraf signifikansi 5% yaitu 0,666. Oleh karena nilai koefisien hitung (0,124) lebih kecil dari harga hitung (0,666) maka butir tes nomor 1 tersebut tidak valid.

## 2. Reliabilitas Instrumen

Reliabilitas merupakan penerjemahan dari kata *reliability* yang mempunyai asal kata *rely* dan *ability*. Pengukuran yang memiliki reliabilitas tinggi disebut sebagai pengukuran yang reliabel. Reliabilitas memiliki istilah atau nama lain seperti keterpercayaan, keterhandalan, keajegan, kestabilan, konsistensi.

Berdasarkan arti kata tersebut, maka instrumen yang reliabel adalah instrumen yang hasil pengukurannya dapat dipercaya. Salah satu kriteria instrumen yang dapat dipercaya jika instrumen tersebut digunakan secara berulang-ulang, hasil pengukurannya tetap. Mistar dapat dipercaya sebagai alat ukur, karena berdasarkan pengalaman jika mistar digunakan dua kali

atau lebih mengukur panjang sebuah benda, maka hasil pengukuran pertama dan selanjutnya terbukti tidak berbeda. Sebuah tes dapat dikatakan reliabel jika tes tersebut digunakan secara berulang terhadap peserta didik yang sama hasil pengukurannya relatif tetap sama.

#### a. Reliabilitas untuk Instrumen yang Berbentuk Dikotomi.

Reliabilitas untuk instrumen yang berbentuk dikotomi yaitu instrumen dengan pemberian skor 0 dan 1 maka pengujiannya dapat dilakukan dengan menggunakan rumus Kuder Richardson 20 (KR. 20) dan Kuder Richardson 21 (KR. 21).

Penggunaan rumus KR. 20 digunakan apabila alternatif jawaban pada instrumen bersifat dikotomi, misalnya benar-salah dan pemberian skor = 1 dan 0. Rumus KR. 20 adalah:

$$r_{kk} = \frac{k}{k-1} \left[ 1 - \frac{\sum p q}{S_t^2} \right]$$

Keterangan:

$r_{kk}$  = koefisien reliabilitas

$k$  = banyaknya butir

$p$  = proporsi jawaban benar

$q$  = proporsi jawaban salah

$S_t^2$  = varians skor total

Contoh penggunaan rumus KR 20.

1. Langkah pertama tes hasil uji coba diberi skor-skor, kemudian didistribusikan ke dalam tabel kerja sebagai berikut:

**Tabel V.12. Skor Tes**

Nama	Butir soal / item										Skor Total
	1	2	3	4	5	6	7	8	9	10	
Ani	1	1	1	1	1	1	1	1	1	1	10
Badu	1	1	1	0	1	0	0	1	1	1	7
Caca	1	1	1	1	1	1	1	1	1	0	9
Danu	1	1	0	0	1	1	0	0	1	0	5
Eka	1	1	1	1	1	1	0	0	0	0	6

Fatur	1	0	1	1	1	1	1	1	0	1	9
Gogon	1	1	1	1	1	1	1	0	0	0	7
Hamid	0	1	0	1	1	1	1	1	1	1	8
p q p q	7	7	6	6	8	7	5	5	5	4	61
	0,875	0,875	0,750	0,750	1,000	0,875	0,625	0,625	0,625	0,500	
	0,125	0,125	0,250	0,250	0,000	0,125	0,375	0,375	0,375	0,500	
	0,109	0,109	0,188	0,188	0,000	0,109	0,234	0,234	0,234	0,250	

Dari tabel di atas diketahui:

k = 10

p.q = 1,656

2. Menghitung harga varians total sebagai berikut:

$$S_t^2 = \frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N}$$

Untuk menyelesaikan rumus di atas dibutuhkan harga  $X^2$  dan  $(X)^2$  yang diperoleh dari tabel berikut:

**Tabel V.13. Tabel Penolong**

No	X	X <sup>2</sup>
1	10	100
2	7	49
3	9	81
4	5	25
5	6	36
6	9	81
7	7	49
8	8	64
	Σ X = 61	Σ X <sup>2</sup> = 485

Dari data di atas dapat dihitung varians total sebagai berikut:

$$\begin{aligned}
 S^2 &= \frac{485 - \frac{(61)^2}{8}}{8} \\
 &= \frac{485 - 465,12}{8} \\
 &= \frac{19,88}{8} \\
 &= 2,48
 \end{aligned}$$

3. Menghitung koefisien reliabilitas sebagai berikut:

$$\begin{aligned}
 r_{kk} &= \frac{10}{10-1} \left[ 1 - \frac{1,656}{2,48} \right] \\
 &= \frac{10}{9} [1 - 0,66] \\
 &= 11 [0,34] \\
 &= 0,38
 \end{aligned}$$

4. Langkah terakhir adalah menentukan kriteria reliabilitas tes. Merujuk kepada Sudijono (2002) suatu instrumen dikatakan memiliki nilai reliabel apabila koefisien reliabilitas adalah e" 0,70. Oleh karena diperoleh harga koefisien reliabilitas 0,38 lebih kecil dari ketentuan tersebut maka dapat disimpulkan bahwa instrumen tidak reliabel.

### **b. Reliabilitas untuk Instrumen yang Berbentuk Kontinum.**

Reliabilitas untuk instrumen yang berbentuk kontinum yaitu instrumen dengan pemberian skor yang skornya merupakan rentangan 0 - 10, 0 - 100 atau berbentuk skala 1 - 3, 1 - 5 atau 1 - 10, maka pengujiannya dapat dilakukan dengan menggunakan rumus Alpha Cronbach yaitu:

$$r_{kk} = \left[ \frac{k}{k-1} \right] \left[ 1 - \frac{\sum S_b^2}{S_t^2} \right]$$

Keterangan:

$r_{kk}$  = reliabilitas instrumen

$k$  = jumlah butir angket

$\sum S_b^2$  = jumlah varians butir

$S_t^2$  = varians total



Contoh:  
Terdapat data sebagaimana tersaji sebagai berikut:

**Tabel V.14. Data Hasil Tes**

Nomor	Butir Soal					Skor Total	Kuadrat Skor Total
	1	2	3	4	5		
1	10	8	9	10	8	45	2025
2	8	7	8	9	7	39	1521
3	6	5	6	8	7	32	1024
4	5	4	3	0	2	14	196
5	9	10	8	7	6	40	1600
6	7	5	3	4	7	26	676
7	3	4	4	5	6	22	484
8	4	3	5	5	5	22	484
9	6	2	2	2	3	15	225
10	7	6	1	5	4	23	529
Jumlah	65	54	49	55	55	278	8764
Jumlah Kuadrat	465	344	309	389	337	1844	

1. Menentukan varians masing-masing butir:

$$\begin{aligned}\sigma^2(1) &= \frac{465 - \frac{(65)^2}{10}}{10} \\ &= \frac{465 - 422,5}{10} \\ &= 4,25\end{aligned}$$

$$\begin{aligned}\sigma^2(2) &= \frac{344 - \frac{(54)^2}{10}}{10} \\ &= \frac{344 - 291,6}{10} \\ &= 5,24\end{aligned}$$

$$\begin{aligned}\sigma^2(3) &= \frac{309 - \frac{(49)^2}{10}}{10} \\ &= \frac{309 - 240,1}{10} \\ &= 6,89\end{aligned}$$

$$\begin{aligned}
 \sigma^2(4) &= \frac{389 - \frac{(55)^2}{10}}{10} \\
 &= \frac{389 - 302,5}{10} \\
 &= 8,65
 \end{aligned}$$

$$\begin{aligned}
 \sigma^2(5) &= \frac{337 - \frac{(55)^2}{10}}{10} \\
 &= \frac{337 - 302,5}{10} \\
 &= 3,50
 \end{aligned}$$

Dengan demikian diperoleh total varian butir adalah:

$$\begin{aligned}
 \sigma^2 &= 4,25 + 5,24 + 6,89 + 8,65 + 3,50 \\
 &= 28,8
 \end{aligned}$$

2. Varians total dihitung sebagai berikut:

$$\begin{aligned}
 \sigma^2(t) &= \frac{8764 - \frac{(278)^2}{10}}{10} \\
 &= \frac{8764 - 7728,4}{10} \\
 &= 103,56
 \end{aligned}$$

3. Menghitung koefisien reliabilitas sebagai berikut:

$$\begin{aligned}
 r_{kk} &= \left[ \frac{5}{5-1} \right] \left[ 1 - \frac{28,8}{103,56} \right] \\
 &= \left[ \frac{5}{4} \right] [1 - 0,27] \\
 &= 1,25 \times 0,73 \\
 &= 0,91
 \end{aligned}$$

4. Langkah terakhir adalah menentukan kriteria reliabilitas tes.  
Merujuk kepada Sudijono (2002) suatu instrumen dikatakan memiliki nilai reliabel apabila koefisien reliabilitas adalah  $\geq 0,70$ . Oleh karena diperoleh harga koefisien reliabilitas 0,91 lebih kecil dari ketentuan tersebut maka dapat disimpulkan bahwa instrumen tidak reliabel.

## BAB VI

### ANALISIS DAN INTERPRESTASI DATA

#### A. Jenis Data

**K**ata “data” berasal dari bahasa Latin yang berarti keterangan atau kumpulan keterangan. Data adalah kata dalam bentuk jamak, sedangkan dalam bentuk tunggal adalah *datum*. Data yang merupakan bahan mentah yang perlu diolah sehingga menghasilkan informasi atau keterangan baik kualitatif maupun kuantitatif disebut data mentah.

Data merupakan keterangan-keterangan tentang sesuatu hal, dapat berupa sesuatu yang diketahui atau dianggap. Jadi data dapat diartikan sebagai sesuatu yang diketahui atau yang dianggap (anggapan). Sesuatu yang diketahui biasanya didapat dari hasil pengamatan atau percobaan dan hal itu berkaitan dengan waktu dan tempat. Anggapan atau asumsi merupakan suatu perkiraan atau dugaan yang sifatnya masih sementara, sehingga belum tentu benar. Oleh karena itu, anggapan atau asumsi perlu diuji kebenarannya.

Atas dasar sifat dan bentuknya maka data dapat dibedakan atas dua jenis sebagaimana dijelaskan Riduwan (2002), yaitu:

1. Data kuantitatif.

Data kuantitatif adalah data dalam bentuk angka sebagai hasil pengamatan atau pengukuran yang dapat dihitung dan diukur. Misalnya data tentang berat badan, harga barang-barang, yang dapat diukur dan dinyatakan dalam bentuk angka. Contohnya: Fatur beratnya 30 kg, Annisa tingginya 120 cm, sepatu itu harganya Rp. 75.000, Salsa dapat menyelesaikan tugas itu dalam waktu 1 jam.

2. Data kualitatif.

Data kualitatif yaitu data yang berhubungan dengan kategorisasi, karakteristik yang berwujud pernyataan atau berupa kata-kata. Misalnya baik-buruk, senang-sedih, harga minyak naik, rumah itu besar, pohon itu rindang, laut itu dalam dan sebagainya.

Menurut waktu pengumpulannya, data dapat dikelompokkan atas dua jenis yaitu:

1. Data berkala (*time series*).

Data berkala adalah data yang terkumpul dari waktu ke waktu untuk memberikan gambaran perkembangan suatu kegiatan. Misalnya, data perkembangan harga bahan pokok selama 10 bulan terakhir yang dikumpulkan setiap bulan.

2. Data *cross section*.

Data *cross section* adalah data yang terkumpul pada suatu waktu tertentu untuk memberikan gambaran perkembangan keadaan atau kegiatan pada waktu itu. Misalnya hasil ujian semester 1, data sensus penduduk tahun 2010.

Menurut sumber pengambilannya, data dapat dikelompokkan atas dua jenis yaitu:

1. Data primer.

Data primer adalah data yang diperoleh atau dikumpulkan oleh orang yang melakukan penelitian atau yang bersangkutan memerlukannya. Data primer disebut juga data asli atau data baru.

2. Data skunder.

Data skunder adalah data yang diperoleh atau dikumpulkan dari sumber-sumber yang telah ada. Data itu biasanya diperoleh dari perpustakaan atau laporan-laporan, dokumen peneliti yang terdahulu. Data skunder disebut juga data tersedia.

Menurut susunannya, data dapat dikelompokkan atas dua jenis sebagai berikut:

1. Data acak atau data tunggal.

Data acak atau tunggal adalah data yang belum tersusun atau dikelompokkan ke dalam kelas-kelas interval.

## 2. Data kelompok.

Data kelompok adalah data yang sudah tersusun atau dikelompokkan ke dalam kelas-kelas interval. Data kelompok disusun dalam bentuk distribusi frekuensi atau tabel frekuensi.

Data kelompok dibedakan atas:

### a) Data kelompok diskrit.

Data kelompok diskrit adalah data yang diperoleh dari hasil menghitung, seperti jumlah penduduk, jumlah siswa laki-laki dan siswa perempuan.

### b) Data kelompok kontiniu

Data kelompok kontiniu adalah data yang diperoleh dari hasil mengukur, seperti berat badan, tinggi badan, hasil belajar, motivasi belajar, kecerdasan inteligensi dan sebagainya.

Arikunto dan Jabar (2009:130) memaparkan data mentah yang diperoleh dari lapangan akan bervariasi, tergantung pada alat pengumpul data yang digunakan oleh evaluator sebagai berikut:

1. Data yang diperoleh dengan menggunakan angket/kuesioner maka data yang diperoleh berupa centangan atau tanda *check list* (✓) pada pilihan-pilihan, lingkaran-lingkaran pada angka atau huruf yang disediakan dalam instrumen, atau kalimat-kalimat jawaban yang sifatnya kualitatif.
2. Data yang diperoleh dengan wawancara, wujud data yang diperoleh berbentuk centangan, lingkaran dan kalimat jawaban yang diberikan oleh responden dan dicatat oleh petugas pengumpul data.
3. Data yang diperoleh dengan observasi maka wujud data yang diperoleh berbentuk centangan atau *check list*, lingkaran, dan kalimat-kalimat catatan petugas pengumpul data.
4. Data yang diperoleh dengan menggunakan dokumentasi berupa angka-angka atau simbol-simbol yang menunjuk peringkat kondisi objek yang ditelaah.
5. Data yang diperoleh dengan tes atau inventori berupa angka-angka yang menunjukkan skor nilai.

Berdasarkan kelima bentuk data sebagaimana dijelaskan di atas, maka Arikunto dan Jabar (2009:130) mengelompokkan jenis data menjadi 3 (tiga) kelompok yaitu:

1. Nilai jadi, berupa nilai angka yang dibuat dari interpretasi kriteria dan tes.
2. Kode-kode atau simbol-simbol yang bisa berupa tanda centang dan lingkaran atau memberikan tanda silang pada pilihan-pilihan.
3. Informasi dalam bentuk paparan kalimat yang memuat data kuantitatif dan kualitatif.

Dalam evaluasi program maupun penelitian lainnya, data secara umum dapat dibedakan menjadi data kualitatif dan data kuantitatif (Purwanto dan Suparman, 1999:182). Berikut penjelasannya:

1. Data kualitatif adalah data yang sifatnya tidak numerik. Data ini biasanya dikumpulkan untuk menjaring informasi yang tidak dapat ditangkap secara kuantitatif. Contoh data kualitatif adalah data tentang alasan peserta pelatihan tidak termotivasi untuk aktif mengikuti pelatihan.
2. Data kuantitatif adalah data yang berbentuk angka-angka (numerik). Contohnya data usia dan data hasil tes peserta pelatihan. Data kuantitatif menurut sifatnya dapat dibedakan atas beberapa jenis data sebagai berikut:
  - a. Data kontiniu adalah data yang secara teoritis dapat memiliki nilai pada tiap titik kemungkinan pada suatu kontinum. Contoh, nilai tes akhir pelatihan yang menggunakan rentang nilai 0 sampai dengan 100. Data nilai tersebut bersifat kontiniu karena dimungkinkan seseorang peserta akan memperoleh nilai antara 0 sampai 100 tersebut dalam bentuk angka ganjil atau genap, pecahan atau utuh, misalnya 40, 45 maupun 5,5 dan seterusnya.
  - b. Data jenjang atau *rank* adalah data yang melukiskan posisi seseorang atau sesuatu objek dibandingkan dengan keseluruhan keadaan. Contoh, data tentang peringkat prestasi hasil tes akhir peserta pelatihan tersusun mulai dari peringkat pertama, peringkat kedua dan seterusnya.
  - c. Data dikotomi adalah data yang hanya mempunyai dua macam nilai. Data dikotomi dibedakan menjadi dua macam:
    - 1) Benar-benar dikotomi (*true dichotomy*) yaitu data yang benar-benar dikotomi contohnya data jenis kelamin yaitu laki-laki dan perempuan, data kelulusan yaitu lulus dan tidak lulus.
    - 2) Dikotomi buatan (*artificial*) contohnya pelatihan peserta dikelompokkan menjadi peserta senior dan junior.

- d. Data kategori adalah data yang bisa memiliki lebih dari dua kutub yang bukan dikotomi dan juga bukan kontiniu. Contoh data mengenai tingkat partisipasi peserta pelatihan dalam kegiatan diskusi yaitu sangat aktif, aktif, kurang aktif, pasif.

## B. Prinsip dan Kriteria Menganalisis Data

Beberapa prinsip dan kriteria dalam menganalisis data yang harus diketahui seorang evaluator. Prinsip dan kriteria tersebut dijelaskan oleh Purwanto dan Suparman (1999:187) sebagai berikut:

1. Jangan terlalu menyederhanakan.  
Pertanyaan evaluasi selalu berkaitan dengan gejala yang bersifat dinamis dan kompleks. Prosedur analisis harus peka terhadap hal-hal yang kompleks tersebut, dan janganlah mengurangi kompleksitas tersebut untuk memudahkan dan menyederhanakannya.
2. Perhatikan perbedaan-perbedaan tentang kondisi dan yang mempengaruhinya.
3. Gunakan teknik yang beragam. Jika memungkinkan gunakanlah analisis yang beragam.
4. Yakinkan diri bahwa asumsi-asumsi untuk mempergunakan teknik analisis tersebut telah dipenuhi dan sesuai dengan data yang diperoleh.
5. Gunakanlah metode yang cocok dengan audien dan tujuan.
6. Gunakan metode praktis dan terjangkau biayanya.
7. Upayakan analisis yang sesederhana mungkin.
8. Janganlah menuntut yang serba sempurna, setiap teknik analisis memiliki keterbatasan.

## C. Mengorganisasi Data

Menurut Purwanto dan Suparman (1999:184) ada tiga kegiatan utama dalam pengorganisasian data yaitu:

### 1. Pengkodean atau Pembuatan Daftar Kode Data (*Coding*).

Setiap variabel dan indikator yang akan dianalisis dibuatkan kodenya dalam bentuk angka dan menentukan jumlah kolomnya. Berhubung pilihan-pilihan dalam instrumen terkadang tidak seragam ada yang berwujud nominal, interval ataupun ordinal maka harus menyeragamkannya dalam



bentuk yang sama agar dapat diolah dengan mudah. Artinya harus dilakukan manipulasi atas variabel dan indikator-indikator tersebut dengan menggunakan kodel/label.

Tujuan dari pemberian atau pengkodean data adalah untuk memudahkan pengorganisasian data, terutama apabila data tersebut akan diolah menggunakan komputer.

Contoh: daftar kode atau petunjuk pengkodean sebagai berikut:

No	Variabel/Indikator	Kode	Kolom	Keterangan
A	Identitas peserta			
1	Nama	01-99		
2	Jenis kelamin :		1	
	L	1		
	P	2		
3	Jenis pekerjaan:		2	
	PNS	1		
	Pegawai Swasta	2		
	Buruh	3		
B	Hasil tes			
1	Pretes	000-100	3-5	Rentangan 0-100
2	Postes	000-100	6-8	

2. Membuat Matriks Tabulasi.

Tabulasi dapat diartikan menyusun menjadi tabel, pengertian lain tabulasi adalah pengolahan atau pemrosesan menjadi tabel (Arikunto dan Jabar, 2009:129). Tabulasi merupakan *coding sheet* yang memudahkan evaluator dalam mengolah dan menganalisisnya, baik secara manual maupun dengan menggunakan komputer. Tabulasi berisikan variabel dan indikator-indikator yang akan diteliti dan angka-angka sebagai simbolisasi (label) dari kategori berdasarkan variabel dan indikator-indikator tersebut.

Setelah menentukan kode serta jumlah kebutuhan kolom untuk pengorganisasian data, maka selanjutnya membuat matriks tabulasinya. Matriks tabulasi data kuantitatif berbeda dengan untuk data kualitatif.

Contoh: matriks tabulasi data kuantitatif.

No Kode Peserta	Identitas		Skor Pretes			Skor Postes		
	1	2	3	4	5	6	7	8

Contoh: matriks tabulasi data kualitatif.

Subjek/Peserta	Komentar

### 3. Memasukkan Data yang Telah Dikode ke Dalam Matriks Tabulasi.

Setelah selesai dengan matriks tabulasi, kemudian dilanjutkan dengan memasukkan data yang telah dikumpulkan ke dalam matriks tabulasi. Data yang dimasukkan ke dalam matriks adalah kode datanya atau narasinya.

Contoh: hasil pengisian matriks tabulasi untuk data kuantitatif

No Kode Peserta	Identitas		Skor Pretes			Skor Postes		
	1	2	3	4	5	6	7	8
01	1	2	0	5	9	0	6	7
02	1	2	0	7	0	0	7	1
03	2	1	0	5	8	0	5	4

Contoh: hasil pengisian matriks tabulasi untuk data kualitatif

No Peserta	Komentar
01	Instruktur kurang tanggap terhadap pertanyaan peserta. Instruktur cukup menguasai materi.
02	Instruktur mampu menjelaskan dengan baik, tetapi kurang memberikan contoh.

