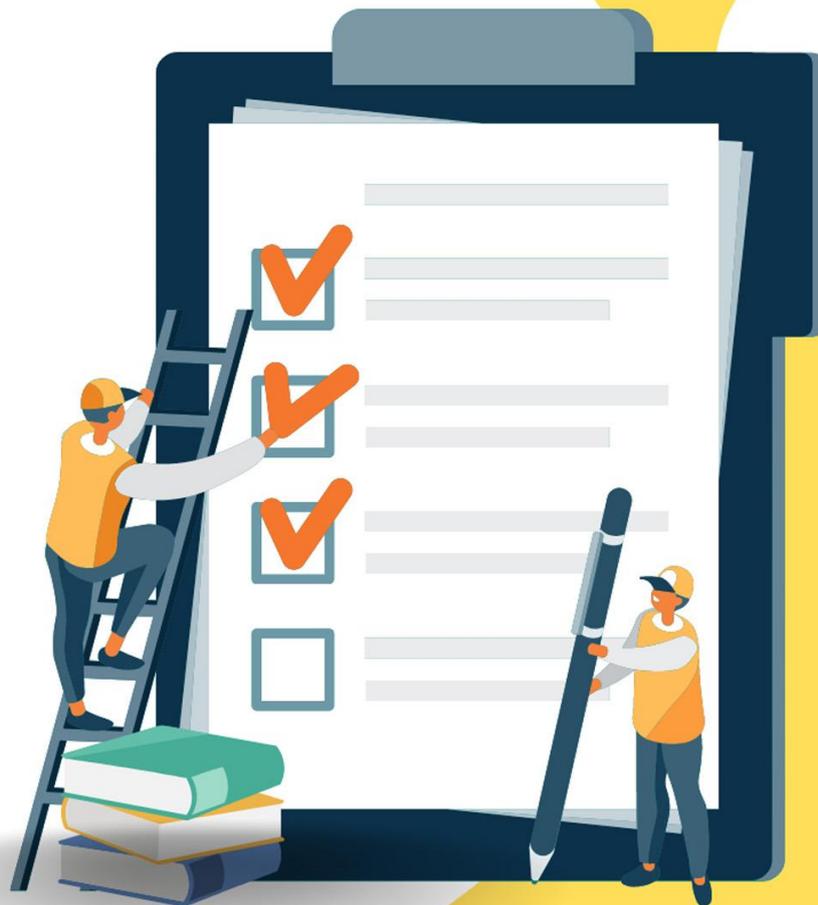


Modul Pembelajaran

EVALUASI PROGRAM PENDIDIKAN



Dr. Musringudin, M. Pd
Prof. Dr. Abd. Rahman A. Ghani, M. Pd
Dr. Dwi Priyono, M. Pd

**MODUL PEMBELAJARAN
EVALUASI PROGRAM PENDIDIKAN**

UU No 28 tahun 2014 tentang Hak Cipta

Fungsi dan sifat hak cipta Pasal 4

Hak Cipta sebagaimana dimaksud dalam Pasal 3 huruf a merupakan hak eksklusif yang terdiri atas hak moral dan hak ekonomi.

Pembatasan Pelindungan Pasal 26

Ketentuan sebagaimana dimaksud dalam Pasal 23, Pasal 24, dan Pasal 25 tidak berlaku terhadap:

- i Penggunaan kutipan singkat Ciptaan dan/atau produk Hak Terkait untuk pelaporan peristiwa aktual yang ditujukan hanya untuk keperluan penyediaan informasi aktual;
- ii Penggandaan Ciptaan dan/atau produk Hak Terkait hanya untuk kepentingan penelitian ilmu pengetahuan;
- iii Penggandaan Ciptaan dan/atau produk Hak Terkait hanya untuk keperluan pengajaran, kecuali pertunjukan dan Fonogram yang telah dilakukan Pengumuman sebagai bahan ajar; dan
- iv Penggunaan untuk kepentingan pendidikan dan pengembangan ilmu pengetahuan yang memungkinkan suatu Ciptaan dan/atau produk Hak Terkait dapat digunakan tanpa izin Pelaku Pertunjukan, Produser Fonogram, atau Lembaga Penyiaran.

Sanksi Pelanggaran Pasal 113

1. Setiap Orang yang dengan tanpa hak melakukan pelanggaran hak ekonomi sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf i untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 1 (satu) tahun dan/atau pidana denda paling banyak Rp100.000.000 (seratus juta rupiah).
2. Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf c, huruf d, huruf f, dan/atau huruf h untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 3 (tiga) tahun dan/atau pidana denda paling banyak Rp500.000.000,00 (lima ratus juta rupiah).

**MODUL PEMBELAJARAN
EVALUASI PROGRAM PENDIDIKAN**

Dr. Musringudin, M. Pd
Prof. Dr. Abd. Rahman A. Ghani, M. Pd
Dr. Dwi Priyono, M. Pd

Penerbit



CV. MEDIA SAINS INDONESIA
Melong Asih Regency B40 - Cijerah
Kota Bandung - Jawa Barat
www.medsan.co.id

Anggota IKAPI
No. 370/JBA/2020

**MODUL PEMBELAJARAN
EVALUASI PROGRAM PENDIDIKAN**

Dr. Musringudin, M. Pd
Prof. Dr. Abd. Rahman A. Ghani, M. Pd
Dr. Dwi Priyono, M. Pd

Editor :
Rintho R. Rerung

Tata Letak :
Rizki R. Pratama

Desain Cover :
Rintho R. Rerung

Ukuran :
A4: 21 x 29,7 cm

Halaman :
viii, 228

ISBN :
978-623-362-446-6

Terbitan:
Maret, 2022

Hak Cipta 2022 @ Media Sains Indonesia dan Penulis

Hak cipta dilindungi undang-undang. Dilarang keras menerjemahkan, memfotokopi, atau memperbanyak sebagian atau seluruh isi buku ini tanpa izin tertulis dari Penerbit atau Penulis.

PENERBIT MEDIA SAINS INDONESIA
(CV. MEDIA SAINS INDONESIA)
Melong Asih Regency B40 - Cijerah
Kota Bandung - Jawa Barat
www.medsan.co.id

PRAKATA

Pada kesempatan ini, penyusun ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu penyusun, baik secara langsung maupun tidak langsung dalam penyusunan modul ini sampai selesai. Dengan kerendahan hati, perkenankan penyusun menyampaikan rasa terima kasih yang sebesar-besarnya kepada;

1. Prof. Dr. Gunawan Suryoputro, M. Hum, selaku Rektor Universitas Muhammadiyah Prof. Dr. Hamka.
2. Prof. Dr. Abd. Rahman Ghani, M. Pd selaku wakil Rektor 1 Universitas Muhammadiyah Prof. Dr. Hamka.
3. Dr. Zamah Sari, M. Ag selaku wakil Rektor 2 Universitas Muhammadiyah Prof. Dr. Hamka
4. Dr. Tri Wintolo Apoko, M. Pd selaku Ketua Lembaga Pengembangan Pendidikan dan Pengajaran Universitas Muhammadiyah Prof. Dr. Hamka
5. Prof. Dr. Ade Hikmat, M. Pd selaku Direktur Sekolah Pasacasarjana UHAMKA
6. Dr. Ihsana El Khuluko, M. Pd selaku Kaprodi MAP Sekolah Pascasarjana UHAMKA

Penulis

DAFTAR ISI

PRAKATA	i
DAFTAR ISI.....	ii
DESKRIPSI MATA KULIAH.....	v
MODUL 1 KONSEP PENELITIAN ILMIAH, KEBENARAN ILMIAH, METODE ILMIAH, EVALUASI	1
A. Konsep Penelitian Ilmiah	1
B. Kebenaran Ilmiah.....	4
C. Metode Ilmiah	5
D. Evaluasi	7
MODUL 2 TAHAPAN EVALUASI, JENIS RISET DAN EVALUASI VS RISET	17
A. Tahapan Evaluasi.....	17
B. Jenis Riset.....	23
C. Evaluasi Vs Riset.....	24
MODUL 3 PERAN DAN TUJUAN EVALUASI.....	31
A. Peran Dan Tujuan Evaluasi.....	31
B. Arti Evaluasi.....	35
C. Fungsi Evaluasi.....	37
MODUL 4 OBJEK EVALUASI, DIMENSI OBJEK EVALUASI, KRITERIA DALAM EVALUASI DAN BEBERAPA ISTILAH DALAM EVALUASI	43
A. Obyek Evaluasi	43
B. Dimensi Obyek Evaluasi.....	45
C. Menentukan Kriteria Dalam Evaluasi.....	46
D. Beberapa Istilah Dalam Evaluasi.....	50
MODUL 5 PENERIMA MANFAAT EVALUASI, PROSEDUR EVALUASI, METODE DAN PENDEKATAN EVALUASI, PIHAK YANG MELAKUKAN EVALUASI.....	57
A. Penerima Manfaat Evaluasi	57
B. Prosedur Evaluasi	58
C. Metode dan Pendekatan Evaluasi.....	60
D. Pihak yang Melakukan Evaluasi.....	61
MODUL 6 TEORI EVALUASI DAN PERAN TEORI DALAM EVALUASI .	71
A. Konsep Teori Evaluasi	71
B. Peran Teori dalam Evaluasi	81