**D. Prosedur Analisis Data**

Prosedur analisis data dalam evaluasi program dikelompokkan menjadi dua yaitu prosedur kuantitatif dan prosedur kualitatif (Purwanto dan Suparman, 1999:188). Prosedur kuantitatif untuk menganalisis data yang berupa angka-angka, sedangkan prosedur kualitatif untuk menganalisis data yang bersifat naratif. Berikut penjelasannya:

**1. Prosedur Kuantitatif**

Setidaknya terdapat dua jenis prosedur analisis kuantitatif yang kerap kali digunakan dalam evaluasi program yaitu statistik deskriptif dan statistik inferensial. Statistik deskriptif (*descriptive statistics*) yaitu statistik yang mempelajari tata cara mengumpulkan, menyusun, menyajikan dan menganalisa data penelitian yang berwujud angka-angka, agar dapat memberikan gambaran yang teratur, ringkas dan jelas mengenai suatu gejala, keadaan peristiwa, sehingga dapat ditarik pengertian atau makna tertentu. Analisis data yang tergolong statistik deskriptif, terdiri dari tabel, grafik, mean, median, modus, pengukuran variasi data, dan teknik statistik lain yang bertujuan hanya mengetahui gambaran atau kecenderungan data tanpa bermaksud melakukan generalisasi.

Sugiyono (2004) menjelaskan statistik deskriptif adalah statistik yang digunakan untuk menggambarkan atau menganalisa suatu statistik hasil penelitian, tetapi tidak digunakan untuk membuat kesimpulan yang lebih luas (generalisasi/ inferensi). Lebih lanjut dijelaskan Sugiyono bahwa penelitian yang tidak menggunakan sampel, maka analisisnya akan menggunakan statistik deskriptif. Demikian juga dengan penelitian yang menggunakan sampel tetapi peneliti tidak bermaksud untuk membuat kesimpulan untuk populasi dari mana sampel diambil, maka statistik yang digunakan adalah statistik deskriptif.

Ruang lingkup kajian pada analisis statistik deskriptif dijelaskan Djarwanto dan Subagyo (1998) yaitu:

- a. Distribusi frekuensi serta pengukuran nilai-nilai statistiknya seperti pengukuran nilai sentral, dispersi, skewness dan kurtosis, dan grafiknya seperti poligon, histogram dan ogive.
- b. Angka indeks.
- c. Time series atau deret waktu.
- d. Koefisien regresi dan koefisien korelasi sederhana.

Hal senada dijelaskan Supardi (2013) mengenai ruang lingkup kajian statistik deskriptif yaitu:

- a. Penyajian data dalam bentuk tabel seperti tabel tunggal, tabel kontigensi maupun tabel distribusi frekuensi.
- b. Penyajian data dalam bentuk grafik seperti diagram batang, diagram garis, diagram lingkaran, diagram pencar, diagram peta, diagram simbol maupun diagram yang disajikan dari tabel distribusi frekuensi yaitu histogram, poligon frekuensi dan ogive.
- c. Ukuran nilai pusat dan letak, seperti rerata, median, modus, varian, simpangan baku, kuartil, desil, persentil.
- d. Ukuran dispersi atau simpangan seperti jangkauan atau rentang, rerata simpangan, variansi, simpangan baku.
- e. Model distribusi data yaitu kemencengan dan keruncingan kurva distribusi.
- f. Angka indeks.
- g. Times series/deret waktu atau data berkala.

Statistik inferensial (*inferensial statistics*) yaitu statistik yang mempelajari atau mempersiapkan tata cara penarikan kesimpulan mengenai karakteristik populasi, berdasarkan data kuantitatif yang diperoleh dari sampel penelitian. Penarikan kesimpulan mengenai karakteristik populasi berdasarkan data sampel yang diambil dari populasinya disebut *generalisasi* atau *induksi*. Karena itu statistik inferensial juga dikenal sebagai induktif (*inductive statistics*). Di samping fungsi generalisasi, statistik inferensial juga menyediakan aturan-aturan tertentu dalam rangka penyusunan atau pembuatan ramalan (*prediction*) maupun penaksiran (*estimation*).

Ruang lingkup kajian pada analisis statistik inferensial dijelaskan Djarwanto dan Subagyo (1998) yaitu:

- a. Probabilitas.
- b. Distribusi teoritis.
- c. Sampling dan distribusi sampling.
- d. Estimasi harga parameter.
- e. Uji hipotesis, termasuk uji chi-square dan analisis varians.
- f. Analisis regresi untuk prediksi.
- g. Korelasi dan uji signifikansi.
- h. Koefisien regresi dan koefisien korelasi sederhana.

Selanjutnya Supardi (2013) menjelaskan ruang lingkup kajian statistik inferensial sebagai berikut:

- a. Uji persyaratan analisis (uji pelanggaran klasik) seperti uji normalitas, uji homogenitas, uji kelinieran, uji multikolinieritas.
- b. Uji hipotesis asosiasi seperti uji korelasi, uji regresi, uji analisis jalur dan uji kanonikal.
- c. Uji hipotesis komparasi, seperti uji t, uji beda 2 (dua) kelompok data, uji Tuckey, analisis varian, analisis kovarian, multivarian analisis varians dan multivariat analisis kovarians.

Berikut contoh penggunaan statistik deskriptif:

- Deskripsi program.  
Program ini bertujuan untuk meningkatkan kemampuan guru dalam berkomunikasi interpersonal. Untuk mencapai tujuan ini pelatihan maka diberikan materi secara teori dan praktek, di antaranya dengan cara menonton video yang berisikan penerapan kemampuan berkomunikasi interpersonal. Dalam video dijelaskan cara membedakan antara yang diucapkan dengan apa yang dimaksud. Caranya dengan melihat isi, nada, konteks dan bahasa tubuh. Program ini akan berhasil jika peserta pelatihan, belajar menggunakan isi, nada, konteks dan bahasa tubuh untuk mengintegrasikan pembicaraan.
- Pertanyaan evaluasi.  
Apakah peserta mampu menggunakan isi, nada, konteks dan bahasa tubuh dalam percakapan?

- Rancangan evaluasi.  
Kemampuan menginterpretasikan pembicaraan diukur dengan instrumen tes.
- Rencana pengumpulan data.  
Tes diberikan kepada semua peserta pada akhir program. Skor tes maksimal adalah 100 dan peserta dianggap berhasil jika mencapai skor 70 (kriteria berhasil).
- Rencana analisis.  
Menghitung mean, simpangan baku, rank dan skor peroleh masing-masing peserta di akhir program.
- Analisis data.  
Data hasil tes tergambar pada tabel berikut:

**Tabel VI. 1. Data Skor Tes**

No	Skor Tes
1	50
2	50
3	50
4	55
5	55
6	55
7	60
8	60
9	60
10	70
11	70
12	70
13	70
14	65
15	65
16	65
17	65
18	65
19	75
20	75
21	75

22	75
23	75
24	80
25	80
26	80
27	80
28	85
29	85
30	85
31	85
32	90
33	90
34	90
35	90

Berdasarkan data hasil tes di atas maka dilakukan pengolahan statistik deskriptifnya yaitu untuk mengetahui rata-rata (mean), varians dan simpangan bakunya. Langkah-langkah atau prosedur penghitungan rata-rata (mean) dan varians dan simpangan bakunya dapat dilihat pada tabel kerja sebagai berikut :

**Tabel VI. 2 Tabel Kerja Menghitung Mean, Varians Dan Simpangan Baku**

No	Skor ( $x_i$ )	$f_i$	$f_i x_i$	$x_i^2$	$f_i x_i^2$
1	50	3	150	2500	7500
2	55	3	165	3025	9075
3	60	3	180	3600	10800
4	65	5	325	4225	21125
5	70	4	280	4900	19600
6	75	5	375	5625	28125
7	80	4	320	6400	25600
8	85	4	340	7225	28900
9	90	4	360	8100	32400
		35	2495		183125

Rata-rata (mean) dapat dihitung sebagai berikut ini :

$$\begin{aligned}\bar{X} &= \frac{\sum f_i x_i}{\sum f_i} \\ &= \frac{2495}{35} \\ &= 71,28\end{aligned}$$

Varians dapat dihitung sebagai berikut :

$$\begin{aligned}s^2 &= \frac{n \sum f_i x_i^2 - (\sum f_i x_i)^2}{n(n-1)} \\ &= \frac{35 \times 183125 - (2495)^2}{35(35-1)} \\ &= \frac{6409375 - 6225025}{35(34)} \\ &= \frac{184350}{1190} \\ &= 154,91\end{aligned}$$

Simpangan baku dihitung dengan cara menarik akar dari varians yaitu:  $\sqrt{154,91} = 12,44$

Rangkuman analisis data yaitu mean atau rata-rata diperoleh = 71,28 kemudian simpangan baku 12,44 dengan N (banyak responden) = 35 orang, rank 50-90. Jumlah peserta yang mencapai skor 70 atau lebih sebanyak 26 orang.

- Interpretasi.

Dari keseluruhan peserta yaitu 35 orang, nilai rata-ratanya adalah 71,28 yang berarti jauh lebih tinggi dari kriteria keberhasilan yaitu 70. Sebanyak 26 peserta pelatihan menunjukkan kemampuan menginterpretasikan percakapan dengan baik. Simpangan baku dan rank yang relatif kecil menunjukkan bahwa kemampuan peserta pelatihan relatif sama.

Berikut contoh penggunaan statistik inferensial (dalam hal ini menggunakan uji t).

- Deskripsi program.

Kepala perpustakaan di satu perguruan tinggi negeri mengadakan program



pelatihan bagi operator komputer perpustakaan baik operator komputer di tingkat Fakultas maupun di perpustakaan Induk. Tujuan pelatihan ini adalah untuk meningkatkan pengetahuan dan keterampilan operator dalam mengoperasikan *software* dan *hardware* terbaru.

- **Pertanyaan evaluasi.**  
Apakah peserta mampu mengoperasikan *software* dan *hardware* terbaru tersebut?
- **Rancangan evaluasi.**  
Kemampuan peserta pelatihan untuk mengoperasikan *software* dan *hardware* diukur dengan instrumen tes (dalam hal ini dilakukan *pre-test* dan *post-test*).
- **Rencana pengumpulan data.**  
Tes diberikan kepada semua peserta di awal dan diakhir program pelatihan.
- **Analisis data.**  
Langkah analisis yang dilakukan adalah:

- 1) **Tabulasi hasil tes**  
Dalam hal ini mentabulasi hasil tes sebelum (*pre-test*) yang diberikan diawali program dilakukan dan hasil tes sesudah (*post-test*) yang diberikan diakhir program. Hasilnya sebagai berikut:

**Tabel VI. 3 Data Pre-Test dan Post-Test**

No	Pre-Test	Post-Test
1	40	60
2	48	72
3	44	84
4	48	64
5	52	80
6	52	84
7	60	80
8	36	68
9	50	72
10	32	68
11	40	80
12	44	84
13	36	76

14	56	76
15	28	64
16	36	76
17	40	50
18	34	60
19	44	76
20	28	64
21	24	60
22	34	72
23	36	76
24	24	60
25	32	68
26	32	68
27	48	76
28	28	64
29	34	50
30	40	80
31	40	80
32	56	72

- 2) Melakukan perhitungan statistik dengan menggunakan uji t.

$$t = \frac{Md}{\sqrt{\frac{\sum Xd^2}{n(n-1)}}}$$

Keterangan:

d = selisih skor sesudah dan dan sebelum dari tiap subjek.

Md = rerata dari gain (d)

Xd = deviasi skor gain terhadap reratanya ( $Xd = d - Md$ )

$Xd^2$  = kuadrat deviasi skor gain

n = banyak subjek

Perhitungan untuk menghitung nilai  $t_{hitung}$  terdapat pada tabel sebagai berikut:

Tabel VI. 4 Tabel Kerja Menghitung Harga t

No	Skor		D	Xd	Xd <sup>2</sup>
	Pre-Test	Post-Test			
1	40	60	20	-10,88	118,27
2	48	72	24	-6,88	47,27
3	44	84	40	9,13	83,27
4	48	64	16	-14,88	221,27
5	52	80	28	-2,88	8,27
6	52	84	32	1,13	1,27
7	60	80	20	-10,88	118,27
8	36	68	32	1,13	1,27
9	50	72	22	-8,88	78,77
10	32	68	36	5,13	26,27
11	40	80	40	9,13	83,27
12	44	84	40	9,13	83,27
13	36	76	40	9,13	83,27
14	56	76	20	-10,88	118,27
15	28	64	36	5,13	26,27
16	36	76	40	9,13	83,27
17	40	50	10	-20,88	435,77
18	34	60	26	-4,88	23,77
19	44	76	32	1,13	1,27
20	28	64	36	5,13	26,27
21	24	60	36	5,13	26,27
22	34	72	38	7,13	50,77
23	36	76	40	9,13	83,27
24	24	60	36	5,13	26,27
25	32	68	36	5,13	26,27
26	32	68	36	5,13	26,27
27	48	76	28	-2,88	8,27
28	28	64	36	5,13	26,27
29	34	50	16	-14,88	221,27
30	40	80	40	9,13	83,27
31	40	80	40	9,13	83,27
32	56	72	16	-14,88	221,27
Jumlah			988	-	2551,50

Dari tabel di atas maka dapat dihitung harga  $t$  berikut:

$$\begin{aligned} Md &= \frac{988}{32} \\ &= 30,88 \end{aligned}$$

Perhitungan nilai  $t_{hitung}$  sebagai berikut:

$$\begin{aligned} t &= \frac{30,88}{\sqrt{\frac{2551,50}{32(32-1)}}} \\ &= \frac{30,88}{\sqrt{2,57}} \\ &= 19,3. \end{aligned}$$

- 3) Mengecek nilai  $t_{tabel}$  dan membandingkannya dengan nilai  $t_{hitung}$ . Nilai  $t_{tabel}$  pada  $db = N - 1 = 32 - 1 = 31$  pada  $\alpha = 0,05$  adalah 2,039. Oleh karena harga  $t_{hitung} > t_{tabel}$  maka dapat disimpulkan bahwa terdapat perbedaan yang signifikan antara hasil *pre-test* dan *post-test*.
- 4) Menarik kesimpulan.  
Pelatihan yang dilakukan mencapai tujuannya yaitu meningkatkan pengetahuan dan ketrampilan operator komputer perpustakaan dalam mengoperasikan *software* dan *hardware* terbaru tersebut.

## 2. Prosedur Kualitatif

Prosedur analisis data kualitatif bertujuan pada proses penggalan makna, penggambaran, penjelasan dan penempatan data pada konteksnya masing-masing. Uraian data jenis ini berupa kalimat-kalimat, bukan angka-angka atau tabel-tabel. Untuk itu data yang diperoleh harus diorganisir dalam struktur yang mudah dipahami dan diuraikan, dalam hal ini Faisal (1999:256) menggambarkan prosedur analisis data kualitatif sebagai berikut:

### a. Reduksi Data

Data yang diperoleh dari lapangan direduksi dengan cara memilih dan memilah mana yang sesuai atau satu kelompok dengan kelompok variabel atau penggolongan/kategori yang telah dibuat sebelumnya, yaitu

yang jelas-jelas kategori atau variabel ini harus mengacu pada tujuan evaluasi program yang telah ditentukan. Data dirangkum, dipilih hal-hal pokoknya, difokuskan pada hal-hal penting. Dikelompokkan berdasarkan satuan konsep, tema dan kategori tertentu akan memberikan gambaran yang lebih tajam tentang hasil pengamatan, juga mempermudah evaluator untuk mencari kembali data sebagai tambahan atas data sebelumnya yang diperoleh jika diperlukan.

#### **b. Display Data.**

Data yang diperoleh dikategorisasikan menurut pokok permasalahan dan dibuat dalam bentuk matriks sehingga memudahkan evaluator untuk melihat pola-pola hubungan satu data dengan data lainnya.

#### **c. Menafsirkan Data.**

Dalam menafsirkan data dapat menggunakan model analisis konten. Dalam model ini, kegiatan yang dilakukan adalah mengklarifikasi istilah-istilah, tanda, simbol atau kode yang dipakai dalam komunikasi, dengan menggunakan beberapa patokan dalam klarifikasi dan menggunakan teknik analisis dalam memprediksikan.

#### **d. Menyimpulkan dan Verifikasi.**

Data yang telah ditafsirkan kemudian disimpulkan. Untuk mengecek kebenaran dari apa yang telah ditafsirkan dan disimpulkan maka dilakukan verifikasi. Kegiatan ini adalah mencocokkan kembali apakah semua data telah tercakup dalam kegiatan analisis dan penafsiran, apakah penafsirannya sesuai, apakah perlu ada konfirmasi ulang pada sumber data atau informan, apakah perlu perbaikan format tafsiran, atau perlu data pendukung untuk memperkuat.

#### **e. Meningkatkan Keabsahan Hasil.**

Kegiatan ini adalah untuk menjawab kelemahan yang sering dialami dalam pendekatan kualitatif, berkaitan dengan validitas dan reliabilitas data. Untuk meningkatkan keabsahan hasil, upaya yang evaluator lakukan atas hasil yang diperoleh, ada beberapa prinsip utama yang harus diperhatikan yaitu:

1) *Krebilitas (credibility)*.

Usaha untuk membuat lebih terpercaya (*credible*) proses, interpretasi dan temuan dalam evaluasi ini yaitu dengan cara:

- Keterikatan yang lama peneliti dengan yang diteliti yaitu dilaksanakan dengan tidak tergesa-gesa sehingga pengumpulan data dan informasi tentang situasi sosial dan fokus evaluasi akan diperoleh secara sempurna. Dengan kata lain aktivitas ini akan meningkatkan kualitas keterlibatan evaluator dalam kegiatan pengumpulan data di lapangan.
- Ketekunan pengamatan terhadap aktivitas program guna memperoleh informasi yang sah.
- Melakukan triangulasi (*triangulation*) yaitu informasi yang diperoleh dari beberapa sumber diperiksa silang dan antara data wawancara informan dengan informan lainnya. Antara data wawancara dengan observasi dan sebagainya. Hal ini dilakukan untuk mengecek kebenaran data dan sekaligus nantinya untuk mempertajam analisis evaluator terhadap hubungan sejumlah data.
- Mendiskusikan dengan teman sejawat yang tidak berperan serta dalam evaluasi, sehingga evaluasi ini akan mendapat masukan dari pihak lain. Diskusi juga dapat dilakukan dengan *expert* yang dapat memberikan masukan dan saran-saran dalam proses evaluasi.
- Analisis kasus negatif yaitu menganalisis dan mencari kasus atau keadaan yang menyanggah temuan evaluasi, sehingga tidak ada lagi bukti yang menolak temuan evaluasi.
- Pengujian ketepatan referensi terhadap data temuan dan interpretasi. Dalam hal ini dilakukan *member check* yaitu pengecekan terhadap hasil yang diperoleh guna perbaikan dan tambahan dengan kemungkinan kekeliruan atau kesalahan dalam memberikan data yang dibutuhkan evaluator.

2) *Dapat Ditransfer (transferability)*.

Pembaca laporan evaluasi program ini diharapkan mendapat gambaran yang jelas mengenai latar (situasi) yang bagaimana agar hasil evaluasi dapat diaplikasikan atau diberlakukan kepada konteks atau situasi lain yang sejenis dalam rangka pemecahan masalah sosial. Dengan kata lain rekomendasi yang dihasilkan dari evaluasi yang dilakukan dapat diaplikasikan oleh lembaga pemakai lainnya.

3) Dapat dipercaya kebenarannya (*dependability*).

Evaluator berupaya dan mengusahakan konsistensi dalam keseluruhan proses yang terjadi pada evaluasi ini agar dapat memenuhi persyaratan yang berlaku. Semua aktivitas evaluasi harus ditinjau ulang terhadap data yang telah diperoleh dengan memperhatikan konsistensi dan dapat dipertanggungjawabkan.

4) Dapat dikonfirmasikan (*confirmability*).

Data harus dapat dipastikan keterpercayaannya atau diakui oleh banyak orang (objektivitas) sehingga kualitas data dapat dipertanggungjawabkan sesuai spektrum, fokus dan latar alamiah evaluasi yang dilakukan.

**f. Narasi Hasil Analisis.**

Pembahasan dalam evaluasi yang menggunakan pendekatan kualitatif dalam penggalan datanya, menyajikan informasi dalam bentuk teks tertulis atau bentuk-bentuk gambar mati atau hidup seperti foto, video dan lain-lain. Dalam menarasikan data kualitatif ada beberapa hal yang harus diperhatikan yaitu:

- 1) Tentukan bentuk (*form*) yang akan digunakan dalam menarasikan data.
- 2) Hubungkan bagaimana hasil yang berbentuk narasi itu menunjukkan tipe/bentuk keluaran yang sudah didesain sebelumnya.
- 3) Jelaskan bagaimana keluaran yang berupa narasi itu mengkomparasikan antara teori dan literasi-literasi lainnya yang mendukung topik.

Arikunto dan Jabar (138-140) memaparkan dalam melakukan analisis data kualitatif dapat dilakukan dengan dua cara yaitu:

**1. Data narasi berpotensi tabulasi.**

Data jenis ini mengacu pada jawaban responden yang tingkat kemunculannya tinggi, artinya jawaban yang sering muncul karena diminati oleh responden. Contoh:

**Tabel VI. 5. Data Narasi Berpotensi Tabulasi**

<p>Pertanyaan:</p> <p>Bagaimanakah pendapat Bapak/Ibu tentang pentingnya pelatihan calon kepala sekolah di masa yang akan datang?</p> <p>a. Harus dilakukan, karena .....</p> <p>b. Tidak perlu dilaksanakan, karena .....</p> <p>c. Harus dilakukan dengan perbaikan, yaitu.....</p> <p>d. ....</p>
--

Contoh item pertanyaan di atas akan menghasilkan dua macam data narasi sekaligus. Pilihan alternatif dengan huruf (a), (b) dan (c) merupakan data narasi yang berpotensi tabulasi. Evaluator dapat menghitung berapa responden yang memilih pilihan (a), (b), (c) atau (d) sedangkan alasan yang dituliskan dibelakang pilihan tersebut merupakan data nontabulasi. Dengan demikian, hanya jawaban alternatif saja yang dapat ditabulasi.