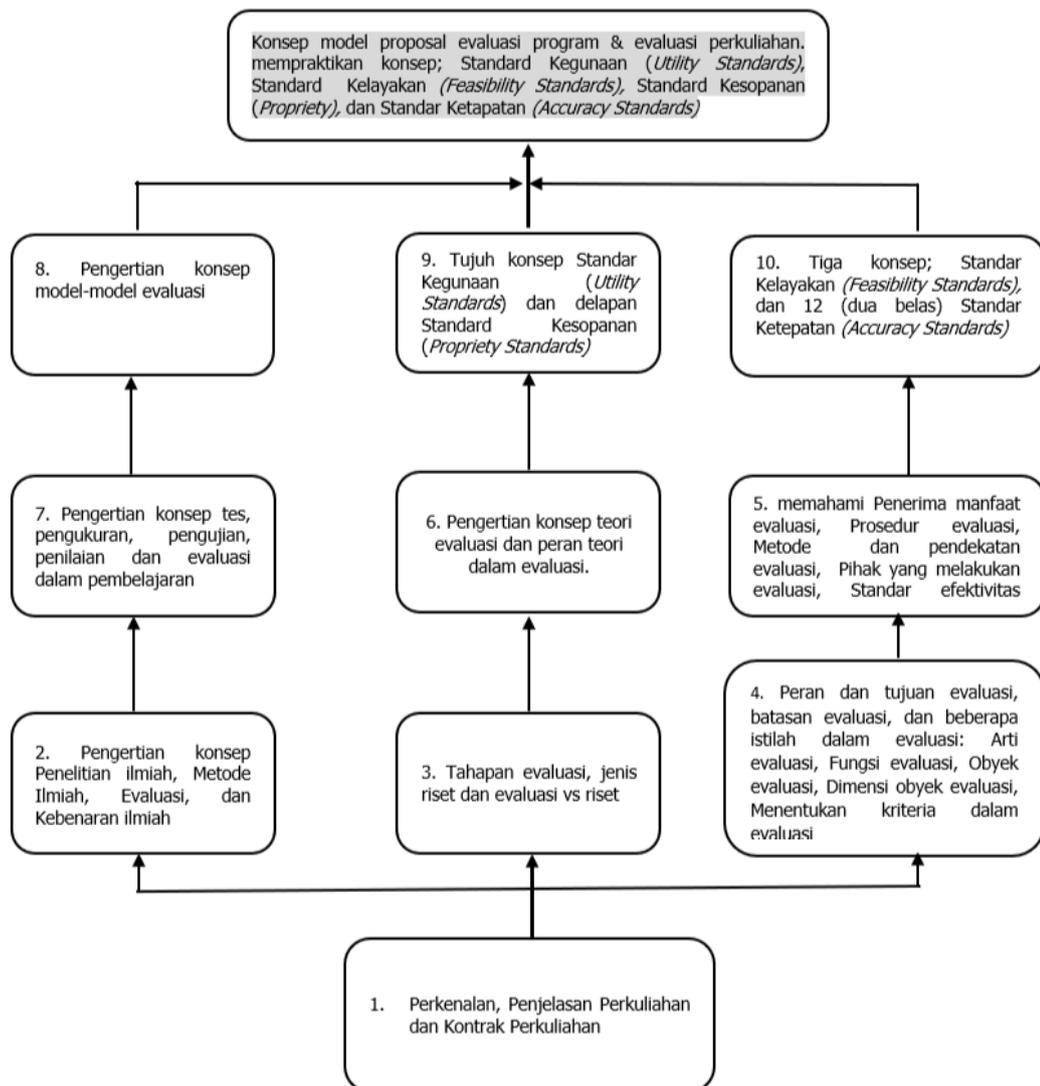
MODUL 7 PENGERTIAN KONSEP TES DAN PENGUKURAN.....	89
A. Pengertian Konsep Tes	89
B. Fungsi Tes.....	91
C. Bentuk Tes.....	92
D. Pengukuran.....	93
MODUL 8 PENGERTIAN KONSEP PENILAIAN DAN EVALUASI DALAM PEMBELAJARAN	103
A. Penilaian	103
B. Evaluasi Dalam Pembelajaran.....	106
MODUL 9 MODEL-MODEL EVALUASI GOAL-BASED EVALUATION, GOAL FREE EVALUATION, FORMATIF & SUMATIF EVALUATION, DAN RESPONSIVE EVALUATION.....	117
A. Model-Model Evaluasi	117
B. Evaluasi Berbasis Tujuan (<i>Goal-Based Evaluation</i>).....	119
C. <i>Goal Free Evaluation Model</i> /Evaluasi Bebas Tujuan.....	124
D. Evaluasi Formatif dan Sumatif.....	132
E. Evaluasi Responsif	136
MODUL 10 MODEL-MODEL EVALUASI CIPP, EVALUASI ADVERSARY, EVALUASI MODEL KETIMPANGAN ATAU DISCREPANCY EVALUATION MODEL, EVALUASI SISTEM ANALISIS	145
A. Evaluasi CIPP.....	145
B. Evaluasi Adversary.....	153
C. Evaluasi Ketimpangan.....	158
D. Evaluasi Sistem Analisis	161
MODUL 11 MODEL EVALUASI BENCHMARKING, EVALUASI KOTAK HITAM (BLACK BOX EVALUATION MODEL), EVALUASI KONOSERSIP DAN KRITIKISME, UTILIZATION-FOCUSED EVALUATION.....	169
A. Evaluasi Benchmarking	169
B. Evaluasi Kotak Hitam (<i>Black Box Evaluation Model</i>)	171
C. Evaluasi Konosersip dan Kritikisme	174
D. Evaluasi Terfokus Utilisasi	177
MODUL 12 MODEL-MODEL EVALUASI AKREDITASI, THEORY-DRIVEN EVALUATION MODEL, MODEL EVALUASI SEMU ATAU PSEUDOEVALUATION, EVALUASI EKSANTE (EX-ANTE)	189
A. Akreditasi.....	189
B. Theory-driven Evaluation Model	195