## 2. Data Narasi Nontabulasi

Data narasi nontabulasi adalah data yang berwujud kalimat atau uraian yang sangat individual dan unik karena merupakan pendapat responden secara individual. Contoh butir pertanyaan di atas menghendaki responden memberikan pendapatnya mengenai perlunya pelatihan di masa yang akan datang dan uraiannya tidak dibatasi. Apapun yang diberikan, evaluator harus dapat menerimanya. Ketepatan dalam memberikan argumen merupakan masukan yang berharga bagi pengambil keputusan di dalam menentukan tindak lanjut.

Walaupun data narasi nontabulasi tidak dapat diubah atau dimodifikasi, tetapi masih dapat disiasati agar mudah diolah. Pemikiran ini berangkat dari keunikan manusia yang tinggi, tetapi jika mendapatkan data dari sekian banyak responden maka ditemui data yang memiliki kesamaan. Kesamaan inilah yang memberikan peluang untuk melakukan pengelompokkan atas data tersebut dan kemudian dapat diolah. Contoh:



**Tabel VI. 6. Pengelompokkan Data**

Pilihan	Alasan	Jumlah
Harus dilakukan	Untuk memimpin perlu bekal	7
	Tidak semua calon mampu	3
	Untuk mendapatkan standar kemampuan	2
Tidak perlu dilaksanakan	Semua guru memiliki pengalaman mengelola sekolah	2
	Materi yang diajarkan tidak asing bagi guru	3
Harus dilakukan dengan perbaikan	Untuk memimpin perlu persiapan, tetapi disesuaikan dengan tugas-tugas peserta	3
	Pelatihan harus disertai dengan tindak lanjut yang jelas	4

Berdasarkan tabulasi di atas, evaluator program dapat memperoleh gambaran yang jelas tentang pendapat responden terhadap pelatihan. Jika ada pendapat lain yang muncul hanya satu kali, tidak perlu ditabulasi, tetapi jika dianggap cukup baik, lebih baik dimasukkan ke dalam tabel sebagai bahan pertimbangan.

Selanjutnya menurut Purwanto dan Suparman (1999:198) setidaknya terdapat dua jenis prosedur analisis kualitatif yang dipergunakan dalam evaluasi program. Prosedur analisis tersebut digunakan untuk mengolah data yang berbentuk deskripsi sebagai hasil wawancara, studi kasus, catatan kunjungan lapangan, atau rekaman dan sebagainya, dan data tersebut biasanya bersifat naratif bukan angka-angka. Kedua prosedur analisis kualitatif tersebut dijelaskan sebagai berikut:

**a. Analisis isi dan catatan.**

Analisis isi dan catatan dilakukan dengan cara mengelompokkan data ke dalam kategori-kategori dan variabel yang telah ditentukan sebelumnya. Kemudian dilanjutkan dengan cara mengidentifikasi kata-kata kunci yang menunjukkan inti pendapat, saran atau kritik. Khusus untuk data yang dikumpulkan melalui wawancara, rekamannya atau catatannya harus dibuat transkripnya. Transkrip ini tidak boleh memuat inteprestasi si pembuatnya, agar pengidentifikasian kata kunci terhindar dari bias.

**b. Pendekatan gejala sosial.**

Pendekatan ini menyediakan enam macam cara untuk memeriksa dan menginterpretasikan yang didasarkan kepada suatu kontinum

tentang tingkatan gejala sosial dari yang paling kecil sampai yang paling besar. Tabel berikut memberikan penjelasan tentang keenam macam cara tersebut dan sekaligus menunjukkan dua macam analisisnya yang digunakan untuk membantu pemeriksaan data:

**Tabel VI. 7 Pendekatan Gejala Sosial**

	Penggambaran statis tentang gejala	Penggambaran tentang fase, langkah dan urutan
1. Perbuatan: tindakan spesifik dan singkat	<ul style="list-style-type: none"> <li>• Terdiri dari tindakan apa?</li> <li>• Menunjukkan kategori apa?</li> <li>• Label apa yang bisa diberikan?</li> <li>• Tindakan apa yang berulang?</li> </ul>	<ul style="list-style-type: none"> <li>• Apa tujuan suatu tindakan?</li> <li>• Apa urutannya dari sebuah tindakan?</li> <li>• Apa tahapan dari tindakan yang berulang?</li> </ul>
2. Kegiatan: tindakan yang panjang atau berkelanjutan, khususnya oleh kelompok	<ul style="list-style-type: none"> <li>• Apa pola dan kategori dari kegiatan yang ada?</li> <li>• Adakah ciri/jenis karakteristik dari kegiatan aktor dan hasilnya?</li> </ul>	<ul style="list-style-type: none"> <li>• Adakah langkah-langkah dan fase yang dapat diidentifikasi untuk seluruh atau partisipan yang berbeda?</li> </ul>
3. Makna: simbol verbal dari orang yang menjelaskan dan mengarahkan tindakan	<ul style="list-style-type: none"> <li>• Apa ide dasar dari pelaku?</li> <li>• Apa norma dan harapan?</li> <li>• Apa penjelasan yang diberikan partisipan?</li> </ul>	<ul style="list-style-type: none"> <li>• Apakah terjadi perubahan makna?</li> <li>• Perubahan apa yang terjadi dalam pandangan, norma, harapan dan sebagainya?</li> <li>• Adakah pola tahapan perubahannya?</li> </ul>
4. Partisipasi: bagaimana orang berbuat, bereaksi dalam suatu kondisi yang diatur	<ul style="list-style-type: none"> <li>• Orang seperti apa yang terlibat?</li> <li>• Apa tanda-tanda partisipan?</li> <li>• Apa bentuk tingkatan dan pola keterlibatannya?</li> </ul>	<ul style="list-style-type: none"> <li>• Adakah siklus dan tahapan partisipanya?</li> </ul>
5. Hubungan: di antara beberapa orang pada waktu dan kondisi yang ditentukan	<ul style="list-style-type: none"> <li>• Hubungan dan interaksi macam apa yang ada?</li> <li>• Apa ada hirarki?</li> </ul>	<ul style="list-style-type: none"> <li>• Apakah hubungan bergerak mengikuti pola dari tahapan perubahan?</li> </ul>
6. Pengaturan	<ul style="list-style-type: none"> <li>• Apa tanda-tanda pengaturannya?</li> <li>• Apa jenis pengaturan yang dapat dilihat?</li> </ul>	<ul style="list-style-type: none"> <li>• Tahapan atau fase pengaturan apa yang dilalui?</li> <li>• Pada tahap atau fase mana pengaturan sekarang?</li> </ul>

Pendekatan lain yang dapat digunakan untuk melakukan analisis data kualitatif dalam evaluasi program adalah pendekatan *key incident* yang dikemukakan oleh Wilcox. Menurut Wilcox, analisis data kualitatif tergantung pada hakikat data dan kerangka konsep yang dipakai dalam analisis (Tayibnapis, 2000:123).

Pendekatan *key incident* adalah mengambil catatan tentang kejadian penting dari lapangan, menghubungkannya dengan kejadian lain, fenomena, teori, dan menulisnya sehingga orang lain dapat melihat secara umum, universal dalam kenyataan tentang hubungan antar bagian-bagian dan keseluruhannya. Melalui pendapatan *key incident* memungkinkan untuk memasukkan sejumlah besar kesimpulan, bermacam-macam data dari berbagai sumber, termasuk *field notes*, dokumen informasi tentang demografis, wawancara yang tidak terstruktur dan sebagainya.

## **E. Pengolahan Data Wawancara**

Data atau informasi yang telah dikumpulkan melalui wawancara biasanya dicatat atau direkam. Selanjutnya data atau informasi tersebut disebut catatan lapangan perlu diolah menjadi informasi yang dapat digunakan untuk pembuatan keputusan.

Hasil wawancara berupa catatan atau rekaman tersebut mula-mula dibuat transkripnya. Pokok-pokok jawaban atas pertanyaan dikumpulkan dan dikelompokkan menjadi beberapa kategori sesuai dengan sifat dan permasalahannya. Transkrip wawancara tersebut dideskripsikan. Tabulasi kadang-kadang diperlukan untuk mengolah data tentang pendapat atau sikap sejumlah responden sehingga diperoleh angka-angka. Selanjutnya diadakan telaah dan analisis terhadap informasi tersebut sehingga diperoleh suatu kesimpulan. Telaah atau pemeriksaan data tersebut dimaksudkan untuk memperoleh keabsahan dan kecukupan data. Dalam proses analisis informasi atau data hasil wawancara harus dibedakan dan dipilah-pilah jenis datanya.

Berdasarkan suatu hasil wawancara yang cukup lengkap akan dapat dibuat kesimpulan-kesimpulan tentang masalah yang berkaitan dengan kelebihan dan kelemahan dalam perencanaan dan pelaksanaan suatu program. Hasil wawancara pada dasarnya adalah suatu kesimpulan yang akurat dan dapat dipercaya dan diharapkan dapat dijadikan dasar dan bahan untuk pembuatan keputusan. Informasi yang akurat tersebut dirumuskan berdasarkan

pengolahan data wawancara. Informasi tersebut disajikan dalam bahasa yang lugas, sederhana, mudah dipahami dan ditafsirkan.

Contoh:

Hasil wawancara dengan peserta program pelatihan diperoleh kesimpulan sebagai berikut: (1) perlu ada materi pelatihan baru yaitu tentang peraturan dan prosedur terbaru dalam penguasaan pengetahuan dan keterampilan yang dilatihkan, dan (2) *briefing* merupakan alternatif untuk menjelaskan peraturan dan prosedur baru, apabila penambahan jam pelatihan tidak mungkin dilakukan.

Hasil wawancara tentang program pelatihan di atas berupa rekomendasi tentang penyempurnaan penyelenggaraan pelatihan. Rekomendasi tersebut disimpulkan dan disusun berdasarkan data berupa pendapat, penilaian, pandangan dan gagasan dari pihak-pihak yang diwawancarai dan dianggap kompeten.

## F. Interpretasi Data

Hasil pengolahan data selanjutnya perlu diinterpretasikan untuk memudahkan dalam penyusunan kesimpulan dan rekomendasi. Membuat interpretasi berarti membuat hasil analisis tersebut menjadi sesuatu yang dapat dibandingkan dengan standar tertentu, harapan atau referensi lain. Membandingkannya dengan tingkat signifikansi adalah salah satu cara menginterpretasikan hasil analisis kuantitatif.

Membuat interpretasi juga bermakna menafsirkan temuan dari kegiatan evaluasi di lapangan. Evaluator perlu mencurahkan segenap kemampuan analitis dan kritisnya dalam menguraikan dan merangkai permasalahan yang menjadi fokus evaluasi. Berbagai referensi yang dapat dijadikan dasar untuk interpretasi data menurut Purwanto dan Suparman (1999:201) adalah:

1. Pertimbangan ahli (*expert judgement*).
2. Harapan staf.
3. Pendapat masyarakat.
4. Standar institusi.
5. Peraturan, hukum.
6. Laporan penelitian.
7. Norma.

Selanjutnya menurut Worthen dkk sebagaimana dikutip Tayibnapis (2000:129) memberikan petunjuk untuk melakukan menafsirkan hasil pengolahan data evaluasi sebagai berikut:

1. Menentukan apakah tujuan sudah dicapai.
2. Menentukan apakah hukum, norma-norma, demokrasi, aturan dan prinsip-prinsip tidak dilupakan.
3. Menentukan apakah analisis kebutuhan telah dikurangi.
4. Menentukan nilai pencapaian.
5. Bertanya kepada kelompok penilai, melihat kembali data, menilai keberhasilan, dan kegagalan, menilai kelebihan dan kelemahan tafsiran.
6. Membandingkan variabel-variabel penting dengan hasil yang diharapkan.
7. Membandingkan analisis yang dilaporkan oleh program lain yang usahanya sama.
8. Menafsirkan hasil analisis dengan prosedur yang menghasilkannya.

Purwanto dan Suparman (1999:202) memaparkan pedoman dalam melakukan interpretasi atau penafsiran hasil pengolahan data sebagai berikut:

1. Selesaikan dengan bukti-bukti yang lengkap. Hal ini penting agar dapat dihindari kesalahan interpretasi.
2. Jangan berasumsi bahwa signifikansi statistik adalah berarti signifikan secara nyata, atau sebaliknya jika tidak signifikan, bukan berarti secara praktis tidak ada signifikansi.
3. Waspada efek regresi. Dalam hal ini skor atau nilai ekstrim yang di luar batas (terlalu tinggi atau terlalu rendah) dibandingkan dengan nilai dalam kelompok.
4. Carilah informasi dan konsisten dari sumber lain. Data tunggal sering tidak cukup, misalnya peserta menyatakan dengan *rating* bahwa ia puas dengan program pelatihan yang diikutinya. Janganlah lantas segera menyimpulkan, sebaiknya dilihat apakah peserta tersebut mengikuti keseluruhan proses dan kegiatan pelatihan. Apa kata instruktur, benarkah demikian halnya. Jadi pada hal-hal tertentu perlu melakukan triangulasi sumber data.
5. Ketahuilah kapan harus berhenti. Ingatlah bahwa tidak akan pernah memperoleh hasil analisis yang sempurna. Batas berhenti untuk

data kuantitatif adalah jika telah yakin dapat mempertahankannya, sedangkan untuk data kualitatif adalah ketika dihadapkan kepada pengulangan (*redundancy*).

6. Hati-hatilah dengan keterbatasan teknik analisis data yang digunakan. Ada keterbatasan yang melekat pada setiap teknik analisis data yang digunakan, dan asumsi-asumsi yang mendasari penggunaannya. Misalnya koefisien korelasi tidak memperhitungkan semua varians, dan ia hanya cocok untuk menyimpulkan hubungan.
7. Hati-hatilah dengan keterbatasan data yang diperoleh. Tidak ada penemuan dalam kondisi vakum, semua harus dihadapkan pada data yang lain.
8. Hati-hatilah dengan kemungkinan keterlibatan audien dalam interpretasi.

## BAB VII

# KESIMPULAN DAN REKOMENDASI EVALUASI PROGRAM

### A. Kesimpulan

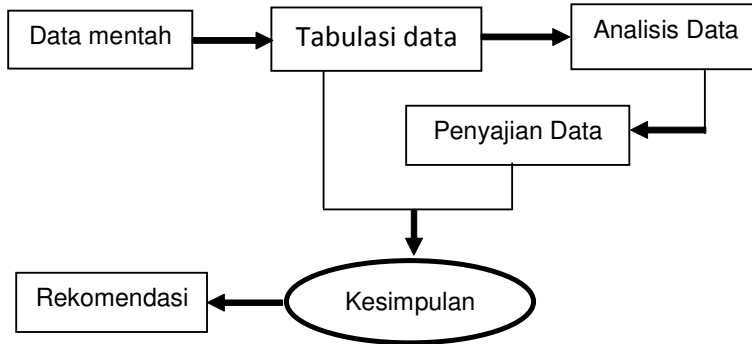
**K**esimpulan atau simpulan adalah sesuatu yang merupakan inti dari sederetan informasi atau sajian yang menyatakan tentang status program yang sedang dievaluasi (Arikunto dan Jabar, 2009:191). Dalam evaluasi program, simpulan diambil dari atau dibuat berdasarkan hasil analisis data yang sudah disajikan dalam bentuk yang sudah sistematis, ringkas dan jelas.

Simpulan berbentuk kalimat pernyataan kualitatif yang menunjukkan keadaan atau sifat sesuatu sehingga di dalam gerak kegiatan program dengan cepat dapat diketahui di mana posisi hasil kegiatan tersebut dalam mencapai tujuan yang telah ditetapkan. Dengan rumusan pernyataan yang bersifat kualitatif ini diharapkan bahwa di dalam kesimpulan tersebut tidak lagi mencantumkan angka-angka rasio, presentase, lebih-lebih lagi masih dalam bentuk frekuensi.

Urgensi pencantuman kesimpulan dalam laporan evaluasi program dijelaskan oleh Arikunto dan Jabar (2009:192) bahwa sifat kegiatan evaluasi program adalah kegiatan yang berorientasi pada pengambilan keputusan. Untuk dapat menentukan suatu kebijakan para pengambil keputusan memerlukan masukan yang lengkap, jelas, rinci dan operasional, berupa rekomendasi dari evaluator.

Evaluator tidak dapat membuat rekomendasi apabila tidak ada informasi yang jelas, ringkas, padat dan sistematis berdasarkan data yang andal dan dapat dipercaya. Jadi kesimpulan harus ada karena berada dalam rangkaian proses, dan hal itu sangat dibutuhkan bagi para pengambil keputusan untuk tindak lanjut dari kegiatan sebuah program.

Kesimpulan dalam laporan evaluasi adalah intisari dari suatu proses. Di dalam suatu rangkaian proses, kesimpulan berada dibagian paling akhir (sebelum digunakan untuk merumuskan rekomendasi) setelah melalui proses awal yakni data mentah, tabulasi, analisis data dan penyajian data. Alur proses kesimpulan dipaparkan oleh Arikunto dan Jabar (2000:193) pada diagram berikut:



Gambar 7.1 Bagan Alur Pemrosesan Data

Berdasarkan gambar 7.1 di atas maka dicermati alur pemrosesan data mentah sampai pada rumusan rekomendasi, sebagai berikut:

1. Data mentah dari instrumen yang terkumpul diolah dalam proses tahapan pertama yang tabulasi.
2. Hasil dari proses tabulasi berguna untuk tiga proses yang mengikutinya meskipun tidak berurutan dengannya yaitu analisis data, penyajian data dan membuat kesimpulan.
3. Proses penyajian data mendapatkan sumber bahan dari tabulasi dan analisis data.
4. Kesimpulan dilakukan berdasarkan sumber bahan dari penyajian data dan tabulasi data.
5. Proses terakhir yaitu membuat rekomendasi, didasarkan atas sumber bahan dari kesimpulan. Kaitan antara kesimpulan dan rekomendasi, atau dengan kata lain adalah proses melanjutkan kesimpulan menjadi rekomendasi ini di dalam proses evaluasi program merupakan langkah yang paling penting dan menunjukkan seberapa tinggi kemampuan evaluator dalam bertugas mengevaluasi program.



Oleh karena urgensinya proses dari kesimpulan dilanjutkan menjadi rekomendasi tersebut maka kesimpulan sendiri merupakan sesuatu yang juga sangat penting kedudukannya maupun isi rumusannya. Istilah populer yang banyak digunakan dalam memuat isi rumusan kesimpulan yaitu menggigit, aktraktif dan menantang, sehingga pembaca, terutama sekali pengambil keputusan dapat berpikir dengan masalah tersebut dan dengan senang hati mengarahkan pusat perhatiannya untuk melahirkan rekomendasi yang sifatnya positif tentang program yang dievaluasi.

Arikunto dan Jabar (2009:194) memberikan arahan untuk membuat kesimpulan secara tepat dan bersifat menggigit yaitu rumusan kesimpulan harus singkat, jelas dan ringkas. Rumusan tersebut tidak menggunakan angka-angka lagi, tetapi dimaknai dengan kata-kata predikat seperti baik, berkualitas tinggi, cukup mengesankan, tidak menarik dan sebagainya, semuanya merupakan kata sifat atau keadaan yang menunjukkan posisi tertentu.

Tayibnapis (2000:172) memberikan tips dalam merumuskan simpulan dari evaluasi program yang dilakukan sebagai berikut:

1. Apa kesimpulan umum tentang keefektifan program secara keseluruhan? Sampai seberapa jauh kebenaran kesimpulan?
2. Apakah perlu ada penilaian terhadap beberapa aspek kebijaksanaan program?
3. Apakah ada hal-hal yang diperkirakan terjadi sebagai akibat program yang dipertimbangkan untuk memperoleh gambaran yang lengkap tentang *impact* program?

Berikut dipaparkan suatu ilustrasi mengenai kesimpulan dari program bantuan operasional sekolah (BOS) di kecamatan X sebagai berikut:

1. Ketercapaian tujuan program BOS.  
Secara umum, tujuan dari program BOS ini sudah tercapai dengan baik, yaitu masyarakat mendapatkan bantuan keringanan biaya pendidikan yang biasanya dikenakan kepada mereka dalam bentuk biaya SPP atau uang sekolah, daftar ulang di samping itu alokasi dana BOS juga diperuntukan untuk membayar honor guru khususnya guru di sekolah swasta.
2. Sasaran program.  
Sasaran program dana BOS diperuntukkan kepada seluruh siswa yang terdaftar dan terverifikasi di sekolah masing-masing baik sekolah

negeri maupun sekolah swasta. Besarnya berdasarkan tingkatan sekolah yaitu besaran untuk siswa sekolah dasar berbeda dengan siswa sekolah menengah pertama.

3. Ketercukupan dana program BOS.

Dana BOS yang disalurkan melalui sekolah-sekolah baru mencakup sebagian kecil kebutuhan sekolah dalam penyelenggaraan pelayanan pendidikan dan pembelajaran. Jika dikaitkan dengan besarnya dana yang harus ditanggung masyarakat, maka program BOS ini baru mencakup sebagian kecil saja kebutuhan biaya langsung pendidikan anak-anak mereka.

4. Efektivitas penggunaan dana BOS oleh sekolah.

Semua sekolah di kecamatan X telah memanfaatkan dan menggunakan dana BOS yang berasal dari pemerintah dengan peruntukan yang telah ditetapkan dalam petunjuk teknis pelaksanaan penggunaan dana BOS.

5. Sistem pengawasan.

Pengawasan yang dilakukan oleh internal sekolah melalui komite sekolah dan pihak Dinas Pendidikan juga cukup efektif dalam membimbing dan mengawasi jalannya pelaksanaan program BOS.

Berdasarkan kesimpulan spesifik di atas, maka dapatlah ditarik kesimpulan umum dari program BOS yang dievaluasi, yaitu program BOS di kecamatan X tersebut sudah cukup baik dilaksanakan, namun demikian besaran dana yang diterima belum mampu mencukupi kebutuhan sekolah, di sisi lain masyarakat terbantu dengan adanya program BOS ini khususnya terkait dengan pembiayaan langsung biaya pendidikan anak-anak mereka.

## B. Rekomendasi

Rekomendasi berasal dari bahasa Inggris yaitu *recommendation* yang bermakna pujian atau pengajuan pendapat sebagai pujian atas kebaikannya. Dalam konteks evaluasi program dapatlah dipahami bahwa rekomendasi berisikan pendapat-pendapat ataupun saran-saran sehubungan dengan evaluasi terhadap program yang dilakukan.

Urgensi rekomendasi dalam laporan evaluasi program dipaparkan oleh Arikunto dan Jabar (2009:195) yaitu sebagai upaya untuk memperoleh informasi sehubungan dengan dijalankannya suatu kebijakan atau program.

Hasil dari evaluasi program diharapkan berupa saran-saran bagi dikeluarkannya kebijakan lanjutan apabila sudah diperoleh informasi yang dikumpulkan oleh evaluator program. Informasi yang terkumpul dan sudah dirumuskan menjadi kesimpulan, agar menjadi jelas arah kebijakan yang harus dilakukan pengambil keputusan, perlu diubah menjadi suatu rumusan rekomendasi. Dalam hal ini yang diberikan kepada pengambil keputusan bukan sekedar saran-saran, tetapi dinyatakan pernyataan yang cenderung memuji program atau bagian yang perlu diperbaiki agar pengambil keputusan semakin berkeyakinan untuk menetapkan kebijakan lanjutan.

Pihak pemesan atau sponsor biasanya membaca hasil laporan evaluasi fokus kepada rekomendasi yang disampaikan evaluator. Oleh sebab itu pemaparan rekomendasi hendaknya disiapkan evaluator program dengan hati-hati. Rekomendasi atau saran-saran biasanya diberikan untuk satu tindak lanjut, dilengkapi dengan perbaikan atas kelemahan-kelemahan atau kekurangan-kekurangan program. Namun lebih baik memberikan beberapa rekomendasi untuk alternatif tindakan, dan setiap alternatif tindakan didukung oleh penemuan dan data evaluasi.

Evaluator dapat membuat rekomendasi berdasarkan kriteria yang relevan dengan pertanyaan evaluasi yaitu validitas-reliabilitas, dan objektivitas, evaluator menilai data yang ditemukan signifikan atau tidak. Dari penilaian ini evaluator membuat rekomendasi khusus untuk tindak lanjut. Rekomendasi dan alternatif harus langsung berhubungan dengan tujuan evaluasi dan aspek-aspek program yang dimintai pemakai terutama para pemegang keputusan dalam program.

Rekomendasi atau saran bukan berupa daftar keinginan, tetapi harus sesuai dengan logika dari penilaian yang dibuat berdasarkan data. Rekomendasi harus langsung mengarah kepada aspek-aspek khusus program atau atas tindakan khusus namun cukup umum bagi pemegang keputusan untuk memutuskan sesuai dengan situasi dan kondisi program.

Arikunto dan Jabar (2009:199) memaparkan hal-hal yang perlu diperhatikan evaluator dalam menyusun rekomendasi yaitu mengenai perlunya melihat dengan lebih cermat terhadap alasan-alasan atau cara-cara yang diusulkan oleh responden tentang upaya peningkatan kualitas program yang dievaluasi di masa yang akan datang. Dikarenakan yang memanfaatkan rekomendasi adalah pengambil keputusan, maka perlu dijelaskan kepada siapa rekomendasi tersebut ditujukan, apa yang harus dilakukan, dalam bentuk apa perlakuan itu dilakukan dan sebagainya. Selain itu evaluator perlu hati-hati memilih

urutan cara-cara yang diusulkan oleh responden, tetapi juga sebaiknya jangan diminta untuk berpikir lebih jauh, tetapi diusahakan tinggal memilih alternatif saja.

Selanjutnya Tayibnapis (2000:173) memberikan beberapa tips dalam menyusun rekomendasi sebagai berikut:

1. Berdasarkan data khusus, apa saran-saran dan pilihan yang dapat diberikan terhadap program? Apa kelebihan-kelebihan program dan aspek-aspek apa saja yang perlu atau yang dapat dikembangkan dan diperbaiki?
2. Apakah tujuan evaluasi juga memberikan rekomendasi dan saran-saran pilihan? Apakah pemakai ingin mengetahui efektivitas atau keefektifan program atau apakah ingin juga mengetahui kelemahan-kelemahan program?
3. Perkiraan, hipotesis atau dugaan apa yang diberikan data? Apa rekomendasi yang diberikan untuk program selanjutnya atau penelitian yang akan dilakukan?

Berikut dipaparkan suatu ilustrasi mengenai rekomendasi dari program bantuan operasional sekolah (BOS) di kecamatan X sebagai berikut:

1. Rekomendasi untuk ketercapaian tujuan program BOS.  
Oleh karena kesimpulannya berbunyi: secara umum, tujuan dari program BOS ini sudah tercapai dengan baik, yaitu masyarakat mendapatkan bantuan keringanan biaya pendidikan yang biasanya dikenakan kepada mereka dalam bentuk biaya SPP atau uang sekolah, daftar ulang di samping itu alokasi dana BOS juga diperuntukan untuk membayar honor guru khususnya guru di sekolah swasta.

Maka rekomendasi yang diberikan sebagai berikut:

Beberapa aspek penunjang yang belum tercapai dalam tujuan program perlu ditinjau ulang dan diselesaikan pada kegiatan sejenis di masa yang akan datang. Program dana BOS ini ternyata cukup signifikan membantu sekolah dan orang tua siswa dalam meringankan beban biaya oleh karena itu perlu dilanjutkan di tahun-tahun berikutnya.

2. Rekomendasi untuk sasaran program.  
Oleh karena kesimpulannya berbunyi: sasaran program dana BOS diperuntukkan kepada seluruh siswa yang terdaftar dan terverifikasi di sekolah masing-masing baik sekolah negeri maupun sekolah swasta.

Besarnya berdasarkan tingkatan sekolah yaitu besaran untuk siswa sekolah dasar berbeda dengan siswa sekolah menengah pertama.

Maka rekomendasi yang diberikan sebagai berikut:

Sebaiknya sasaran program dana BOS pada nominal jumlah yang diberikan kepada setiap siswa pada tahun berikutnya ditingkatkan mengingat tingkat inflasi semakin tinggi.

3. Rekomendasi untuk ketercukupan dana program BOS.

Oleh karena kesimpulannya berbunyi: dana BOS yang disalurkan melalui sekolah-sekolah baru mencakup sebagian kecil kebutuhan sekolah dalam penyelenggaraan pelayanan pendidikan dan pembelajaran. Jika dikaitkan dengan besarnya dana yang harus ditanggung masyarakat, maka program BOS ini baru mencakup sebagian kecil saja kebutuhan biaya langsung pendidikan anak-anak mereka.

Maka rekomendasi yang diberikan sebagai berikut:

Perencanaan program dana BOS yang dilakukan sebaiknya mengidentifikasi secara lengkap kebutuhan penyelenggaraan pendidikan dan pembelajaran yang vital bagi sekolah dan juga pembiayaan yang dikeluarkan oleh orang tua siswa, sehingga besaran dana yang dianggarkan benar-benar mencerminkan kebutuhan sekolah dan kebutuhan orang tua.

4. Rekomendasi untuk efektivitas penggunaan dana BOS oleh sekolah.

Oleh karena kesimpulannya berbunyi: semua sekolah di kecamatan X telah memanfaatkan dan menggunakan dana BOS yang berasal dari pemerintah dengan peruntukan yang telah ditetapkan dalam petunjuk teknis pelaksanaan BOS.

Maka rekomendasi yang diberikan sebagai berikut:

Sistem pengelolaan yang selama ini dilakukan Dinas Pendidikan, sebaiknya tetap digunakan untuk program yang akan datang, dengan sedikit penyesuaian dan perbaikan/penambahan agar kualitas program menjadi lebih baik. Sekolah perlu diberi bimbingan terkait dengan perencanaan pembiayaan lebih lanjut oleh Dinas Pendidikan, atau instansi yang berwenang.

5. Rekomendasi untuk sistem pengawasan.

Oleh karena kesimpulannya berbunyi: pengawasan yang dilakukan oleh internal sekolah melalui komite sekolah dan pihak Dinas Pendidikan

juga cukup efektif dalam membimbing dan mengawasi jalannya pelaksanaan program BOS.

Maka rekomendasi yang diberikan sebagai berikut:

Sistem pengawasan yang sekarang ada sudah cukup baik dan dapat digunakan untuk pengawasan program di masa yang akan datang, hanya perlu ada beberapa penyesuaian agar sistem pengawasan yang diterapkan dapat merangsang kreativitas sekolah dalam mengembangkan program penyelenggaraan pendidikan dan pembelajaran.

Berdasarkan ilustrasi di atas maka pihak pengambil keputusan (*decision maker*) dapat menentukan atau mengambil keputusan. Dalam hal ini terdapat empat kemungkinan kebijakan atau keputusan yang diambil berdasarkan rekomendasi tersebut yaitu:

1. Menghentikan program.

Oleh karena dipandang bahwa program tersebut tidak bermanfaat atau kurang bermanfaat atau tidak dapat terlaksana dengan baik dan sesuai dengan harapan, atau terdapat berbagai kendala yang besar menghambatnya dan berbagai alasan lainnya yang kuat, maka program dihentikan.

2. Merevisi program.

Oleh karena ada bagian-bagian yang kurang tepat dan sesuai dengan harapan dalam hal ini terdapat kesalahan (tidaklah terlalu fatal) dan memungkinkan untuk diperbaiki dengan memberikan pedoman yang lebih rinci dan pengawasan yang lebih baik maka program dilanjutkan tetapi dilakukan revisi terhadap bagian-bagian yang lemah tersebut.

3. Melanjutkan program.

Oleh karena pelaksanaan program yang dilakukan sebelumnya telah menunjukkan segala sesuatu yang sudah berjalan sesuai dengan pedoman dan harapan dari *stakeholder* dan hasil atau dampak yang dari pelaksanaan program memberikan hasil yang tepat guna dan tepat sasaran maka program dilanjutkan pada tahun berikutnya.

4. Menyebarluaskan program.

Oleh karena program tersebut berhasil dengan baik pada satu wilayah tertentu di mana program diluncurkan maka sangatlah baik jika dilaksanakan lagi di wilayah yang berbeda dan lebih luas lagi jangkauannya, sehingga implementasi program dapat dimanfaatkan dan dirasakan masyarakat lebih luas lagi.

## BAB VIII

# PELAPORAN DAN PEMANFAATAN LAPORAN EVALUASI PROGRAM

### A. Pengertian

Laporan evaluasi adalah media komunikasi antara evaluator dengan pihak-pihak yang berkepentingan dan ingin mengetahui hasil evaluasi, oleh karena itu laporan harus mampu berperan sebagai media komunikasi yang baik (Purwanto dan Suparman, 1999:210).

### B. Tujuan dan Manfaat Laporan Evaluasi

Tujuan laporan evaluasi berhubungan langsung dengan tujuan pemakaiannya. Pada evaluasi formatif tujuan utamanya yaitu untuk memperbaiki dan mengembangkan program, dan laporannya harus diserahkan secepatnya kepada pihak yang meminta untuk dilakukan evaluasi program, diinformasikan pula tentang bagaimana program berfungsi dan perubahan-perubahan apa yang harus dilakukan untuk tujuan tersebut (Tayibnapis, 2000:134-135).

Selanjutnya dijelaskan Tayibnapis apabila evaluasi dilakukan adalah evaluasi sumatif, maka laporan harus berisi informasi dan penilaian tentang kegunaan program yang dilaporkan kepada orang-orang yang ingin memakainya, orang yang akan menentukan alokasi sumber-sumber untuk melanjutkan program dan orang-orang yang berhak mengetahui tentang program untuk tujuan-tujuan yang lain.

Fitzpatrick dkk (2004) menjelaskan tujuan laporan evaluasi program dapat dilihat dari dua perspektif yaitu:

1. Tujuan laporan evaluasi formatif berkaitan perbaikan dan mengembangkan program, dan laporannya diserahkan kepada pihak pengguna program.

Laporan evaluasi berisikan bagaimana program berfungsi dan perubahan-perubahan apa yang harus dilakukan untuk mencapai program.

2. Tujuan laporan evaluasi sumatif adalah laporan evaluasi yang berisi informasi dan penilaian (*judgement*) tentang kegunaan program. Laporan disampaikan kepada:
  - a. Pihak-pihak yang ingin mengadopsi program.
  - b. Pihak-pihak yang akan menentukan alokasi-alokasi sumber untuk melanjutkan program.
  - c. Pihak-pihak yang berhak menentukan tentang program untuk tujuan-tujuan yang lain.

Selanjutnya Henry dan Mark (2003), Brinkerhoff, Brethower, Hluchyj, dan Nowakowski (1983), Patton (1986), Coussin dan Leithwood (1986), dan King (1988) memaparkan tujuan laporan evaluasi yaitu:

- Menunjukkan akuntabilitas.
- Membantu dalam membuat keputusan.
- Membawa isu yang menarik perhatian orang lain.
- Membantu pihak atau mempersempit pendapat mereka atas suatu masalah.
- Meyakinkan orang lain untuk mengambil tindakan.
- Mendalami masalah dan menyelidiki.
- Melibatkan *stakeholder* dalam perencanaan program atau pengembangan kebijakan.
- Mendapatkan dukungan untuk program.
- Mempromosikan pengertian dari masalah-masalah.
- Mengubah sikap.
- Mengubah perilaku individu.
- Mengubah sifat atau interaksi antara kelompok-kelompok dialog.
- Kebijakan yang mempengaruhi.
- Memperkenalkan orang-orang yang terlibat untuk cara-cara berpikir baru melalui evaluasi.

Laporan evaluasi program dapat memberikan banyak manfaat namun yang paling penting yaitu menyampaikan pesan, memberi informasi yang tepat kepada audiensi tentang penemuannya dan kesimpulan hasil pengumpulan



informasi, analisis dan penafsiran informasi evaluasi. Dalam hal ini Brinkerhoff sebagaimana dikutip Tayibnapis (2000:135) menjelaskan bahwa manfaat laporan evaluasi bagi pengambil keputusan sebagai berikut: (1) pertanggung-jawaban, (2) menjelaskan, meyakinkan, (3) mendidik, (4) meneliti, (5) dokumen, (6) turut terlibat, (7) mendapat dukungan, (8) menambah pengertian, dan (9) hubungan masyarakat.

### **C. Pendekatan dan Jenis Laporan**

Fitzpatrick dkk (2004) menawarkan suatu pendekatan alternatif dalam melaporkan hasil evaluasi, pendekatan tersebut dinamakan pendekatan evaluasi berorientasi lawan/musuh. Pendekatan evaluasi berorientasi lawan/musuh sebagai cara kreatif dan efektif untuk mempresentasikan informasi kepada audiens dengan memperhatikan pendapat-pendapatnya terkait penilaian program yang dievaluasi.

Makna luas dari pendekatan ini adalah evaluator memperoleh informasi mendengarkan dari sudut pandang yang berlawanan sebagaimana argumentasi yang disampaikan pihak lain. Sehingga nantinya laporan evaluasi yang dibuat memiliki keseimbangan dan perspektif yang berbeda karena terlihat argumentasi dari pihak lain dalam melihat kekuatan dan kelemahan dari program yang dievaluasi. Melalui pendekatan ini proses penilaian terhadap program yang dievaluasi akan berguna dalam hal:

1. Paparan tentang aspek positif dan aspek negatif dari sebuah program.
2. Memuaskan kebutuhan informasi dari berbagai pihak dengan cara yang elegan dan menarik.

Bagi evaluator selain mempersiapkan dan menyajikan informasi dalam laporan evaluasi, maka harus diingat adalah evaluator jangan melupakan dampak dari pesan-pesan informasi tersebut. Dalam hal ini yang dimaksud adalah dampak psikologis kemanusiaan yang mungkin timbul dari laporan evaluasi tersebut. Oleh karena itu evaluator harus memiliki kepekaan terhadap persoalan ini.

Terkait dengan masalah ini, Fitzpatrick dkk (2004) menyarankan dua hal yaitu:

### 1. *Delivering negative messages (mengirimkan pesan negatif).*

Zaman dahulu, pembawa pesan yang menyampaikan berita kepada raja dibayangi-bayangi resiko, jika berita yang disampaikan buruk dan tidak diterima raja, maka mungkin saja pembawa pesan dapat kehilangan lidahnya bahkan kepalanya. Namun saat ini, pembawa pesan evaluasi buruk hanya dikatakan sebagai pihak pengacau saja atau bahasa kerennya provokator. Oleh karena itu evaluator harus peka dalam menyampaikan pesan atau penemuan negatif dari hasil temuan evaluasi program yang dilakukan.

Terkait dengan masalah ini Van Mondfrans (1985) memberikan saran-saran untuk membantu orang-orang yang terlibat dalam evaluasi supaya jangan menelan pil pahit:

- a. Dalam sebuah percakapan singkat, di mana peristiwa utama evaluasi ditinjau dan temuan utama dilihat cuplikannya, informasi negatif dinyatakan dalam konteks yang positif sebanyak mungkin. Hal ini seolah-olah lebih mudah bagi klien untuk menerima informasi negatif dalam bentuk percakapan yang bersahabat.
- b. Sebuah laporan tertulis awal dipersembahkan di mana informasi negatif digambarkan dalam suatu diskusi dengan cara faktual tetapi dari sebagai sebuah perspektif yang mungkin positif. Dengan sering melakukan kunjungan pribadi dalam rangka membahas laporan awal yang dibahas dan klien diizinkan untuk mengusulkan perubahan jika informasi yang dilihatnya tidak tepat. Jika perubahan yang diperlukan dalam laporan awal, maka evaluator tidak boleh mengaburkan informasi negatif atau memungkinkan untuk menjadi disalahtafsirkan. Namun mungkin saja dengan membahas informasi negatif, klien akan membawa faktor-faktor lain yang tidak diketahui evaluator pada waktu laporan awal ditulis. Faktor-faktor lain ini mungkin disertakan dalam laporan akhir sehingga memungkinkan interpretasi yang lebih baik.
- c. Sebuah laporan tertulis akhir disiapkan di mana informasi negatif benar-benar disajikan. Setelah mengalami langkah-langkah sebelumnya, klien lebih siap untuk berurusan dengan informasi negatif, mengalami kesempatan untuk meninjau beberapa hal dalam rangka memikirkan faktor-faktor lain yang relevan. Di samping itu evaluator memiliki kesempatan yang untuk meninjau faktor-faktor lain dan melibatkan klien dalam laporan jika diperlukan bantuannya dalam memaknai informasi negatif.

Benkofske (1994) juga menyarankan kepada semua pihak tentang cara-cara untuk menyiapkan diri menerima penemuan negatif tentang program mereka dan meningkatkan kinerja, terutama apabila temuan yang negatif dari hasil data kualitatif. Oleh karena itu penting untuk dilakukan evaluasi sejak awal melibatkan klien terkait dengan pembahasan tentang apa yang akan mereka lakukan jika data kualitatif menunjukkan negatif.

## **2. *Providing an opportunity for review of the report* (menyediakan kesempatan untuk meninjau laporan).**

Fitzpatrick dkk (2004) menegaskan bahwa sebuah kesombongan apabila evaluator menganggap bahwa laporan evaluasi yang dibuat telah menyajikan hal-hal yang akurat dan wajar dalam semua hal. Intinya yang dimaksud disini adalah melakukan kegiatan untuk meninjau kembali laporan evaluasi sebelum diserahkan kepada klien dengan melakukan peninjauan ulang akan menghilangkan faktor-faktor kesalahan penafsiran dan bias yang mungkin saja dapat dilakukan evaluator.

Untuk alasan-alasan tersebut maka sebaiknya evaluator menyampaikan draft laporan evaluasi kepada klien dan *stakeholder* untuk dimintai pendapatnya antara lain:

- a. Kesalahan kecil, misalnya kesalahan ejaan.
- b. Kesalahan faktual misalnya nama-nama yang salah atau *tittles* (judul), kesalahan dalam bilangan orang yang ikut serta.
- c. Kesalahan interpretasi.

Melalui penyampaian draft awal laporan evaluasi tersebut maka evaluator dapat melihat pendapat klien terkait dengan isi laporan evaluasi berupa pandangan klien tentang kesalahan yang ada maupun dukungan atas fakta-fakta atau interpretasi yang terdapat di dalamnya. Namun demikian evaluator berhak untuk mengabaikan saran-saran dan membuat perubahan-perubahan dalam laporan evaluasi. Melalui penyampaian draft awal laporan juga dapat dilihat meningkatnya keterlibatan individu yang membaca laporan dan tanggung jawab bersama untuk laporan yang baik dan tingkat akurasinya.

Apabila evaluator menolak untuk menerima suatu usulan perubahan dalam laporan, tetapi pembahas (pembaca) dapat menyarankan perbaikan atas ketidakakuratan data, menyesatkan atau tidak wajar. Untuk itu, evaluator dapat mengundang pembahas (pembaca) untuk berbagi pandangan, ditulis dan memasukkannya ke dalam laporan evaluasi.