C. Evaluasi Semu	199
D. Evaluasi Eksante (<i>EX-ante</i>).....	201
MODUL 13 DELAPAN KONSEP STANDAR KEGUNAAN (UTILITY STANDARDS) DAN TUJUH STANDARD KESOPANAN (PROPRIETY STANDARDS). EMPAT KONSEP; STANDAR KELAYAKAN (FEASIBILITY STANDARDS).....	209
A. Standar Evaluasi.....	209
B. <i>Utility standards</i>	211
C. <i>Feasibility Standards</i>	212
D. <i>Propriety Standards</i>	212
MODUL 14 DELAPAN STANDAR KETEPATAN (ACCURACY STANDARDS), TIGA STANDAR META EVALUASI DAN STANDAR ETIKA DALAM EVALUASI	217
A. <i>Accuracy standards</i>	217
B. <i>Meta-Evaluation</i>	218
C. Standar Etika dalam Evaluasi	218
GLOSARIUM	225
DAFTAR PUSTAKA.....	227

DESKRIPSI MATA KULIAH

Mata Kuliah Evaluasi Program Pendidikan ini memberikan wawasan kepada mahasiswa tentang konsep Evaluasi Program sebagai riset dan menjadi dasar pengetahuan untuk menyusun proposal penelitian Evaluasi.

Peta Kompetensi



MODUL 1

KONSEP PENELITIAN ILMIAH, KEBENARAN ILMIAH, METODE ILMIAH, EVALUASI

Metode Pembelajaran	Estimasi waktu	Capaian Pembelajaran
-Presentasi virtual -Diskusi Virtual interaktif -Tanya jawab	120 Menit X 1	-Perkenalan -Penjelasan perkuliahan dan kontrak perkuliahan -Mahasiswa mampu memahami pengertian konsep Penelitian ilmiah, Metode Ilmiah, Evaluasi dan Kebenaran ilmiah

Materi 1. Pengertian konsep Penelitian ilmiah, Kebenaran Ilmiah, Metode Ilmiah, dan Evaluasi.

A. Konsep Penelitian Ilmiah

Penelitian ilmiah merupakan salah satu cara mendapatkan pengetahuan yang dianggap memenuhi syarat kebenaran yang diakui secara logis. Hal ini dapat dipahami karena ada cara lain dalam memperoleh pengetahuan tetapi dengan pendekatan yang secara akal kurang bisa ditoleransi. Sebagai contoh bagi sebagian orang yang percaya pada wangsit maka mereka akan berusaha mendapatkan pengetahuannya melalui ritual tertentu (bersemedi) atau kegiatan lainnya. Pendekatan ilmiah juga menjadi salah satu cara yang paling dapat dipertanggungjawabkan hasilnya dalam hal memperoleh pengetahuan. Tentu saja karena metode ilmiah dilakukan dengan cara-cara yang sistematis dan terukur sehingga memiliki tingkat akurasi yang baik.

McMilan mengatakan *“The primary purpose of scientific inquiry is to explain natural phenomena and understand the underlying relationship and then, using this information, to predict and influence behavior”*.¹ Bahwa tujuan utama penelitian ilmiah adalah untuk menjelaskan

¹ James H. McMilan. *Educational Research: Fundamentals for the Consumer*. (USA: Pearson Education, Inc. 2008), hlm. 6.

fenomena alam dan memahami hubungan yang mendasarinya, kemudian menggunakan informasi yang ada untuk memprediksi dan memengaruhi perilaku.

Mertens menyatakan bahwa *“research is one of many different ways of knowing or understanding.in that it is a process of systematic inquiry that is designed to collect, analyze, interpret, and use data to understand, describe, predict, or control an educational or psychological phenomenon or to empower individual in such contexts”*.² Penelitian merupakan salah satu cara untuk mendapatkan pengetahuan atau cara untuk memahami. Penelitian adalah proses yang sistematis yang di disain untuk mengumpulkan, menganalisisi, menginterpretasikan, dan menggunakan data yang terkumpul untuk memahami, mendeskripsikan, memperkirakan, atau mengontrol fenomena Pendidikan atau psikologi atau untuk memberdayakan individu-individu dalam kontek yang relevan dengan objek penelitian.

Disana ada proses yang sistematis dalam mengumpulkan data atau informasi sehingga data dan informasi yang terkumpul layak untuk diolah menjadi pengetahuan yang dapat dipertanggungjawabkan kebenarannya. Menjadi penting untuk dijadikan catatan bahwa data atau informasi yang terkumpul harus melalui proses dan tahapan yang benar karena data dan informasi yang tidak benar akan menyesatkan.