Selanjutnya terkait dengan jenis laporan evaluasi, Purwanto dan Suparman (1999:208-209) mencatat terdapat 4 (empat) jenis laporan evaluasi sebagai berikut:

**1. Pengumuman atau berita.**

Laporan berupa pengumuman, *press release* atau berita ini ditujukan kepada masyarakat umum, isinya tentang sedang atau telah dilakukannya evaluasi. Di dalamnya memuat alasan, tujuan, dan berbagai hal yang dianggap penting tentang evaluasi. Laporan jenis ini dilakukan jika program yang dievaluasi merupakan program yang sangat menarik dan mendapat perhatian luas dari masyarakat.

**2. Laporan kemajuan dan laporan sementara.**

Jenis laporan ini menyajikan hal-hal penting yang terjadi dalam evaluasi pada kurun waktu tertentu. Hal-hal penting tersebut berkenaan dengan temuan-temuan sementara, kemajuan yang dicapai, perubahan rencana, dan hasil-hasil kegiatan evaluasi lainnya yang bersifat sementara. Penyampaiannya bisa dalam bentuk memorandum, dan hanya ditujukan bagi kalangan yang sangat terbatas, yang berkepentingan langsung dengan kegiatan evaluasi.

**3. Laporan akhir.**

Laporan akhir adalah laporan lengkap yang berisi informasi yang komprehensi tentang tujuan, objek, kegiatan, temuan-temuan dan kesimpulan evaluasi. Laporan ini juga memuat rekomendasi yang dibuat oleh evaluator atas program yang dievaluasi.

**4. Laporan intern.**

Laporan intern adalah laporan yang dimaksudkan untuk memenuhi keperluan intern dan biasanya ditujukan kepada staf evaluasi atau partisipan. Laporan intern ini berisikan dengan hal-hal khusus yang berkaitan dengan proses atau pelaksanaan kegiatan evaluasi. Laporan ini dibuat untuk konsumsi audien yang khusus pula yaitu pihak-pihak yang memiliki otoritas atas program yang dievaluasi, sehingga memperlancar proses evaluasi yang sedang berlangsung.

Selanjutnya menurut Arikunto dan Jabar (2009:211-213) jenis laporan dapat dibedakan atas 5 (lima) jenis. Kelima jenis laporan tersebut dipaparkan berikut ini:

### **1. Laporan untuk memberi keterangan.**

Laporan jenis ini dibedakan dalam laporan berkala dan laporan khusus. Laporan berkala dibuat untuk periode waktu tertentu. Ada laporan berkala yang dibuat setiap bulan, setiap triwulan, setiap semester, setiap tahun dan sebagainya. Keterangan yang dimasukkan ke dalam laporan bersifat rutin. Bentuk serta susunannya biasanya telah ditentukan.

Sedangkan laporan khusus dapat menyajikan hasil-hasil pengujian, percobaan, atau pemeriksaan. Laporan ini dapat juga menyampaikan sesuatu yang berhubungan dengan jalannya suatu program. Misalnya jika pada suatu kabupaten dijumpai kondisi anak usia sekolah ternyata sudah tidak tertampung semua di sekolah negeri, sehingga mengharuskan peninjauan kembali pembangunan ruang belajar baru atau mendirikan sekolah baru maka semua itu perlu dilaporkan.

### **2. Laporan untuk memulai suatu tindakan atau pekerjaan.**

Pusat perhatian pada jenis laporan ini adalah tindakan atau pekerjaan dan dengan alasan apa tindakan atau pekerjaan itu dilakukan. Kedua hal tersebut dapat ditampilkan pada saran yang tercantum dalam laporan. Laporan jenis ini harus tegas, jelas, dan terperinci. Tekanannya diberikan pada aspek apa, bagaimana, siapa, bilamana dan di mana.

### **3. Laporan untuk mengoordinasi proyek.**

Mengordinasikan sesuatu memiliki arti mengatur atau menempatkan sesuatu pada tempat atau susunan yang sebaik-baiknya atau wajar. Mengkoordinasi memerlukan keterangan mutakhir, semua itu mesti dikemukakan secara jelas, tetapi padat. Hanya pokok yang berhubungan dengan tindakan yang harus dikoordinasi saja yang perlu dimasukkan dalam laporan, selebihnya ditinggalkan. Misalnya, laporan untuk mengoordinasi guru dalam satu propinsi. Untuk itu diperlukan data menurut kenyataan dan perkiraan yang dapat diandalkan. Dengan demikian, pemanfaatan guru dan pengangkatan guru dapat dilakukan dengan sebaik-baiknya. Jelas untuk laporan ini faktor waktu sangat penting, keterlambatan dalam penyampaian data mutakhir dapat menyebabkan kekeliruan dalam perkiraan yang akibatnya merugikan proyek.

### **4. Laporan untuk menyarankan suatu langkah atau tindakan.**

Bentuk laporan yang termasuk dalam jenis ini adalah laporan pelaksanaan (*feasibility report*). Pada dasarnya laporan ini mengemukakan pendapat,

apakah suatu langkah dibenarkan atau tidak. Langkah di sini tidak harus berupa hal yang besar tetapi permasalahan kecil dapat dibuat laporannya. Saran dapat juga menyangkut sesuatu hal yang sudah berjalan, namun memerlukan perluasan.

Hal yang diinginkan pihak pemakai laporan jenis ini adalah langkah atau tindakan apa yang harus dilakukan, mengapa harus dilakukan, apa manfaat yang akan diperoleh, dan berapa biayanya jika pada tindakan itu memang diperlukan pembiayaan. Pemakai laporan tentu ingin tahu, misalnya apakah hasil saran itu sudah pasti ataukah hanya dugaan, hal seperti itu berisiko, jika saran itu diterima atau ditolak.

## **5. Laporan untuk merekam kegiatan.**

Jenis laporan ini adalah laporan kemajuan dan laporan akhir. Laporan kemajuan dapat dibuat menurut kebutuhan, ada yang setiap bulan, setiap triwulan, atau setiap semester. Untuk proyek jangka pendek memerlukan laporan kemajuan yang dibuat setiap minggu atau setiap dua minggu. Selain itu, ada juga laporan kemajuan yang tanpa kaitan langsung dengan jangka waktu, penekanannya pada volume pekerjaan yang telah terselesaikan yaitu tergantung pada jangka waktu yang tercakup pada laporan tersebut. Di dalamnya disajikan semua kegiatan selama masa pelaporan, termasuk semua perincian yang perlu dikemukakan sebagai konsekuensi pelaksanaan pekerjaan. Dilihat dari segi penyampaian informasinya, laporan kemajuan yang dikeluarkan secara teratur, pada hakikatnya termasuk laporan berkala.

Laporan akhir berbeda dengan laporan kemajuan. Laporan akhir merangkum segala segi pekerjaan setelah semuanya selesai. Cara membahasnya pun berbeda, setiap bagian perlu disesuaikan dalam susunan baru dan segala sesuatunya diintegrasikan dalam cakupan yang menyeluruh, dilihat dari pandangan bahwa semua itu telah lewat. Meskipun demikian, laporan akhir tidak terlepas dari laporan kemajuan, penyusunannya mengacu kepada laporan kemajuan.

Jenis-jenis laporan yang disampaikan di atas dapat berbentuk laporan evaluasi tertulis maupun laporan evaluasi lisan. Berikut penjelasannya:

1. Laporan evaluasi tertulis.
2. Laporan evaluasi lisan.

Laporan lisan yang didukung oleh alat bantu visual yang sesuai lebih efektif untuk pelaporan evaluasi. Audiens yang mendengarkan laporan

lisan memerlukan pengenalan yang menjelaskan tujuan prosedur evaluasi yang digunakan. Menyajikan kesimpulan positif diikuti dengan kesimpulan negatif dan rekomendasi, dan evaluator hendaknya memperhatikan point-point sebagai berikut:

- a. Akurasi, keseimbangan, dan kewajaran.
- b. Komunikasi dan persuasi.
- c. Tingkat rincian.
- d. Penggunaan bahasa sederhana, terarah, tepat dan menarik.
- e. Hindari jargon dan bahasa teknis yang tidak perlu.
- f. Gunakan contoh-contoh, anekdot, dan ilustrasi-ilustrasi.
- g. Sensitif terhadap hak-hak dan perasaan pihak-pihak yang terlibat dalam evaluasi program.

Laporan lisan juga memerlukan perhatian khusus dalam menggunakan audiovisual untuk presentasi informasi. Tak pelak lagi bahwa saran yang sering diajukan dalam berbicara dan berkomunikasi serta teks-teks relevan diperlukan dalam laporan lisan. Berikut beberapa tips terkait untuk membuat laporan evaluasi lisan yang efektif:

- a. Menyiapkan garis besar, berurutan dan bertautan ketika berkomunikasi dengan audiens, sehingga pemaparannya nantinya dapat berjalan dengan baik dan audiens dapat memahami apa yang dipaparkan.
- b. Memutuskan siapa yang akan menyampaikan informasi. Ketua evaluator menentukan siapa dari tim evaluator yang akan menyampaikan informasi laporan evaluasi, jika dalam hal ini ketua evaluator memiliki kemampuan berkomunikasi yang baik maka ketua tim tersebut adalah pilihan terbaik. Namun demikian dapat saja menggunakan anggota tim atau menggunakan pihak lain di luar tim evaluasi jika dirasakan tim evaluasi tidak memiliki kemampuan berkomunikasi yang baik karena apabila dipaksakan akan merusak presentasi evaluasi. Namun demikian apabila menggunakan pihak lain sebaiknya ketua tim evaluator dapat terlibat dalam sekurang-kurangnya pada satu bagian dari presentasi sehingga kekakuan dapat dikurangi.
- c. Pilih media dalam menyampaikan laporan lisan (narasi verbal, videotape, debat dan lainnya). Buatlah presentasi menarik dan beragam format penggunaan berbagai media, beberapa presenter, atau variasi lainnya. Melalui hal tersebut dapat mempertahankan perhatian audiens terhadap jalannya pemaparan yang dilakukan.

- d. Membuat visual yang menyertai presentasi. Namun demikian diperhatikan agar visual yang digunakan tidak mendominasi presentasi sehingga tidak menjadi efektif. Penggunaan *power point* yang didominasi oleh penggunaan *numbering* dan *bullets* terkadang membuat tidak efektifnya presentasi. Penggunaan *power point* yang tepat dengan menampilkan grafis kreatif, gambar, *flowcharts*, atau menyisipkan kartun dan warna dapat membangkitkan perhatian terhadap temuan-temuan evaluasi yang disampaikan.
- e. Mengembangkan presentasi yang natural dan menyentuh perasaan dan tentunya nyaman digunakan akan memberikan kemanfaatan kepada penyampai pesan. Dalam hal ini dapat digunakan *pointer laser* atau *computer stylus*. Namun perlu diperhatikan dalam menggunakan *pointer laser* jangan sampai menyentuh mata klien karena hal ini dapat mengganggu jalannya pemaparan.
- f. Melibatkan audiens dalam presentasi melalui tanya-jawab, tunjuk tangan atau interaksi lainnya dengan memberikan kesempatan kepada audiens selama beberapa menit untuk berbicara bertiga atau berpasangan untuk mengenali masalah.
- g. Mengembangkan dan mengikuti agenda tanpa mengganggu waktu istirahat audiens. Laporan lisan yang berlarut-larut akan membuat orang-orang akan meninggalkan ruang presentasi.

## D. Faktor-faktor Dalam Perencanaan Laporan

Faktor-faktor penting yang harus diperhatikan dalam perencanaan laporan evaluasi dipaparkan oleh Fitzpatrick dkk (2004) yaitu:

### 1. *Tailoring reports to their audiences* (mengenal audiens laporan evaluasi).

Mengenal audiens merupakan faktor penting yang harus diperhatikan evaluator dalam menulis laporan evaluasi. Hal ini dikarenakan audiens memiliki pengetahuan dan kebutuhan informasi yang berbeda-beda, untuk itu evaluator dalam menulis laporan evaluasi haruslah memperhatikannya sehingga laporan evaluasi tersebut dapat diterima audiens karena komunikatif.



## **2. *Tailoring reports content to the evaluation's audiences* (mengenal konten laporan evaluasi).**

Evaluators harus memahami bahwa audiens bermacam-macam latar belakang, kepentingan, preferensi, dan motivasi, bahkan audiens menerima dan menggunakan laporan evaluasi secara berbeda pula. Oleh karena itu evaluator dalam membuat laporan evaluasi berisikan konten-konten yang tidak mengabaikan perbedaan audiens tersebut sehingga nantinya laporan evaluasi tersebut tidak diabaikan audiens.

Berkaitan dengan faktor ini, Fitzpatrick dkk (2004) menegaskan bahwa seorang evaluator harus memperhitungkan perbedaan audiens dalam menafsirkan dan menerima laporan evaluasi. Evaluator harus memahami jika ada pihak yang dapat menemukan informasi laporan tersebut di samping ada yang tidak dapat menerimanya. Ditegaskan lebih lanjut oleh Fitzpatrick dkk (2004) bahwa terkait dengan faktor ini dapat dirujuk penelitian yang dilakukan Weiss dan Bucuvalas (1980) bahwa laporan evaluasi harus memenuhi dua kriteria yaitu nilai-nilai kebenaran dan nilai-nilai kebermanfaatan.

## **3. *Tailoring report format, style, and language to the evaluation audiences* (mengenal format laporan, style, dan bahasa).**

Laporan evaluasi tidak hanya dalam bentuk dokumen tertulis, saat ini laporan evaluasi dapat ditulis atau disampaikan dalam berbagai alternatif format seperti *videotape*, *website*, *power point*. Selain faktor format yang menjadi perhatian adalah penyesuaian bahasa dan tingkat kecanggihannya sehingga hasil/temuan evaluasi dapat dipahami dengan jelas oleh audiens.

Cousin dan Leithwood (1986) memaparkan bahwa kombinasi lisan dan tulisan dalam laporan evaluasi mempunyai dampak kepada audiens sehingga audiens memiliki apresiasi yang besar terhadap hasil/temuan evaluasi. Berbagai media dan model ditampilkan dalam pelaporan evaluasi diantaranya adalah: laporan tertulis, esai foto, laporan pita rekaman, slide presentasi, film, video tape, presentasi multimedia, dialog-dialog, pendapat atau praktek ujian, tampilan produk, simulasi, skenario, gambar, studi kasus, grafik dan peta, ringkasan skor tes dan tanya-jawab.

Selain media dan model tersebut, maka saat ini penggunaan *e-mail* merupakan dimensi terbaru dalam pelaporan evaluasi. Laporan evaluasi dapat dikomunikasikan secara cepat kepada berbagai audiens. Terkait penggunaan format, *style*, bahasa dalam laporan evaluasi maka Patton (1986)

menegaskan bahwa hendaknya evaluator membahas dan merundingkannya dengan audiens sehingga laporan evaluasi nantinya sesuai dengan kebutuhan audiens.

Bahkan Brinkerhoff dkk (1983) menjelaskan terkait dengan laporan evaluasi, audiens dapat menyarankan kepada evaluator beberapa hal terkait dengan laporan evaluasi yaitu: (1) catatan saran tentang informasi yang diperlukan, (2) informasi-informasi yang menjadi interest, (3) rekomendasi-rekomendasi tertentu, dan (4) tampilan grafik.

#### **4. *Timing of evaluation reports (waktu pelaporan evaluasi).***

Waktu merupakan hal yang amat penting diperhatikan apabila informasi akan dipakai. Serahkan hasil/temuan evaluasi sesuai dengan waktu dan pola kerja pembuat keputusan. Apabila informasi datang terlalu dini, pengaruhnya mungkin hilang, dan kalau terlambat jelas tak akan ada gunanya. Di samping itu penting juga bagi evaluator untuk tidak menyerahkan data prematur sebelum data cukup terkumpul untuk menolong pengambilan keputusan yang tepat.

Dalam hal ini Cousin dan Leithwoods (1966) menyarankan agar penggunaan hasil evaluasi disempurnakan dengan melakukan komunikasi berkelanjutan dan atau melakukan pendekatan antara evaluator dan pembuat kebijakan.

Terkait dengan faktor pemilihan waktu pelaporan, Fitzpatrick dkk (2004) memaparkan tiga hal yaitu:

##### **a. *Scheduled interim report* (laporan interim terjadwal).**

Selama perencanaan dan pelaksanaan evaluasi, evaluator melakukan penjadwalan untuk bertemu dengan pihak-pihak terkait untuk menyampaikan hasil/temuan dan reaksinya. Pertemuan yang dijadwalkan secara teratur dan berkala dapat membantu persiapan laporan akhir dan memberikan kesempatan kepada evaluator untuk mendalami persepsi audiens terhadap hasil/temuan, perubahan sikap, meningkatkan kredibilitas evaluasi dan evaluator serta meningkatkan pengaruh evaluasi. Laporan dapat dijadwalkan pada pencapaian terbaik dari evaluasi atau program atau secara berkala terkait dengan pertemuan rutin kecil atau *stakeholder*.

##### **b. *Unscheduled interim reports* (laporan interim tak terjadwal).**

Kebutuhan untuk laporan evaluasi interim tidak selalu dapat dilihat sebelumnya, namun akan ada waktu tambahan apabila tersedia informasi yang harus digunakan secara bersama. Informasi yang diperoleh dari

pertemuan yang tak terjadwal membantu evaluator dalam membuat laporan interim.

## **E. Format Laporan Evaluasi**

Tidak ada satu kerangka penulisan laporan terbaik yang tepat untuk semua jenis laporan evaluasi tertulis. Hal ini dikarenakan peran, objek, dan konteks evaluasi yang beragam, sehingga masing-masing berisikan kerangka dengan kekhasannya tersendiri, dan laporan-laporan tersebut dirancang untuk mencerminkan kekhasan tersebut.

Namun demikian ada beberapa item penting yang harus terdapat dalam setiap laporan evaluasi tertulis baik untuk laporan evaluasi final maupun laporan intern. Menurut Arikunto dan Jabar (2009:205-209) detail susunan/organisasi laporan evaluasi sebagai berikut:

### **1. Ringkasan Eksekutif.**

Ringkasan eksekutif dituntut dapat memberikan informasi lugas sehingga dapat cepat dipahami dan dipertimbangkan dalam perumusan kebijakan yang dilakukan oleh pengambil keputusan/ kebijakan. Oleh karena itu yang diperlukan adalah pokok-pokok permasalahan kebijakan dan alternatif rekomendasi kebijakannya dengan dukungan kuat dari informasi empiris yang akurat serta nilai normatif yang tajam, sehingga diketahui kelayakannya dan peluang keberhasilannya.

### **2. Pendahuluan.**

Pada bagian ini berisikan komponen-komponen penting sebagai berikut:

#### **a. Latar belakang masalah.**

Bagian ini menguraikan latar belakang empiris misalnya berupa kasus aktual, konseptual ideologis, atau keduanya, untuk menunjukkan adanya permasalahan evaluasi. Tidak setiap permasalahan pendidikan adalah masalah evaluasi. Kriteria permasalahan evaluasi yang paling menonjol adalah adanya orientasi ke arah pencarian alternatif untuk melakukan peningkatan atau perbaikan sistem yang ada.

#### **b. Rumusan masalah.**

Permasalahan evaluasi dirumuskan sedemikian rupa sehingga mencerminkan misi pencarian alternatif rekomendasi yang dibutuhkan untuk melakukan perbaikan.

c. Tujuan evaluasi.

Tujuan evaluasi menggambarkan berbagai alternatif rekomendasi kebijakan yang diperlukan dan layak untuk memecahkan permasalahan kebijakan.

d. Manfaat evaluasi.

Oleh karena evaluasi program berorientasi pada pencarian alternatif rekomendasi kebijakan maka perumusan manfaat evaluasi harus dapat menunjukkan sasaran strategi yang menjadi pusat perhatian evaluasi program.

e. Batasan pengertian.

Apabila diperlukan maka evaluator dapat menyertakan beberapa batasan pengertian konsep kunci dari kegiatan evaluasi yang dilakukan.

### **3. Kajian Pustaka.**

Kajian pustaka diperlukan untuk mempertajam permasalahan evaluasi, mendasari pengembangan strategi, rancangan dan model evaluasi, mendasari instrumentasi dan penafsiran makna dari data yang akan diperoleh dan mendasari analisis dan perumusan alternatif kebijakan. Oleh karena itu, kajian pustaka hendaknya dapat menunjukkan kebijakan dan peraturan yang menjadi konteks permasalahan evaluasi, keluasan dan kedalaman konsep yang mendasari evaluasi, serta informasi empiris untuk mendukung argumentasi yang dikembangkan dalam kegiatan evaluasi tersebut.

### **4. Metodologi Evaluasi.**

Komponen penting dalam laporan evaluasi adalah terkait dengan metodologi yang meliputi:

a. Cakupan wilayah evaluasi.

Pada bagian ini dipaparkan pembatasan cakupan seberapa jauh dapat diberlakukannya temuan evaluasi dan alternatif rekomendasinya.

b. Rancangan evaluasi.

Bagian ini dipaparkan rancangan yang dimaksudkan dalam melakukan evaluasi beserta penjelasannya. Dalam hal ini kegiatan evaluasi sangat dimungkinkan diterapkannya berbagai pendekatan evaluasi seperti eksploratori, eksplanatori, deskripsi dan sebagainya bergantung pada peran evaluasi dalam proses kebijakan. Evaluasi dapat dilakukan dengan maksud untuk menjadi dasar perumusan kebijakan untuk menunjang

implementasi kebijakan, atau untuk mengetahui kinerja dan dampak dari kebijakan.

c. Pengumpulan data.

Data evaluasi dapat berada pada berbagai sumber data, karenanya tidak tertutup kemungkinan suatu evaluasi menggunakan berbagai metode dan alat pengumpulan data. Dalam rancangan harus jelas data apa saja yang diperlukan dan/atau dikumpulkan, masing-masing perlu jelas sumber data, metode serta instrumen pengumpul datanya. Keterkaitan antar jenis data satu dengan lainnya dapat ditata dalam suatu kerangka sistemik yang diturunkan berdasarkan kajian teoretis. Alat pengumpul data dalam evaluasi harus dapat menjamin bahwa informasi yang dihasilkan sah dan andal, sehingga dapat menjadi dasar untuk perumusan alternatif rekomendasi kebijakan.

d. Triangulasi.

Triangulasi merupakan suatu cara memandang permasalahan/ objek yang dievaluasi dari berbagai sudut pandang, serta dapat dipandang dari banyak metode atau sumber data. Tujuannya agar dapat melihat objek evaluasi dari semua sisi.

Triangulasi antar metode dan antar sumber data dilakukan untuk mengejar kualitas data yang dapat dipertanggungjawabkan. Dalam kaitannya dengan luasnya data dan sumber data, maka suatu evaluasi program juga dimungkinkan dilakukan secara sampling. Untuk itu, harus jelas batas-batas populasi dan prosedur samplingnya, serta menggunakan kaidah sampling yang tepat.

e. Analisis data.

Pada bagian ini dipaparkan cara melakukan analisis data yaitu analisis untuk menghasilkan kesimpulan atas data empiris dan analisis untuk menghasilkan alternatif rekomendasi kebijakan. Analisis yang pertama untuk menemukan apa yang perlu direkomendasi, sedangkan analisis kedua menjadi dasar untuk merumuskan alternatif rekomendasi kebijakan yang operasional.

## **5. Hasil Evaluasi.**

Bagian ini memuat tiga komponen utama yaitu:

a. Dekripsi data.

Pada bagian ini dipaparkan secara singkat konteks kelembagaan dan

karakteristik lain tentang konteks dari evaluasi program yang dilakukan. Pemaparan juga dilengkapi dengan sajian deskriptif dari masing-masing ubahan pokok yang menjadi fokus evaluasi program.

b. Analisis data dan pembahasan.

Analisis data sangat ditentukan oleh sifatnya evaluasi. Untuk evaluasi formulasi kebijakan akan dilakukan prakiraan kondisi yang memerlukan kebijakan alternatif, untuk evaluasi implementasi kebijakan akan banyak dilakukan eksplanasi fenomena yang memerlukan optimasi, sedangkan untuk evaluasi hasil kebijakan akan banyak dilakukan evaluasi kinerja. Hasil serta dampak yang positif maupun negatif diupayakan sustainabilitasnya. Bagian ini dengan sendirinya akan secara fungsi melibatkan sejumlah ubahan lain dalam rangka melakukan prediksi, eksplanasi atau evaluasi. Pembahasan merupakan upaya untuk memaknai semua temuan hasil analisis data, dari berbagai perspektif seperti teoretis, teknis, legalistik, sosial kultural dan sebagainya.

c. Analisis rekomendasi.

Sifat perumusan kebijakan adalah memerlukan kriteria ganda seperti untuk peningkatan efektivitas, efisiensi, pemerataan, edukasi, serta kepekaan terhadap belum terpenuhinya kebutuhan ataupun belum termanfaatkannya kesempatan. Rekomendasi bersifat prospektif, memandang informasi empiris sama pentingnya dengan nilai formatif. Oleh karena itu, bagian ini mencerminkan seberapa sarat nilai dan informasi dari suatu evaluasi. Pengajuan setiap alternatif kebijakan dituntut dapat mengindentifikasikan dan menggunakan tujuan, konsekuensi biaya, kendala, dampak lanjutan atau sampingan, waktu, resiko, dan juga peluang keberhasilan.

## **6. Kesimpulan dan Rekomendasi.**

Pada bagian ini dipaparkan kesimpulan yang diperoleh dari analisis data dan alternatif rekomendasi yang dirumuskan berdasarkan analisis rekomendasi.

## **7. Daftar Pustaka.**

Daftar pustaka disusun sesuai dengan bahan acuan yang digunakan dalam evaluasi, baik mengenai substansi isi maupun metodologi evaluasi.

Menurut Tayibnapis (2000:158-172) menyatakan *outline* laporan evaluasi terdiri dari:

**a. Cover Depan.**

Cover depan atau kulit laporan berisikan informasi sebagai berikut:

- Judul program dan lokasinya.
- Nama evaluator.
- Periode waktu yang dilalui.
- Tanggal laporan diserahkan.

**b. Bab 1 Ringkasan.**

Ringkasan atau *executive summary* yang berisi laporan pendahuluan evaluasi, menerangkan mengapa evaluasi dilakukan, memuat kesimpulan, dan saran-saran.

Ringkasan dibuat untuk orang-orang yang tidak mempunyai waktu panjang untuk membaca laporan evaluasi program secara utuh, oleh sebab itu, sebaiknya ringkasan ini tidak lebih dari dua atau tiga halaman. Walaupun ringkasan dalam laporan evaluasi diletakkan di depan, tetapi ringkasan ini dibuat paling akhir.

Apabila masih diperlukan tambahan informasi sebagai berikut:

- Apakah ada keputusan yang akan dibuat berdasarkan hasil evaluasi? Kalau ada, apa keputusan tersebut?
- Kepada siapa laporan akan diberikan?
- Siapa-siapa yang berminat atas laporan tersebut?
- Apa kesulitan-kesulitan yang dihadapi dalam proses evaluasi?

**c. Bab 2 Latar Belakang Evaluasi.**

Pada bagian ini dipaparkan asal mula mengapa program dibuat dan apa yang harus dilakukan. Sejauhmana informasi tergantung kepada penerima hasil evaluasi (untuk siapa evaluasi dibuat). Kalau audiensi tidak mengetahui apa-apa tentang program, maka harus diterangkan dengan lengkap. Kalau evaluasi hanya diperuntukkan bagi keperluan intern dan pembaca sudah banyak mengetahui program maka bagian ini dapat dipersingkat, informasi dibuat untuk direkam.

Apabila laporan yang dibuat merupakan satu-satunya laporan untuk program maka laporan harus berisi penjelasan rinci tentang tujuan umum, tujuan khusus, dan komponen-komponen penting lainnya. Penjelasan awal tentang program diuraikan ketika evaluasi direncanakan secepatnya sesudah mengetahui peran sebagai evaluator, atau masukkan ke dalam usulan evaluasi. Informasi untuk bagian ini dapat diperoleh dari orang-orang program, catatan-catatan hasil rapat, memo, *outline kurikulum*, daftar tujuan umum, perkiraan anggaran, dan sebagainya.

Bagian khusus dalam bab ini adalah:

1. Hakikat program.

- Di mana program dikerjakan? Pada masyarakat atau kelompok yang bagaimana? Siapa dan berapa banyak orang yang dipengaruhi?
- Berapa jumlah orang yang ikut berpartisipasi? Bagaimana pengolahan peserta? Sekolah, kelas atau individu?
- Bagaimana program dimulai?
- Apakah dilakukan *need assessment* formal atau informal, apabila ada, bagaimana hasilnya?
- Apa dorongan utama program dilaksanakan atau apa latar belakang pendidikan program? Karena kebutuhan masyarakat, daerah atau nasional? Bagaimana kesempatan untuk memperoleh dana? Bagaimana inisiatif staf/karyawan?

2. Tujuan umum program.

- Apa yang dicapai desain program?
- Apa tujuan umum dan tujuan khusus yang dirumuskan?
- Apa prioritas yang diutamakan apabila ada?
- Khusus untuk evaluasi kualitatif, apakah ada tujuan umum yang tidak disebutkan, atau akibat-akibat yang tak dapat diperkirakan yang diidentifikasi oleh evaluator?

3. Klien yang terlibat dalam program.

- Apa ciri-ciri klien program yang dituju misalnya umur, latar belakang, ekonomi/pendidikan, pengalaman, kebutuhan khusus, atau tingkat kemampuannya?
- Atas dasar apa peserta program dipilih?
- Bagaimana bentuk penampilan program yang diinginkan?



- Apakah peserta tinggal di dalam program sesuai dengan waktu yang direncanakan? Apabila tidak, kriteria apa yang menentukan masuk atau keluarnya?
4. Ciri-ciri, materi, kegiatan, dan persiapan administrasi program.
- Apa materi yang dipakai dan bagaimana? Apakah harus dibeli atau dibuat?
  - Apa sumber-sumber program? Berapa dana, sarana fisik, transportasi yang harus ada, dan siapa yang menyediakan?
  - Pada kegiatan-kegiatan apa, peserta program diharapkan mengambil bagian?
  - Apa prosedur khusus yang harus diikuti?
  - Apa rasional yang mendasari program? Mengapa perencana program merasa bahwa materi dan kegiatan program akan mengarah kepada pencapaian tujuan umum program?
  - Sampai sejauhmana program diimplementasikan? Apa perbedaan yang diharapkan dari waktu ke waktu?
  - Bagaimana program dikelola dan dijalankan? Apa yang dikembangkan atau diperlukan? Siapa yang melakukannya?
5. Karyawan dan orang lain yang terlibat dalam program.
- Berapa banyak personal khusus seperti administrator, konsultan, sekretaris, spesialis, sukarelawan, dan lain-lain yang aktif dalam program? Apa proses yang mereka lakukan?
  - Apakah mereka memerlukan latihan khusus sebelum atau selama program?
  - Berapa waktu yang diberikan mereka kepada program?

#### **d. Bab 3 Penjelasan tentang Apa yang Dievaluasi.**

Bab ini memaparkan dan membatasi ruang lingkup yang diterima evaluator, yaitu menjelaskan mengapa evaluasi dilakukan, apa yang akan dicapai, dan apa yang tidak. Siapkan penjelasan tentang maksud dan tujuan evaluasi segera setelah menerima pekerjaan sebagai evaluator. Konsep untuk itu harus disetujui oleh semua orang yang berminat dan harus disimpan di dalam arsip.

Bagian selanjutnya menerangkan tentang metodologi evaluasi. Bagaimana program dievaluasi, penting dijelaskan sampai rinci supaya kesimpulan diterima dan dipercaya orang lain. Pemakai tentu ingin mengetahui bagaimana informasi diperoleh. Namun buatlah uraian tentang hal-hal yang teknis secara komprehensif sedemikian rupa sehingga dapat terbaca dan dimengerti oleh rata-rata pembaca.

Isi bab ini berupa:

1. Tujuan evaluasi.

- Siapa yang meminta evaluasi?
- Apa evaluasi yang dilakukan, formatif atau sumatif?
- Kalau evaluasi dilakukan untuk audiensi khusus, siapa mereka, apakah mereka karyawan program, legislator, kelompok masyarakat, dewan direksi atau orang tua?
- Informasi apa yang diperlukan audiensi?
- Apakah evaluasi dilakukan untuk pengambilan keputusan? Bila ya, siapa pemegang keputusan ini? Apa hakikat keputusan yang akan diambil? Keputusan mungkin berupa pilihan untuk melanjutkan atau tidak melanjutkan program atau mengubah, dan memperbaiki program seluruh atau bagian-bagian tertentu dari program?
- Apakah evaluasi bertujuan menjawab pertanyaan penelitian? Kalau ya, nyatakan dalam bab ini.
- Di mana evaluasi dilakukan? Apakah ada hambatan dari segi waktu, atau biaya dalam melakukan evaluasi? Apakah ada isu-isu khusus yang tidak diinginkan atau tidak disetujui evaluatornya?

2. Desain evaluasi.

- Apakah satu desain evaluasi menjadi dasar bagi semua evaluasi? Adakah desain lain yang ditambahkan? Atau kombinasi antara pendekatan kuantitatif dan kualitatif?
- Mengapa desain ini yang dipilih?
- Apa keterbatasan desain yang dipakai?
- Apa kontaminasi dan *confounding* yang diperkirakan?
- Adakah kondisi tertentu menghalangi pemakaian desain yang lebih tepat?
- Apa alasan yang paling kuat anda memilih desain dan metode itu?

3. Pengukuran hasil (*outcome*).

Bagaimana instrumen atau pendekatan yang dipakai.

- Apa hasil program, akibat atau kegiatan yang diukur, dijelaskan, atau diobservasi? Apakah hasil atau kegiatan tersebut diterakan dalam tujuan umum, atau apakah pilihan pengukuran berdasarkan alasan lain?
- Untuk setiap hasil yang diminati, data apa yang dikumpulkan/ apa instrumen yang dipakai?
- Apakah instrumen dikembangkan oleh program atau dibeli? Bila dikembangkan bagaimana? Bila dibeli berdasarkan pertimbangan apa?
- Bagaimana instrumen dites reliabilitas, validitas, ketepatan, dan relevansinya terhadap program?

4. Prosedur pengumpulan data.

- Bagaimana prosedur pengumpulan? Kapan instrumen dipakai, observasi atau wawancara dilakukan, dan siapa yang mengumpulkan data? Hal ini dapat dibuat dalam tabel. Bila perlu perhatikan kualifikasi orang-orang yang mengumpulkan data.
- Apakah penataran diberikan kepada mereka yang melakukan berbagai pengukuran atau observasi? Apabila ya, penataran apa?
- Apakah setiap peserta dalam setiap kelompok diukur atau diobservasi, atau apakah prosedur sampling dipakai? Apa rasionalisasi yang mendasari prosedur sampling?

5. Implementasi pengukuran.

Bagaimana instrumen dan pendekatan pengumpulan data yang dipakai.

- Mengapa implementasi dijelaskan? Untuk pertanggungjawaban atau untuk melengkapi proposal, rencana, atau hanya untuk menjelaskan apa yang terjadi dan yang telah dilakukan?
- Aspek penting apa dalam program yang diobservasi dan tidak diukur. Apa alasan pilihan itu, mengapa bukan yang lain?
- Bagaimana data penunjang dikumpulkan? Dengan membaca dokumen program, observasi, wawancara formal atau percakapan dengan karyawan? Atau adakah metode pengumpulan data lain yang dipakai?
- Bagaimana instrumen dikembangkan? Dengan pinjam atau beli? Dari mana dan atas dasar apa, mengapa?

- Apa keterbatasan-keterbatasan pada instrumen atau pada alat ukur lain yang dipakai?
  - Apakah validitas dan reliabilitas sudah diukur sehubungan dengan situasi dan kondisi program?
6. Prosedur pengumpulan data.
- Bagaimanakah jadwal pengumpulan data, siapa yang melakukannya? Untuk menjadwalkannya biasanya memakai tabel.
  - Apakah ada pelatihan yang diberikan? Apa alasan pemakaian atau alat pengumpul data lainnya?
  - Apakah sampel yang diukur representatif? Apakah instrumen diberikan kepada semua orang atau pada anggota representatif tertentu?
  - Apa hambatan-hambatan dan kesulitan-kesulitan yang dihadapi dalam prosedur yang dipakai untuk mengukur implementasi?

#### **e. Bab 4 Hasil Evaluasi.**

Bab ini memaparkan bermacam-macam pengukuran, observasi dan lain-lain metode pengumpulan data yang dijelaskan pada bab sebelumnya. Sebelum menulis bab ini, data sudah harus dianalisis, direkam dalam tabel, grafik dan telah diukur signifikansinya. Nilai tes biasanya disajikan dalam grafik dan tabel yang memperlihatkan *mean* dan standar deviasi setiap kelompok. Hasil kuesioner ditabulasi. Hakikat evaluasi adalah penggunaan metode pengumpulan data kualitatif seperti *in-depth*, *open ended interview*, observasi langsung atau studi kasus. Informasi yang kaya dan rinci harus diatur atau difokuskan sesuai dengan bidangnya dan dikategorikan.

Bab ini berisikan:

1. Hasil studi implementasi.
  - Apakah orang-orang program memberikan hasil seperti yang direncanakan?

Apakah program diimplementasikan sesuai rencana seperti yang diharapkan pemakai? Bila tidak, apa yang terjadi? Adakah komponen yang dihapus. Atau diubah? Apakah semua materi tersedia dan apakah semua dipakai? Apakah program diberikan kepada peserta yang telah ditentukan? Apakah semua kegiatan penting dilakukan?

- Jelaskan secara rinci mungkin keadaan yang sebenarnya (penampilan program) dan tabel-tabel yang biasa digunakan. Semua kegiatan dijelaskan secara serinci mungkin karena mereka menggambarkan pengalaman khusus program. Dalam evaluasi kualitatif, penjelasan ini berupa narasi tertulis yang memberikan pandangan yang menyeluruh tentang program.
- Kalau terjadi perubahan dalam program, apa akibatnya terhadap perilaku karyawan dan peserta, efisiensi atau aspek lain dalam program?
- Apakah ada variasi yang terjadi dalam program?

2. Hasil studi *outcome*.

- Berapa orang dan siapa yang mengambil pretes?
- Berapa dari yang mengalami pretes yang masih tinggal dalam program?
- Bagaimana hasil pretes? Apakah ada perbedaan nilai pretes di antara program dan kelompok pembanding?
- Bagaimana hasil final setiap pengukuran yang dilakukan terhadap peserta program dan peserta kelompok pembanding? Bagaimana perbandingannya? Apakah secara statistik perbedaan tersebut signifikan? Tabel dan grafik dapat menggambarkan ringkasan hasil kuantitatif.
- Kalau tidak ada kelompok kontrol, bagaimana perubahan penampilan dari tes yang satu ke tes yang lain?
- Kalau tidak ada kelompok kontrol, dengan apa hasil dibandingkan untuk menilai mutunya? Misalnya, dengan tes atau dengan penampilan yang lain, sejumlah standar untuk mencerminkan tingkat kompetensi dapat dipakai untuk perbandingan. Dapatkah perbandingan digambarkan dengan grafik?
- Bagaimana hasil desain tes statistik yang menjawab pertanyaan tentang hubungan antara peserta atau program?

**f. Bab 5 Diskusi tentang Hasil Evaluasi.**

Penafsiran atau intepretasi bagi setiap hasil evaluasi ditulis dalam bab sebelumnya, yang menyajikan hasil evaluasi. Namun bila program atau evaluasi sulit diinterpretasikan (*complicated*), maka dibuat bab khusus

untuk penafsiran diskusi hasil evaluasi membuat laporan menjadi lebih jelas.

Hasil evaluasi harus didiskusikan dengan referensi khusus, dihubungkan dengan tujuan yang tersebut dalam bab sebelumnya. Bab ini berisi dua isu umum yaitu: sampai seberapa jauh kebenaran hasil evaluasi dan bagaimana suatu hasil program.

- Bagaimana hasil proyek apabila dibandingkan dengan hasil evaluasi apabila tidak ada program?
- Kalau memakai kelompok kontrol, apakah hasil program lebih baik daripada hasil kelompok kontrol? Apakah perbedaannya signifikan secara statistik? Apabila tidak signifikan apakah hasil program tampaknya berhasil? Kalau analisis kualitatif yang dipakai, evaluator harus melakukan penilaian yang substansial (nyata) tentang kelebihan dan keterbatasan bermacam-macam bagian analisis, yang didukung kuat oleh data dan juga kurang didukung.
- Apakah karyawan merasa program akan mencapai perolehan yang lebih signifikan kalau diubah (dimodifikasi) atau diberi waktu lebih lama (dalam periode waktu yang lebih lama)? Apakah bukti dalam data mendukung pendapat tersebut?
- Apa yang dianggap kelebihan dan kelemahan program? Kesimpulan dan saran-saran apa yang dapat diberikan dari penelitian korelasi atau dari observasi? Apakah program tertentu tampaknya lebih efektif dengan kelompok tertentu? Apakah sikap atau pencapaian berhubungan dengan ciri-ciri peserta?

#### **g. Bab 6. Biaya dan Manfaat.**

Dalam bab ini dipaparkan mengenai anggaran program dan bagian-bagian yang berhubungan dengan kontroversi. Menjelaskan kebenaran pendekatan tertentu dengan analisis *cost benefit*, yang telah dipakai. Menjelaskan manfaat dan menafsirkan data.

Apabila *cost benefit* merupakan fokus evaluasi sebaiknya dimasukkan dalam bab 5. Uraian *cost benefit* ini umumnya berupa daftar biaya yang dihubungkan dengan program, kemudian diperluas dengan catatan ringkasan non-rupiah sebagai biaya kualitatif. Manfaat program kemudian diuraikan dan ditimbang dengan biaya. Kalau memungkinkan, berikan dalam bagian ini tabel daftar *cost benefit*.

Bab ini berisikan:

1. Metode yang dipakai untuk menghitung *cost benefit*.

- Bagaimana anda merumuskan *cost* dan *benefit*?
- Metode apa yang dipakai untuk menghitung *cost* dan *benefit*? Apakah menggunakan rumus matematika formal? Atau memakai metode informasi yang membandingkan *cost* dan *benefit*? Apakah alasan memakai metode ini?

2. *Cost* yang dihubungkan dengan program.

Ongkos rupiah.

- Apakah ada uang ekstra diperlukan untuk melaksanakan program? Dari mana uang itu diperoleh?
- Untuk apa uang tersebut kalau tidak dipakai dalam program?
- Berapa persen dari seluruh biaya dipakai untuk operasi program? Berapa biaya yang dipakai ketika memulai program (*start up cost*)? Biaya *start up* tidak diperlukan kalau melanjutkan program. Tabel yang menunjukkan anggaran rupiah harus dimasukkan/dilampirkan.

Ongkos non rupiah.

- Apakah program membebani guru, orang tua, administrator, kesabaran, moral dan semacamnya?
- Apakah karyawan bekerja lembur karena program?
- Apakah ada pekerja sukarela dalam program? Kalau ada, tentu mereka lakukan untuk proyek?
- Apakah partisipasi dalam program menyebabkan peserta kehilangan pengalaman bekerja di bidang lain?
- Adakah biaya-biaya lain yang menyebabkan terjadinya kehilangan alternatif *opportunity*?