Dalam melakukan penelitian ilmiah, peneliti harus mengikuti rambu-rambu yang menjadi standar perolehan pengetahuan ilmiah. Menurut *Committee on Scientific Principles for Education Research, Richard J. Shavelson and Lisa Towne, Editors, National Research Council (2002)* penelitian pendidikan memiliki enam prinsip yang perlu diikuti oleh para peneliti. Berikut enam prinsip yang dimaksud.

1. Mengajukan pertanyaan penelitian yang dapat diinvestigasi secara empiris. Prinsip pertama ini menitikberatkan pada dua hal: (1) pertanyaan penelitian yang berpotensi menguntungkan dan signifikan harus diidentifikasi dan dijawab; (2) berkaitan dengan

² Donna M. Mertens, *Research and Evaluation in Educational and Psychology: Integrating Diversity with Quantitative, Qualitative, And Mixed Methods*, 2nd Edition (California: Sage Publications, Inc. 2005), p. 2.

kebutuhan untuk mendesain studi empiris agar supaya pihak lain dapat menilai kualitas temuan penelitian serta metodenya.

2. Penelitian harus dilandasi dengan teori yang relevan. Dalam penelitian ilmiah, keberlanjutan dan pengujian teori sangat penting untuk membangun pengetahuan yang akan disampaikan secara luas ke publik.
3. Menggunakan metode yang memungkinkan investigasi langsung terhadap pertanyaan penelitian. Prinsip penting dalam melakukan penelitian Pendidikan adalah metode yang di gunakan seharusnya salah satu yang terbaik untuk menjawab pertanyaan penelitian. Sungguhpun tidak ada satu metode yang dapat memberikan jawaban terbaik terhadap pertanyaan penelitian. Metode harus sesuai dengan pertanyaan penelitian. Metode juga harus dipengaruhi oleh situasi dimana penelitian dilakukan dan akses informasinya.
4. Penelitian ilmiah harus menunjukkan serangkaian alasan yang utuh dan eksplisit. Dalam membuat penjelasan dan kesimpulan penelitian membutuhkan serangkaian alasan-alasan logis yang utuh, menyatu, dan persuasive.
5. Dapat direplikasi dan bersifat general. Temuan penelitian harus dapat dicek, di validasi, dan penelitian selanjutnya dibutuhkan untuk mengetahui apakah hasilnya dapat digeneralisasi untuk kepentingan yang lebih luas dan dalam konteks yang berbeda.
6. Menyampaikan hasil penelitian untuk mendapatkan saran, masukan, dan kritik yang professional. Salah satu ciri penelitian ilmiah adalah disebarkan secara luas dan ditunjukkan untuk direview oleh pihak yang berkepentingan. Maka, informasi terkait penelitian tersebut harus bisa diakses dan diusahakan dipublikasikan. Kritik dari masyarakat dan para professional dibutuhkan agar supaya kredibilitas temuan penelitian dapat divalidasi.

B. Kebenaran Ilmiah

Jujun S. Suriasumantri dalam filsafat ilmu sebuah pengantar populer keterkaitan ilmu, agama, dan seni mengemukakan kriteria kebenaran dalam tiga kategori kebenaran koherensi, kebenaran korespondensi, dan kebenaran pragmatism. Koherensi adalah konsistensi dalam keseluruhan suatu system pemikiran sedang konsistensi berlaku hanya dalam satu jalan pikir atau subsistem pemikiran.³ Dicontohkan bahwa $3+4 = 7$, $5+2 = 7$, $6+1 = 7$ adalah pernyataan yang benar karena ketiganya baik pernyataan dan kesimpulannya menunjukkan konsistensi dengan pernyataan dan kesimpulan terdahulu yang dianggap benar. Kebenaran korespondensi menyatakan bahwa suatu pernyataan adalah benar bila materi yang terkandung dalam pernyataan itu berkorespondensi (berhubungan) dengan objek yang dituju oleh pernyataan tersebut.⁴ Jika ada yang mengatakan Universitas Indonesia adalah sebuah perguruan tinggi ternama di Indonesia, maka pernyataan itu adalah benar karena faktanya bahwa Universitas Indonesia merupakan salah satu perguruan tinggi yang memiliki reputasi baik di Indonesia dan bukan nama sebuah makanan atau nama kabupaten.

Ketiga, kebenaran pragmatisme menganggap bahwa kebenaran suatu pernyataan diukur dengan kriteri apakah pernyataan tersebut bersifat fungsional dan berguna dalam kehidupan praktis. Sekiranya ada orang yang mengemukakan sebuah teori baru dalam Pendidikan, dan dari teori tersebut dapat dikembangkan metode mengajar baru yang secara signifikan meningkatkan prestasi belajar maka teori itu dianggap benar.⁵

Berbicara kebenaran ilmiah tentu saja tidak dapat dilepaskan dari serangkaian proses yang sistematis dan terstruktur dalam memperoleh pengetahuan. Ketika kita membahas proses yang sistematis tersebut maka kita akan dibawa ke metode penelitian yang berasaskan *logico-*

³ Jujun S. Suriasumantri, Filsafat Ilmu: Sebuah Pengantar Populer keterkaitan Ilmu, Agama, dan Seni (Jakarta: Pustaka Sinar Harapan, 2017), hlm. 105.