3. *Benefit* yang dihubungkan dengan program.

*Benefit* rupiah.

- Berapa penghasilan yang diperoleh program? Misalnya, apakah program memperoleh subsidi pemerintah untuk biaya pendidikan tertentu?

*Benefit* non rupiah.

- Apakah hasil-hasil positif yang diperoleh program?

- Sampai sejauh mana perkembangan yang diperoleh (pencapaian tujuan umum program)?
- Apakah penghasilan orang-orang program lebih baik daripada orang-orang lain di luar program untuk pekerjaan yang serupa? Bagaimana perbedaannya?
- Bagaimana pandangan orang terhadap program?
- Apakah ada *benefit* yang tidak diduga?

#### **h. Bab 7. Kesimpulan dan Saran-Saran.**

Kesimpulan:

- Apa kesimpulan umum tentang keefektifan program secara keseluruhan? Sampai seberapa jauh kebenaran kesimpulan?
- Apakah perlu ada penilaian terhadap beberapa aspek kebijaksanaan program?
- Apakah ada hal-hal yang diperkirakan terjadi sebagai akibat program, yang dipertimbangkan untuk memperoleh gambaran yang lengkap tentang *impact* program?

Saran-saran:

- Berdasarkan data khusus, apa saran-saran dan pilihan yang dapat diberikan terhadap program? Apa kelebihan-kelebihan program dan aspek-aspek apa saja yang perlu atau yang dapat dikembangkan dan diperbaiki?
- Apakah tujuan evaluasi juga memberi rekomendasi dan saran-saran pilihan? Apakah pemakai ingin mengetahui efektivitas atau keefektifan program atau apakah mereka ingin juga mengetahui kelemahan-kelemahan program?
- Perkiraan, hipotesis, atau dugaan apa yang diberikan data? Apa rekomendasi yang diberikan untuk program selanjutnya atau penelitian yang akan dilakukan?

Komponen-komponen penting yang termuat dalam laporan evaluasi dipaparkan Fitzpatrick dkk (2004) adalah:

- I. Ringkasan eksekutif
- II. Pendahuluan



- A. Tujuan evaluasi
  - B. Laporan evaluasi audiens
  - C. Batasan-batasan evaluasi dan penjelasan
  - D. Tinjauan umum tentang isi laporan
- III. Fokus evaluasi
- A. Deskripsi objek evaluasi
  - B. Pertanyaan evaluasi atau sasaran untuk fokus kajian
  - C. Informasi yang diperlukan untuk menyelesaikan proses evaluasi
- IV. Ringkasan singkat tentang rencana dan prosedur evaluasi
- V. Presentasi hasil evaluasi
- A. Ringkasan temuan evaluasi
  - B. Penafsiran penemuan evaluasi
- VI. Kesimpulan dan saran
- A. Kriteria yang digunakan untuk standar penilaian dan obyek evaluasi.
  - B. Penilaian obyek evaluasi (kekuatan dan kelemahan-kelemahan).
  - C. Rekomendasi
- VII. Laporan minoritas atau *rejoinders* (jika ada)
- VIII. Apendiks
- A. Keterangan mengenai rencana evaluasi/desain, instrumen, dan analisis data dan penafsiran.
  - B. Rincian tabulasi atau analisis data kuantitatif, dan risalah atau ringkasan dari data kualitatif.
  - C. Informasi lainnya, (jika diperlukan).

- **Ringkasan Eksekutif.**

Ringkasan eksekutif adalah informasi esensial terkait dengan tempat, presentasi tentang temuan-temuan evaluasi dan apendiks, sehingga pembaca memperoleh *point-point* penting terkait dengan laporan evaluasi. Ringkasan eksekutif menjadi penting untuk ditulis karena sebagian besar audiens evaluasi tidak memiliki waktu atau energi yang dibutuhkan untuk membaca laporan evaluasi yang tebal yang berisikan informasi, tabel atau perincian.

Untuk menarik perhatian pembaca, ringkasan dapat diletakkan di halaman depan dalam laporan dan dapat dicetak pada kertas warna yang berbeda dengan isi laporan untuk menarik perhatian. Ringkasan eksekutif bisanya

terdiri dari 2 – 6 halaman atau tergantung pada panjang cakupan dan kerumitan evaluasi.

Ringkasan eksekutif berisikan keterangan singkat mengenai: tujuan studi, pengumpulan data, temuan-temuan, penilaian dan rekomendasi, di samping itu ringkasan eksekutif dapat memuat referensi yang mengarahkan pembaca untuk informasi lebih lanjut. Ringkasan eksekutif dapat lebih dipadatkan lagi dalam bentuk abstrak eksekutif. Abstrak ini terdiri dari 1 – 2 halaman yang berisikan *point-point* temuan evaluasi yang utama.

### • **Pendahuluan.**

Pada bagian ini memaparkan rasionalitas dilakukan evaluasi. Rasionalitas evaluasi terkait dengan pertanyaan-pertanyaan mengapa evaluasi dilakukan? Apa tujuan untuk melakukan evaluasi? Pertanyaan apa yang telah diajukan untuk dicari jawabannya? Bagaimana melakukan evaluasi? Pendahuluan juga merupakan salah satu bagian logis untuk menarik perhatian pembaca tentang pembatasan evaluasi yang mempengaruhi koleksi, analisis, atau informasi penafsiran. Pembatasan secara tegas diungkapkan pada pendahuluan atau di bagian prosedur evaluasi.

Demikian pula, pemaparan tanggung jawab yang diletakkan di awal dari laporan (misalnya dalam prakata atau pada halaman judul) untuk mengklarifikasi apa yang dievaluasi atau yang tidak, sehingga melindungi klien dan evaluator dari kritik kesalahpahaman. Pendahuluan juga berguna bagi pembaca sebagai panduan singkat untuk memahami laporan evaluasi program yang dilakukan.

### • **Fokus Evaluasi.**

Fitzpatrick dkk (2004) menegaskan pentingnya menggambarkan obyek evaluasi yang akurat, sehingga semua pihak dapat menyetujui akan objek yang dievaluasi. Dalam hal ini keterangan tentang objek-objek tidak hanya terbatas pada ciri-ciri fisik tetapi termasuk di dalamnya adalah:

- a. Rasional, tujuan, dan tujuan.
- b. Klien, pelanggan, siswa, atau penerima lain/peserta untuk dimanfaatkan.
- c. Struktur program kegiatan, konten, dan karakteristik lain,
- d. Prosedur yang digunakan untuk strategi penerapan.
- e. Konteks operasi.
- f. Manusia dan kebutuhan sumber daya lain yang diperlukan.

Fokus evaluasi juga penting sebagai bagian awal dari daftar pertanyaan evaluasi, tujuan, atau agenda lain digunakan untuk fokus evaluasi, jika diferensiasi prioritas-prioritas telah ditetapkan untuk pertanyaan-pertanyaan atau sasaran, proses yang harus dijelaskan. Di samping itu fokus evaluasi berguna untuk menyertakan subbagian informasi yang diperlukan dalam menjabarkan evaluasi untuk mengumpulkan, menganalisis, dan laporan. Daftar pertanyaan dalam fokus evaluasi tersebut membantu membuat rasional untuk bagian selanjutnya lebih jelas.

- **Ringkasan singkat tentang rencana dan prosedur evaluasi.**

Laporan evaluasi dikatakan lengkap apabila menyertakan presentasi rinci terkait rencana evaluasi, instrumen pengumpulan data, metode-metode dan teknik-teknik yang digunakan untuk menganalisa dan menafsirkan data. Namun semuanya itu terkadang tidak perlu dicantumkan dalam bagian dalam laporan atau dengan kata lain dapat diletakkan pada lampiran.

- **Presentasi hasil evaluasi.**

Bagian ini berisi pemaparan hasil/temuann evaluasi, bagian ini juga merupakan sumber dari penarikan kesimpulan-kesimpulan dan pengajuan saran. Presentasi hasil/temuan evaluasi disajikan dalam bentuk ringkasan lengkap, menggunakan tabel, menampilkan, dan kutipan yang sesuai, dan merujuk data lebih detail, transkrip dalam dukungan apendiks. Walaupun sebagian audiensi melihat terlalu banyak data statistik namun banyak pengambil kebijakan, manajer, dan lain-lain merespon positif untuk data yang disajikan secara mudah dalam bentuk grafik dan peta sehingga mudah dipahami dan mudah mengambil kesimpulan.

Interpretasi hasil/temuan evaluasi sama pentingnya dengan presentasi itu sendiri, dalam hal ini menginterpretasi data tidak boleh sebagai aktivitas kebetulan, tetapi menginterpretasi data harus dilakukan dengan proses kehati-hatian sehingga mencapai penilaian dan rekomendasi yang disampaikan secara tepat.

Fitzpatrick dkk (2004) memaparkan salah satu kekurangan laporan evaluasi adalah kurang penjelasan yang membantu pembaca dalam melihat keterhubungan hasil/temuan evaluasi dengan pertanyaan utama evaluasi karena tanpa pengorganisasian atau pengkategorian, temuan-temuan sering mengaburkan, sehingga menjadi kurang dipahami. Untuk itu Fitzpatrick

dkk menyarankan evaluator untuk mengaitkan temuan-temuan evaluasi kepada pertanyaan atau fokus evaluasi.

- **Kesimpulan dan saran.**

Pada bagian ini evaluator memaparkan standar dan kriteria penilaian yang digunakan dalam melakukan evaluasi sehingga memberikan penilaian yang terbaik terkait dengan kualitas dari apa yang telah dievaluasi, dan memberikan rekomendasi. Standar dan kriteria dicantumkan secara eksplisit, karena data tidak berdiri sendiri, oleh karena itu dalam melakukan evaluasi maka evaluator dan pihak-pihak terkait untuk menetapkan standar dan kriteria dan penilaian atas data yang memenuhi penilaian objek evaluasi efektif atau tidak efektif. Hal ini menjadi urgen karena evaluasi tanpa kriteria yang jelas akan banyak mendapatkan kritikan karena kekuatan penilaian evaluasi berdasarkan pada data.

Fitzpatrick dkk (2004) menegaskan bahwa dalam sajian pada bagian ini untuk menyajikan kekuatan dan kelemahan penilaian, hal ini akan memberikan beberapa keuntungan yaitu:

- a. Perhatian dipusatkan pada penilaian positif dan negatif.
- b. Audiens dapat menemukan bagian praktis dari yang disampaikan evaluator terkait penilaian positif maupun negatif.
- c. Menghadirkan kekuatan yang membantu pihak-pihak yang bertanggung-jawab atas obyek evaluasi untuk menerima kelemahan yang terdapat setelahnya.

Pembahasan terkait dengan kekuatan dan keterbatasan memungkinkan audiens untuk melihat rasional dan rekomendasi yang disampaikan. Format lain yang bermanfaat dalam menyajikan kekuatan dan kelemahan ini adalah dengan menggunakan analisis SWOT (*strengths*/kekuatan, *weakness*/kelemahan, *opportunities*/peluang, dan *threats*/ancaman).

Selanjutnya terkait dengan rekomendasi, Fitzpatrick dkk (2004) menyarankan rekomendasi yang disampaikan kiranya mudah untuk agar mudah diaplikasikan, bahkan dalam beberapa laporan evaluasi, rekomendasi digunakan untuk memulai proses perencanaan strategis. Rekomendasi juga terkait dengan pemaparan evaluator tentang hal-hal yang telah berjalan baik dalam program dan juga hal-hal yang belum berfungsi sebagaimana mestinya dan menyarankan bagaimana melakukan perbaikannya. Dalam

kasus tertentu, terdapat rekomendasi yang disampaikan evaluator mendorong perhatian untuk memperbaiki masalah tanpa menetapkan cara yang tepat untuk mengatasi masalah tersebut.

- **Laporan minoritas atau rejoinders.**

Pada bagian ini dipaparkan (jika ada) anggota tim evaluasi tidak menyetujui atau berbeda pandangan terkait dengan hasil/temuan evaluasi, kesimpulan, atau rekomendasi, hal ini perlu dicantumkan untuk memberi ruang pemaparan dari perspektif yang berbeda sehingga melengkapi laporan evaluasi.

- **Apendiks**

Apendiks dalam laporan evaluasi merupakan pendukung yang dapat diletakkan dalam laporan evaluasi ataupun terpisah. Apendiks sangat membantu pembaca untuk mencari informasi yang dibutuhkan terkait dengan hal-hal yang diperlukan dalam memahami laporan evaluasi seperti prosedur pengambilan sampel yang digunakan, bagaimana informasi yang dikumpulkan untuk menjamin keakurasian data, ataupun statistik tertentu yang digunakan sebagai prosedur analisis.

Apendiks menyediakan metode dan informasi teknis yang diperlukan tidak hanya oleh klien utama tetapi juga oleh rekan-rekan evaluator yang akan memutuskan apakah melakukan studi tersebut sudah cukup untuk membuat hasil-hasilnya yang dapat dipercayai. Apendiks adalah tempat terbaik untuk penjelasan lengkap tentang prosedur evaluasi, data terperinci atau analisis, pengamatan, legalisasi wawancara, dan informasi lain yang relevan tetapi terlalu rumit untuk hadir dalam tubuh laporan.

Apendiks mungkin juga menyertakan data aktual instrumen koleksi dan informasi lainnya (misalnya peta batas unit pengambilan sampel di sebuah survei masyarakat) dianggap menarik dan penting untuk para pendengar tetapi tidak lengkap dan atau terlalu luas untuk dimasukkan dalam tubuh laporan. Penggunaan sesuai apendiks akan membuat laporan sendiri jauh lebih efisien dan lebih mudah dibaca.

## **F. Kemenarikan Penyajian Laporan**

Faktor kemenarikan penyajian informasi dalam penulisan laporan evaluasi penting diperhatikan karena penyajian informasi yang menarik

dapat mengurangi kejenuhan pembaca dalam membaca laporan evaluasi. Terkait dengan hal ini Fitzpatrick dkk (2004) menyarankan beberapa pertimbangan untuk membantu membuat presentasi laporan evaluasi yang dibuat evaluator menjadi menarik dan efektif yaitu:

**1. *Accuracy, balance dan fairness* (akurasi, keseimbangan, dan kesewajaran).**

Evaluator dalam menyajikan informasi dalam laporan evaluasi harus memperhatikan prinsip-prinsip keakurasian, keseimbangan dan kesewajaran. Untuk itu, evaluator harus berhati-hati dalam pengumpulan dan analisis informasi untuk kemudian dipaparkan dalam presentasi laporan tanpa ditambah-tambahi ataupun dikurang-kurangi. Semua informasi, pernyataan publik, dan laporan tertulis evaluasi berpegang teguh keterbukaan dan kelengkapan.

**2. *Communication and persuasion* (komunikasi dan persuasi).**

Komunikasi memainkan peranan penting dalam semua tahap-tahap evaluasi. Komunikasi yang baik sangat penting jika evaluator ingin memahami asal-usul dan konteks evaluasi, menimbulkan pertanyaan evaluasi dan kriteria dari stakeholder, mencapai kesepakatan dengan klien tentang rencana evaluasi, menangani aspek-aspek politik dan interpersonal studi evaluasi, menjalin hubungan selama pengumpulan data, dan seterusnya.

Dalam makna yang lebih luas, komunikasi dapat dianggap sebagai prosedur menggunakan dan menginformasikan sesuatu kepada pihak lain sehingga dalam hal ini evaluator tidak hanya memperhatikan informasi yang benar tetapi bagaimana mengkomunikasikannya kepada audiens. Data dalam bentuk statistik hanyalah sekedar deretan angka saja yang tidak bermakna bagi audiens apabila tidak diberikan pemaparan informasi lebih lanjut.

Komunikasi dan persuasi sangat penting bagi evaluator untuk menyesuaikan setiap presentasi laporan yang dibuat dapat dimengerti audiens, untuk itu diperlukan perhatian dan kreativitas evaluasi. Untuk itu maka penggunaan analogi-analogi, grafik atau *display* gambar, ringkasan hasil temuan, menggunakan format garis bawah pada temuan-temuan penting, dan menggunakan statistik yang sewajarnya saja.

### **3. *Level of detail (tingkat rincian).***

Laporan evaluasi tertulis harus diorganisasi sedetail mungkin sehingga audiens dapat dengan cepat mengakses berbagai informasi yang dibutuhkan.

### **4. *Technical writing style (Teknik gaya penulisan).***

Laporan evaluasi hendaknya tidak menimbulkan ambiguitas bagi audiens, oleh karena itu pemaparannya harus jelas sehingga mudah dipahami. Untuk itu perlu diperhatikan penggunaan bahasa, gambar, grafik, dan tampilan tabel sehingga terjadi komunikasi yang efektif. Melalui pilihan teknis gaya penulisan, maka laporan evaluasi tidak membosankan untuk dibaca.

Fitzpatrick dkk (2004) menyarankan beberapa hal terkait dengan teknis gaya penulisan laporan evaluasi yaitu:

- a. Hindari jargon. Walaupun menggunakan jargon berikan penjelasan sehingga istilah-istilah atau jargon yang digunakan tersebut kredibilitas.
- b. Gunakan bahasa sederhana. Dalam hal ini penggunaan bahasa harus disesuaikan dengan pemahaman audiens dan tidak bertele-tele.
- c. Gunakan contoh-contoh, anekdot, ilustrasi. Ingatlah satu gambar dapat mewakili seribu kata-kata dan yang terpenting adalah jangan pula sembarangan untuk mengilustrasikan suatu teks.
- d. Gunakan tata bahasa dan tanda baca yang benar, di samping ejaan juga harus sesuai dengan penggunaan bahasa di mana laporan evaluasi akan digunakan, menghindari narasi yang mengacaukan catatan referensi.
- e. Gunakan bahasa yang menarik dan tidak membosankan.

### **5. *Appearance of the report (kemunculan tampilan laporan).***

Laporan evaluasi hendaknya dikemas secara atraktif oleh evaluator. Terkait dengan hal ini, Fitzpatrick dkk (2004) menyampaikan hal-hal yang diperhatikan:

- a. Kualitas cetak.  
Kualitas *hardcopy* laporan evaluasi hendaknya menjadi perhatian evaluator, mulai dari cover, penggunaan tinta, intensitas cetakan (terang-gelap) menjadi sesuatu yang penting, termasuk ukuran atau format laporan.

b. Grafis.

Hal-hal terkait dengan grafis (foto, ilustrasi, gambar) menjadi sesuatu yang harus diperhatikan evaluasi dalam membuat *hardcopy* laporan evaluasi.

c. Tampilan halaman.

Laporan evaluasi hendaknya tidak menampilkan halaman-halaman yang monoton sehingga menimbulkan kebosanan untuk membacanya. Untuk itu dapat dilakukan beberapa hal sebagai berikut:

- 1) Ruang putih (untuk memisahkan dan menghilangkan bagian cetak).
- 2) *Heading* yang bervariasi, (*reguler* dan *boldface*).
- 3) Garis bawah atau huruf miring (tidak hanya untuk memberikan penekanan tetapi juga untuk menambahkan variasi).
- 4) Penggunaan *numbering* (nomor) atau *bulleted* (daftar butir).
- 5) Penggunaan visual (grafik, gambar, atau kartun).
- 6) Penggunaan *boxes* (kotak) atau menampilkan-visual lainnya dari bahan-bahan yang dipilih).
- 7) Warna. Penggunaan warna dapat membuat laporan evaluasi yang lebih menarik, serta lebih fungsional. Apabila ringkasan eksekutif muncul sebagai bagian pertama dalam sebuah laporan evaluasi, maka dapat dicetak dalam berbagai variasi warna. Hal ini tidak hanya memberikan beberapa daya tarik visual tetapi juga menarik perhatian ke ringkasan dan membuatnya mudah bagi pembaca untuk menemukan nanti. Demikian juga apendiks dapat dicetak dengan warna berbeda lainnya. Kombinasi warna akan meningkatkan daya tarik visual dari laporan evaluasi tersebut.
- 8) Cover. Sampul depan maupun bagian belakang merupakan bagian yang menjadi perhatian evaluator dalam desain laporan evaluasi. Cover yang menarik menunjukkan laporan tersebut layak presentasi.

Untuk menutup kajian pada bagian ini, Fitzpatrick dkk (2004) membuat daftar cek yang dapat digunakan dalam menilai laporan evaluasi sebagai berikut:

- 
- 
- 
- Isi laporan dirancang untuk audiens.
  - Format laporan dan *style* dirancang untuk audiens.
  - Keterlibatan audiens dalam menentukan format dan gaya laporan.



- \_\_\_\_\_ Ringkasan eksekutif.
- \_\_\_\_\_ Pengenalan yang memadai untuk rancangan laporan.
- \_\_\_\_\_ Penyebutan pembatasan riset.
- \_\_\_\_\_ Presentasi yang memadai dari rencana dan prosedur evaluasi (ada dalam apendiks).
- \_\_\_\_\_ Presentasi hasil/temuan terorganisir.
- \_\_\_\_\_ Semua informasi teknis yang diperlukan disediakan.
- \_\_\_\_\_ Standar spesifikasi dan kriteria untuk penilaian evaluasi.
- \_\_\_\_\_ Penilaian evaluasi.
- \_\_\_\_\_ Daftar kekuatan dan kelemahan-kelemahan.
- \_\_\_\_\_ Rekomendasi tindakan.
- \_\_\_\_\_ Perlindungan untuk klien dan kepentingan *stakeholder*.
- \_\_\_\_\_ Kepekaan terhadap mereka yang dipengaruhi oleh penemuan evaluasi.
- \_\_\_\_\_ Pembiayaan untuk evaluator dan pihak-pihak lain.
- \_\_\_\_\_ Akurasi dan bias presentasi.
- \_\_\_\_\_ Komunikasi yang efektif dan persuasif.
- \_\_\_\_\_ Tingkat detail.
- \_\_\_\_\_ Kurangnya istilah teknis.
- \_\_\_\_\_ Penggunaan, kelengkapan, kemenarikan bahasa.
- \_\_\_\_\_ Penggunaan contoh-contoh ilustrasi.
- \_\_\_\_\_ Tampilan visual dan daya tarik.