⁴ Jujun S. Suriasumantri, Filsafat Ilmu: Sebuah Pengantar Populer keterkaitan Ilmu, Agama, dan Seni (Jakarta: Pustaka Sinar Harapan, 2017), hlm. 105.

⁵ Ibid. hh. 107-109

hypothetico-verifikatif. Penelitian ilmiah merupakan salah satu cara yang dapat dilakukan oleh manusia untuk memperoleh pengetahuan. Pengetahuan yang cara mendapatkannya melalui proses penelitian inilah yang kemudian disebut pengetahuan ilmiah dan kebenarannya pun bersifat ilmiah.

C. Metode Ilmiah

Menurut Peter R. Senn dikutip JUJUN S. Suriasumantri metode ilmiah adalah cara atau prosedur untuk memperoleh pengetahuan ilmiah dengan langkah-langkah yang sistematis. Metodologi adalah pengetahuan tentang metode, atau lebih rinci lagi, kajian yang mempelajari peraturan-peraturan yang terdapat sebuah metode.⁶ Dalam penelitian ilmiah, peneliti perlu mempertimbangkan metode yang akan digunakan dalam penelitiannya. Metode berkaitan dengan strategi pengumpulan, analisis, dan interpretasi data. Sejumlah metode pengumpulan data perlu dipertimbangkan dan diatur secara sistematis agar hasil penelitiannya memiliki kredibilitas.

Pemilihan metode penelitian disesuaikan dengan tujuan peneliti. Apakah ingin menggali informasi secara detail seperti yang diinginkan atau peneliti ingin menganalisis jenis data yang berupa informasi numerik atau informasi tek, atau peneliti ingin menafsirkan hasil statistik atau ingin menafsirkan kecenderungan-kecenderungan dan pola-pola umum yang muncul berdasarkan data penelitian. Secara umum metode penelitian terbagi dalam tiga kategori yaitu; kuantitatif (*quantitative*), kualitatif (*qualitative*) dan campuran (*mixed methods*).

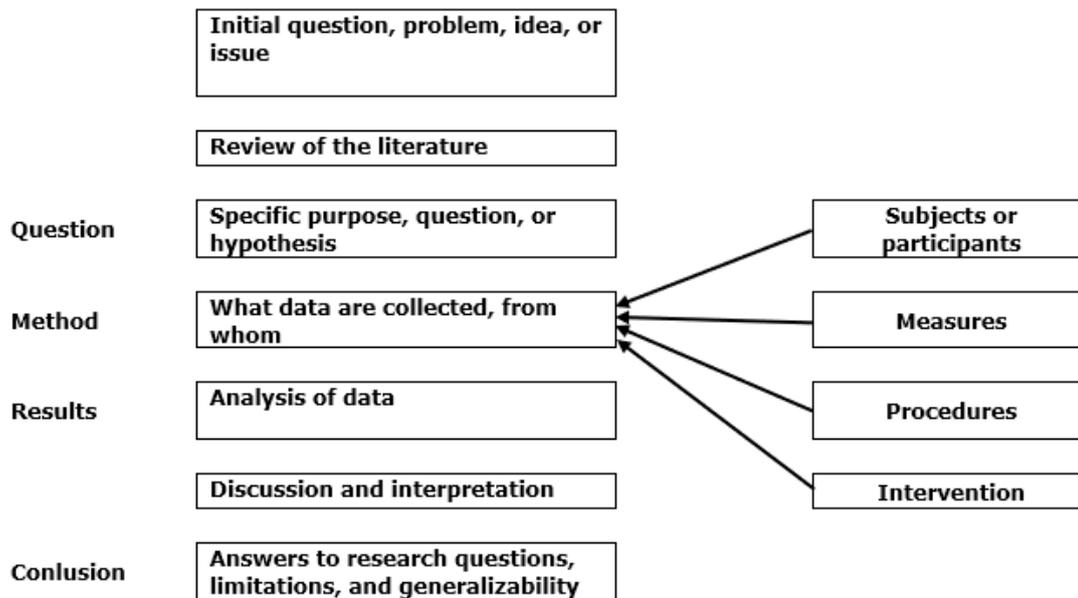
Metode kuantitatif, peneliti yang memilih metode ini ingin menguji suatu teori dengan cara memerinci hipotesis-hipotesis yang spesifik, lalu mengumpulkan data untuk mendukung atau membantah hipotesis tersebut.⁷ Penelitian kualitatif berusaha membangun makna tentang fenomena berdasarkan pandangan-pandangan dari para partisipan. Salah satu metode pengumpulan data dalam strategi ini

⁶ Jujun S. Suriasumantri, *Filsafat Ilmu: Sebuah Pengantar Populer keterkaitan Ilmu, Agama, dan Seni* (Jakarta: Pustaka Sinar Harapan, 2017), h. 276.

⁷ John W. Creswell. *Research Design: Pendekatan Kualitatif, Kuantitatif, dan Mixed*. Yogyakarta: Pustaka Pelajar, 2010), h. 27.

adalah mengobservasi perilaku partisipan dengan melibatkan diri secara langsung kedalam aktivitas partisipan. Pemilihan metode campuran ini dilakukan dengan asumsi bahwa dengan mengumpulkan berbagai jenis data yang dianggap terbaik dapat memberikan pemahaman yang menyeluruh tentang objek penelitian. Penelitian dengan pendekatan ini dapat dimulai dari survei secara luas untuk mendapatkan gambaran yang komprehensif mengenai objek yang akan diteliti. Selanjutnya dilakukan wawancara kualitatif secara terbuka untuk mendapatkan pandangan-pandangan dari partisipan.

Secara umum langkah-langkah penelitian ilmiah dapat dilihat pada gambar berikut;



Gambar 1: Langkah-langkah penelitian ilmiah. Source: McMilan

Penelitian ilmiah secara sederhana terdiri dari empat langkah yaitu dimulai dari sebuah pertanyaan yang harus dijawab, dilanjutkan dengan pengumpulan informasi atau data menggunakan metode-metode tertentu sesuai standar dan menganalisisnya. Berdasarkan hasil analisis maka diperoleh hasil penelitian yang dapat digunakan untuk membuat kesimpulan. Hasil itulah yang kemudian menjadi pengetahuan baru yang disebut ilmiah karena telah melalui proses dan tahapan-tahapan yang sistematis dan terukur untuk mendapatkan pengetahuan.

Wirawan mengemukakan bahwa riset dapat dibedakan menjadi dua kelompok yakni riset murni dan riset terapan. Riset murni adalah riset yang tujuannya hanya untuk mengembangkan ilmu pengetahuan, dengan melakukan penelitian untuk menciptakan teori-teori ilmu pengetahuan baru. Sementara riset terapan adalah riset yang bertujuan untuk menyelesaikan masalah yang dihadapi oleh masyarakat atau organisasi.⁸ Evaluasi adalah salah satu jenis riset terapan karena penelitian evaluasi lebih diarahkan untuk membuat keputusan atas suatu program atau kegiatan yang dievaluasi. Hasil dari evaluasi digunakan untuk mengambil keputusan apakah program atau kegiatan yang dimaksud akan dilanjutkan, distop, atau dilanjutkan dengan perbaikan-perbaikan.

D. Evaluasi

Sebagai salah satu riset terapan evaluasi memiliki tujuan yang berbeda dengan riset murni. Sebelum kita membahas lebih dalam mengenai evaluasi sebagai riset mari kita lihat beberapa definisi evaluasi yang dikemukakan oleh beberapa ahli. Menurut McMilan “*Evaluation research is directed toward making decisions about the effectiveness or desirability of a program. The goal is to make judgements about alternatives in decision-making situations*”.⁹ Penelitian evaluasi dilakukan untuk membuat keputusan mengenai efektifitas sebuah program atau kesesuaian program dengan perencanaan. Tujuannya untuk membuat keputusan terhadap objek yang dievaluasi. Apakah program dapat dilanjutkan, dihentikan, atau diteruskan dengan perbaikan-perbaikan.

Seperti yang dikatakan Merten “*evaluation is more typically associated with the need for information for decision making in specific setting, and research is more typically associated with generating new knowledge that can be transferred to other settings*”.¹⁰ Evaluasi secara khas lebih

⁸ Wirawan. *Evaluasi Teori, Model, Metodologi, Standar, Aplikasi, dan Profesi*. (Jakarta: Grafindo Persada, 2016), hlm. 2-3.

⁹ James H. McMilan, *Educational Research: Fundamentals for the Consumer*. (USA: Pearson Education, Inc. 2008), hlm. 16.

¹⁰ Donna M. Mertens, *Research and Evaluation in Educational and Psychology: Integrating Diversity with Quantitative, Qualitative, And Mixed Methods*, 2nd Edition (California: Sage Publications, Inc. 2005), p. 2.

dikaitkan dengan pengambilan keputusan di tempat tertentu, sementara riset lebih diasosiasikan untuk menghasilkan pengetahuan baru yang dapat di transfer ke tempat lain.

Tyler mengemukakan bahwa *“Evaluation is a process by which the values of an enterprise are ascertained, an appraisal of the degree to which the objectives of an institution are being realized. It is an integral part of the educational process. Evaluation is an ongoing, cyclical process that should lead to a refinement of objectives, learning experiences, and outcomes”*.¹¹ Evaluasi adalah sebuah proses dimana nilai organisasi harus ditingkatkan, target atau sasaran lembaga diwujudkan. Dan semuanya menjadi bagian yang terintegrasi dalam proses Pendidikan. Evaluasi adalah kegiatan yang berkelanjutan, prosesnya terus berulang yang pada akhirnya membawa kepada pemurnian tujuan, pengalaman belajar, dan hasil yang ingin dicapai.