## **G. Pemanfaatan Laporan Evaluasi Program.**

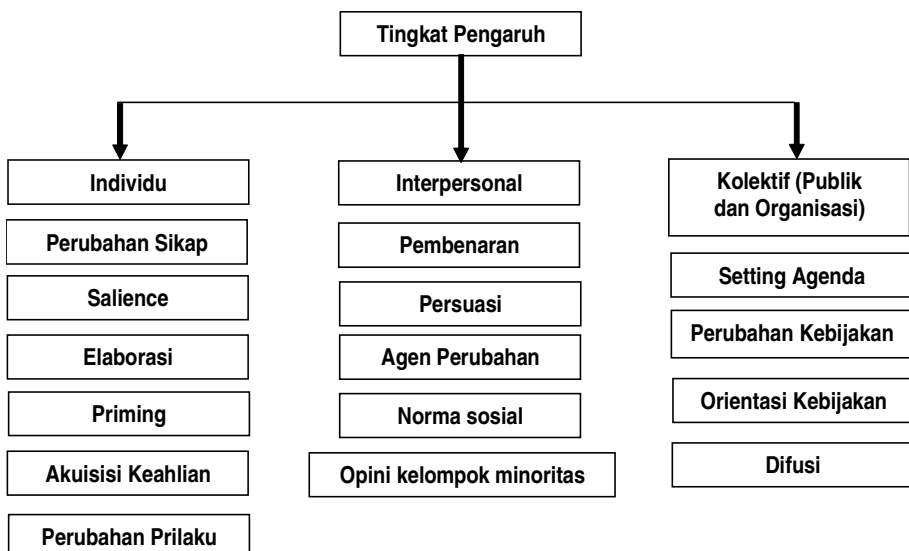
Kirkhart (2000), dan Henry dan Markus (2003) mengembangkan model-model dalam menentukan jenis dan dampak dari evaluasi. Model Kirkhart, ini berdasakan konsep Patton (1986) dan Weiss (1980) yang menekankan bahwa bahasa yang digunakan dalam laporan evaluasi mempengaruhi dalam penggunaannya. Kirkhart menggunakan istilah mempengaruhi (kemampuan atau kuasa orang atau sesuatu untuk menghasilkan dampak pada orang lain secara langsung maupun tidak langsung. Kirkhart juga mengusulkan sebuah model teori terpadu pengaruh untuk menggambarkan kesan evaluasi yang berbeda yang terdiri dari tiga dimensi yaitu: *source* (sumber), *intention* (tujuan) dan *time* (waktu).

Dimensi *source* (sumber) terkait dengan pengaruh laporan evaluasi melalui hasil ataupun prosesnya. Evaluator berfokus pada orientasi atas hasil yang digunakan, dan stakeholder dapat menggunakan hasil tersebut untuk membuat keputusan atau penilaian. Selain untuk efek atau hasil, model Kirkhart menggambarkan bahwa proses evaluasi dapat menyebabkan perubahan manajemen organisasi dan program, dan juga menciptakan dialog antara para pemangku kepentingan yang berbeda.

Dimensi *intention* (tujuan) merupakan tanda untuk mempertimbangkan konsekuensi dari evaluasi yang akan digunakan. Laporan evaluasi berdampak kepada tujuan ataupun perencanaan yang membantu meningkatkan pengaruh penilaian oleh pemangku kepentingan yang berbeda-beda.

Dimensi *time* (waktu). Dampak dari hasil/temuan evaluasi tentunya memerlukan jangka waktu tertentu dalam implementasinya. Mungkin dapat dirasakan pengaruh dalam jangka waktu pendek, sedang maupun dalam jangka waktu panjang. Sebelum menggunakan konsepsi akan membentuk tetapi satu dua kubu dalam model terpadu pengaruh: dimaksudkan, serta-merta atau akhir-akhir dari siklus-penggunaan hasil.

Henry dan Mark (2003) mengusulkan sebuah model yang berbeda, atau yang dimaksudkan kerangka panduan untuk penelitian dan praktek pada dampak-dampak evaluasi. Seperti halnya Kirkhart, maka Henry dan Mark menekankan pada istilah “menggunakan” untuk memeriksa dampak-dampak evaluasi. Terkait dengan dampak penggunaan hasil/temuan evaluasi digambarkan oleh Henry dan Mark (2003) dalam diagram sebagai berikut:



Dari bagan di atas dapat dilihat tingkat pengaruh dari hasil/temuan evaluasi yang berdampak kepada individu, hubungan interpersonal dan publik maupun organisasi. Pada masing-masing tingkatan, berbagai perubahan dapat terjadi. Evaluator dapat mengenali antara tingkat ini dan tipe yang mempertimbangkan bagaimana cara untuk mencapai pengaruh yang diinginkan dengan evaluasi tertentu.

Selanjut terdapat berbagai penelitian terkait dengan kajian dengan pengaruh atau dampak hasil evaluasi yang dilakukan sejak tahun 1980-an. Rangkuman dari berbagai kajian penelitian memaparkan tentang kontribusi penggunaan hasil/temuan evaluasi yaitu: (1) relevansi evaluasi kepada pembuat keputusan dan atau pihak lain, (2) keterlibatan pengguna dalam perencanaan dan melaporkan tahap-tahap evaluasi, (3) reputasi atau kredibilitas evaluator, dan (4) pengembangan prosedur untuk membantu menggunakan rekomendasi untuk melakukan tindakan-tindakan.

## DAFTAR BACAAN

- Alkin, M.C. (1985). *Guide For Evaluation Decision Makers*. Beverly Hills: Sage Publication.
- Arikunto, S. (1988). *Penilaian Program Pendidikan*. Jakarta: Bina Aksara.
- Arikunto, S. (2005). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Arikunto, S., dan Jabar, C.S.A. (2009). *Evaluasi Program Pendidikan. Pedoman Teoretis Praktis Bagi Mahasiswa Dan Praktisi Pendidikan*. Jakarta: Bumi Aksara.
- Briekerhoff, R.O., et-al. (1983). *Program Evaluation. A Source Book*. Boston: Kluwer Nijboff Publishing.
- Briekerhoff, R.O., et-al. (1986). *Program Evaluation, A Practitioner's Guide for Trainer and Educationer*, 4<sup>th</sup> Edition. Boston: Kluwer Nijboff Publishing.
- Briggs, L. (1977). *Instructional Design, Principles and Applications*. New Jersey: Educational Technology Publications.
- Djaali dan Muljono, P. (2004). *Pengukuran Dalam Bidang Pendidikan*. Jakarta: Universitas Negeri Jakarta.
- Djarwanto, PS., dan Subagya, P. (1998) *Statistik Induktif*. Yogyakarta: BPFE.
- Faisal, S. (1999). *Penelitian Kualitatif. Dasar-Dasar dan Aplikasi*. Malang: YA3.
- Fitzpatrick, J.L., Sanders, J.R., dan Worthen, B.R. (2004). *Program Evaluation: Alternative Approaches and Practical Guidelines*. Boston: Pearson Education Inc.
- Feuerstein, M.T. (1990). *Partners in Evaluation*. Alihbahasa: Farid Wadjidi. *Evaluasi Partisipatoris*. Jakarta: P3M.
- Guba, E.G., dan Lincoln, Y.S. (1985). *Effective Evaluation*. San Fransisco: Jossey Bass Publishing.
- Hardjodipuro, S. (1982). *System Planning*. Jakarta: Erlangga.

- Isaac, S., dan Michael, W.B., (1987). *Handbook in Research and Evaluation for Education and the Behavioral Science*, 2<sup>nd</sup> Edition. California: Edits Publishers.
- Kaufman, R., dan English, F.W. (1979) *Need Assessment, Concept And Application*. New Jersey: Educational Technology Publication, Englewood Cliffs.
- Kifer, E. dalam Anglin, G. (1995) *Instructional Technology, Past, Present, Future*, 2<sup>nd</sup> Edition, Englewood: Colorado Libraries Unlimited Inc.
- Kirkpatrick, J.L. (1998). *Evaluating Training Program, The Four Levels*, 2<sup>nd</sup> Edition, San Fransisco: Berret-Kohler Publisher, Inc.
- \_\_\_\_\_. (2005). *Kirkpatrick's Training Evaluating Model*, <http://www.businessball.com/kirkpatrickleaningevaluationmodel.htm>
- Madaus, G.F., Sriven, M.S., dan Stufflebeam. (1987). *Evaluation Models Viewpoints on Educational and Human Services Evaluation*. Boston: Kluwer Nijboff Publishing.
- Mehrens, W.A dan Lehmann, I.J. (1978). *Measurement and Evaluation in Education and Psychology*. 2<sup>nd</sup> New York: Holt, Rinehart and Winston.
- Mutrofin, (2010). *Evaluasi Program, Teks Pilihan Untuk Pemula*. Yogyakarta: Lakesbang Pressindo.
- Oliva, P.F (1992). *Developing the Curriculum*. New York: Harper Collins Publishers.
- Purwanto dan Suparman, A. (1999). *Evaluasi Program Diklat*. Jakarta: STIA-LAN Press.
- Purwanto, M.N. (2001). *Prinsip-Prinsip Dan Teknik Evaluasi Pembelajaran*. Bandung: Remaja Rosdakarya.
- Riduwan. (2002). *Skala Pengukuran Variabel-Variabel Penelitian*. Bandung: Alfabeta.
- Sagala, S. 2010. *Manajemen Strategik Dalam Peningkatan Mutu Pendidikan*. Bandung: Alfabeta.
- Stark, J. S., dan Thomas, A. (1994). *Assessment and Program Evaluation*. Needham Heights: Simon & Schuster Custom Publishing.
- Stufflebeam, D. L. (2003). *The CIPP Model for Evaluation: the Article Presented at the 2003 Annual Conference of the Oregon Program Evaluators Network (OPEN) 3 October 2003* (online). (<http://www.wmich.edu>).
- Stufflebeam, D.L dan Shinkfield, A.J. (2007). *Evaluation Theory, Models and Application*. San Francisco: Jossey Bass.

- Sudidjono, A. (2000). *Pengantar Statistik Pendidikan*, Jakarta: Raja Grafindo Persada.
- Sudjana, D. (2008). *Evaluasi Program Pendidikan Luar Sekolah*. Bandung: Remaja Rosdakarya.
- Sudjana, N., dan Ibrahim. (2004). *Penelitian dan Penilaian Pendidikan*. Bandung: Sinar Baru Algesindo.
- Sugiyono. (2004). *Statistika Untuk Penelitian*. Bandung: Alfabeta.
- Sugiyono. (2010). *Metode Penelitian Pendidikan, Pendekatan Kuantitatif, Kualitatif dan R&D*. Bandung: Alfabeta.
- Suherman, E. dan Sukjaya, Y. (1990). *Petunjuk Praktis Untuk Melakukan Evaluasi Pendidikan Matematika*. Bandung: Wijayakusumah.
- Sukmadinata, N.S. (2006). *Metode Penelitian Pendidikan*. Bandung: Remaja Rosdakarya.
- Supardi U.S. (2013). *Aplikasi Statistika Dalam Penelitian Konsep Statistika Yang Lebih Komprehensif*. Jakarta: Change Publication.
- Tayibnaxis, F.Y. (2000). *Evaluasi Program*. Jakarta: Rineka Cipta.
- Taylor, P et-al. (1996). *Planning a Program Evaluation*. Wisconsin: University of Wisconsin-Extension.
- Weiss, C.H. (1972), *Evaluation Research, Methods for Assessing Program Effectiveness*. Englewood: Prentice Hall.
- Worthern, B.R., dan Sander, J.R. dalam N.L. Smith (1996). *Evaluation Models and Approaches*. Japan: Pergamon Elsevier Science Ltd.

## Lampiran 1

# GLOSSARIUM

### **Analisis kebutuhan**

Proses yang dilakukan untuk mengidentifikasi kebutuhan sekaligus menentukan prioritas diantaranya.

### **Audiens**

Peminat, pemakai, atau pelanggan yang secara langsung atau tidak langsung berurusan dengan evaluasi.

### **Biaya langsung**

Biaya-biaya yang diperlukan untuk mendanai seluruh kegiatan evaluasi seperti gaji/honor, biaya konsultan, biaya ATK dan sebagainya.

### **Biaya tak langsung**

Biaya yang diperlukan untuk mendukung kegiatan evaluasi, tetapi tidak termasuk dalam pos biaya langsung.

### **Data**

Fakta-fakta dan informasi yang dikumpulkan untuk tujuan evaluasi program.

### **Data kualitatif**

Data yang berupa atau berbentuk kata-kata, seperti keterangan tentang kejadian, transkrip wawancara, dan dokumen tertulis.

### **Data kuantitatif**

Data yang berupa angka-angka terkait dengan jumlah dan dapat diukur.

### **Desain evaluasi**

Seperangkat keputusan tentang bagaimana evaluasi akan dilaksanakan, misalnya tujuannya, pengumpulan datanya, pelaporan dan sebagainya.

**Evaluasi**

Penyediaan informasi yang dapat dijadikan sebagai bahan pertimbangan dalam mengambil keputusan.

**Evaluasi formatif**

Evaluasi yang dipakai untuk perbaikan dan pengembangan kegiatan atau program yang sedang berlangsung

**Evaluasi program**

Suatu unit atau kesatuan kegiatan yang bertujuan mengumpulkan informasi tentang realisasi atau implementasi dari suatu kebijakan, berlangsung dalam proses yang berkesinambungan, dan terjadi dalam suatu organisasi yang melibatkan sekelompok orang guna pengambilan keputusan.

**Evaluasi sumatif**

Evaluasi yang dipakai untuk mengambil keputusan apakah kegiatan atau program yang dievaluasi dilanjutkan atau diberhentikan.

**Evaluator**

Individu atau tim yang melakukan pekerjaannya mengevaluasi suatu program.

**Instrumen**

Alat yang digunakan untuk mengumpulkan data dalam evaluasi seperti kuesioner, wawancara, tes.

**Kesimpulan**

Sesuatu yang merupakan inti dari sederetan informasi atau sajian yang menyatakan tentang status program yang sedang dievaluasi.

**Klien**

Orang atau organisasi yang memesan atau meminta kepada evaluator untuk melakukan evaluasi.

**Kontrak evaluasi**

Kesepakatan lisan atau tertulis antara evaluator dengan klien yang berisikan tentang harapan dan tanggungjawab kedua belah pihak, sehingga berfungsi sebagai dasar pelaksanaan pekerjaan evaluasi dan ikatan secara hukum



**Kuesioner**

Pertanyaan tertulis atau tercetak yang digunakan untuk memperoleh informasi dari responden.

**Laporan evaluasi**

Catatan tertulis atau lisan yang dibuat atau disajikan untuk tujuan melaporkan hasil evaluasi program.

**Meta evaluasi**

Evaluasi dari suatu evaluasi.

**Objek evaluasi**

Apa yang akan dievaluasi, dapat berupa program, proyek, training, materi atau bahkan evaluasi lainnya.

**Observasi**

Pengamatan secara langsung terhadap proses yang berlangsung di *setting* program yang di evaluasi.

**Pendekatan ecletik**

Memilih berbagai metode dari beberapa pilihan terbaik sesuai dengan kebutuhan.

**Program**

Suatu unit atau kesatuan kegiatan yang merupakan realisasi atau implementasi dari suatu kebijakan.

**Rangkuman eksekutif**

Pernyataan ringkas yang dirancang untuk menyediakan tinjauan cepat dari laporan evaluasi yang lengkap.

**Rekomendasi**

Saran-saran praktis bagi *stakeholder* program terkait dengan jalannya program yang dievaluasi. Melalui rekomendasi ini pengambil kebijakan dapat membuat kebijakan lanjutan.

**Reliabilitas**

Tingkat keterhandalan suatu instrumen yang hasil pengukurannya dapat dipercaya.

**Sponsor**

Orang atau organisasi yang meminta evaluasi dan membayar untuk itu.

**Statistik deskriptif**

Statistik yang mempelajari tata cara mengumpulkan, menyusun, menyajikan dan menganalisa data penelitian yang berwujud angka-angka, agar dapat memberikan gambaran yang teratur, ringkas dan jelas mengenai suatu gejala, keadaan peristiwa, sehingga dapat ditarik pengertian atau makna tertentu.

**Statistik inferensial**

Statistik yang mempelajari atau mempersiapkan tata cara penarikan kesimpulan mengenai karakteristik populasi, berdasarkan data kuantitatif yang diperoleh dari sampel penelitian.

**Sumber data**

Segala sesuatu yang menunjuk pada asal data diperoleh. Dalam evaluasi program sumber data dikenal dengan istilah 3P yaitu *person, paper, place*.

**Tabulasi**

Menyusun menjadi tabel, dengan kata lain tabulasi adalah pengolahan atau pemrosesan menjadi tabel.

**Triangulasi**

Suatu cara memandang permasalahan/objek yang dievaluasi dari berbagai sudut pandang, dapat dipandang dari banyak metode atau sumber data. Tujuannya agar dapat melihat objek evaluasi dari semua sisi.

**Validitas**

Ketepatan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya.

## TENTANG PENULIS

**Rusydi Ananda**, Lahir di Tanjung Pura Langkat, dengan Ayah yang bernama H. Thaharuddin AG (alm) dan Ibu Hj. Rosdiani. Anak pertama dari 6 bersaudara. Menempuh pendidikan SD di Medan tamat tahun 1984, melanjutkan ke SMP di Medan tamat tahun 1987, kemudian menyelesaikan SMU di Medan tamat pada tahun 1990. Melanjutkan pendidikan strata 1 (S.1) di IAIN SU jurusan Tadris Matematika yang diselesaikan pada tahun 1995. Meraih gelar Magister Pendidikan dari Universitas Negeri Medan dengan konsentrasi studi Teknologi Pendidikan pada tahun 2005. Saat ini menyelesaikan studi S3 di Universitas Negeri Jakarta pada program studi Teknologi Pendidikan.

Menikah dengan Tien Rafida, yang berprofesi sebagai PNS/Dosen di Fakultas Ilmu Tarbiyah dan Keguruan UIN Sumatera Utara. Saat ini dikarunia Allah SWT 3 (tiga) orang anak, yaitu: Annisa Arfitha, Salsabila Hadiyanti dan Faturrahman.

Pengalaman kerja dimulai sebagai tenaga administrasi di PT. Marhamah Medan pada tahun 1995-1996. Guru matematika di SMP Perguruan Bandung tahun 1996-1997. Guru Matematika di SMA UISU Medan Tahun 1997-1999. Sejak tahun 2000 sampai sekarang bekerja sebagai PNS/Dosen di Fakultas Ilmu Tarbiyah dan Keguruan UIN Sumatera Utara. Sejak tahun 2006 – 2008 bertugas di pusat penelitian UIN Sumatera Utara dan tahun 2008 – 2011 dipercaya sebagai ketua program studi Pendidikan Matematika UIN Sumatera Utara.

Aktivitas lainnya yang digeluti adalah sebagai trainer di Widya Pustpita tahun 2003 – 2009, trainer pada kegiatan yang dikelola DBE2 USAID tahun 2006 – 2010, dan trainer di AUSAID sejak tahun 2014 - 2015.

Karya berupa buku yang sudah diterbitkan adalah Evaluasi Pembelajaran (2014), Penelitian Tindakan Kelas (2015), Pengantar Kewirausahaan, Rekayasa Akademik Melahirkan Enterpreneurship (2016).

**Tien Rafida**, lahir di Pematang Siantar, dengan Ayah yang bernama (Alm) H. Arifin dan Ibu Hj. (Alm). Hj. Zubaidah (Alm). Anak kedua dari 2 bersaudara. Menempuh pendidikan SD di Siantar tamat tahun 1983, melanjutkan ke SMP di Siantar tamat tahun 1986, kemudian menyelesaikan SMU di Siantar tamat pada tahun 1989. Melanjutkan pendidikan strata 1 (S.1) di IAIN SU jurusan Tadris Bahasa Inggris yang diselesaikan pada tahun 1995. Tahun 2000 meraih gelar Magister Humaniora Program studi Linguistik dari Universitas Sumatera Utara dan meraih gelar Doktor Bidang Linguistik dari Universitas Sumatera Utara tahun 2014.

Pengalaman kerja dimulai sebagai pegawai kontrak kerja di BNI Pematang Siantar tahun 1994-1996, Tahun 1997 – 2000 bertugas sebagai guru di MTsN Bandar Simalungun. Sejak tahun 2001 menjadi beralih fungsi menjadi Dosen di UIN Sumatera Utara. Tahun 2012-2016 dipercaya sebagai ketua program studi Pendidikan Bahasa Inggris IAIN Sumatera Utara yang saat ini menjadi UIN Sumatera Utara. Aktif juga pada beberapa kegiatan USAID Prioritas sebagai instruktur.

Mendapat kesempatan mengikuti program Sandwich di Nanyang Technology University NIE Singapura selama 3 bulan pada tahun 2011, Doktoral Program pada Leiden University Eropa tahun 2013 dan Pada tahun 2015 Duta Dosen UIN Medan program speech ke Fatani University Thailand.

Karya berupa buku yang sudah diterbitkan adalah Psikolinguistik (2007), Penelitian Tindakan Kelas (2015), Pengantar Kewirausahaan, Rekayasa Akademik Melahirkan Enterpreneurship (2016).