Stufflebeam *“Evaluation means a study designed and conducted to assist some audience to assess an object’s merit and worth”*.¹² Evaluasi bermakna penelitian yang di desain dan dilakukan untuk membantu audien memberi nilai dan manfaat mengenai objek yang dievaluasi. Peneliti dalam hal memberi nilai melakukan penilaian kualitas dari objek yang di evaluasi. Serta memberi makna manfaat terhadap objek yang dievaluasi, yang berarti apakah suatu program atau kegiatan yang telah berjalan memiliki kontribusi positif bagi sasaran program atau tidak. Seluruh nilai dan makna manfaat tersebut dapat diidentifikasi setelah melalui proses evaluasi.

Wirawan mendefinisikan evaluasi sebagai riset untuk mengumpulkan, menganalisis dan menyajikan informasi yang bermanfaat mengenai objek evaluasi, menilainya dengan membandingkannya dengan indicator evaluasi dan hasilnya dipergunakan untuk mengambil keputusan mengenai nilai dan manfaat objek evaluasi.¹³

¹¹ Ralph W. Tyler, *Educational Evaluation*, (Boston: Kluwer Academic Publishers, 1989), pp. 90-91.

¹² Daniel L. Stufflebeam. *Evaluation Models Viewpoints on Educational and Human Services Evaluation Second Edition, Foundational Models for Century Program Evaluation** (New York: Kluwer Academic Publishers, 2002), p. 35.

¹³ Wirawan. *Evaluasi Teori, Model, Metodologi, Standar, Aplikasi, dan Profesi*. (Jakarta: Grafindo Persada, 2016), hlm. 9.

Suharsimi Arikunto mendefinisikan evaluasi sebagai kegiatan untuk mengumpulkan informasi tentang bekerjanya sesuatu, yang selanjutnya informasi tersebut digunakan untuk menentukan alternative yang tepat dalam mengambil sebuah keputusan.¹⁴

Dari beberapa definisi evaluasi yang telah dijabarkan di atas, evaluasi adalah proses pengumpulan data dan informasi secara terukur dan memaknainya sebagai bahan pertimbangan untuk membuat keputusan mengenai objek evaluasi. Secara umum sepakat bahwa akhir dari evaluasi itu untuk membuat keputusan terhadap objek evaluasi. Dalam hal ini, maka hasil evaluasi akan sangat menentukan apakah suatu program atau kegiatan akan dilanjutkan diberhentikan atau diperbaiki. Karena itulah yang kemudian membedakan antara penelitian murni dengan penelitian evaluasi.

*Evaluations are conducted on the merit and worth of programs in the public domain— programs that are, themselves, responses to prioritized individual and community needs that resulted from political decisions.*¹⁵

Evaluasi dilakukan berdasarkan nilai dan manfaat program bagi masyarakat-program-programnya yang dapat merespon terhadap prioritas individu dan kebutuhan komunitas yang dihasilkan dari keputusan politik. Merten menekankan bahwa evaluasi harus dilakukan dengan memahami nilai dan manfaat program yang akan dievaluasi sehingga hasil evaluasi dapat mengukur ketercapaian program dan pengaruhnya bagi masyarakat yang menjadi sasaran program.

*One of the most distinctive differences is that the purpose of research is generally considered to be the creation of new knowledge and theory construction, whereas the purpose of evaluation is to support decision making.*¹⁶ Satu hal yang menjadi pembeda adalah bahwa tujuan penelitian secara umum untuk menciptakan pengetahuan baru dan membangun teori baru, sementara tujuan evaluasi adalah untuk

¹⁴ Suharsimi Arikunto, Cepi Safruddin Abdul Jabar. *Evaluasi Program Pendidikan* (Jakarta: Bumi Aksara, 2009), hlm. 2.

¹⁵ Donna M. Mertens and Amy T. Wilson. *Program Evaluation Theory and Practice Second Edition: A Comprehensive Guide*, (New York: The Guilford Press A Division of Guilford Publications, Inc, 2019), P. 11.

¹⁶ *Ibid.* p. 11.

mendukung dalam membuat keputusan. Perbedaan yang mendasar antara evaluasi dan riset terletak pada tujuannya. Bahwa evaluasi memiliki tujuan memberikan data dan informasi yang akurat dan valid sehingga dapat digunakan sebagai dasar dalam membuat keputusan terkait program yang dievaluasi. Tetapi riset lebih kepada untuk membangun teori baru dan pengetahuan baru, yang berarti ada sifat generalisasi.

Merujuk pada tulisan Suharsimi Arikunto & Cepi Safruddin Abdul Jabar menyebutkan perbedaan antara penelitian dengan evaluasi program yaitu;

1. Dalam kegiatan penelitian, peneliti ingin mengetahui gambaran tentang sesuatu kemudian hasilnya dideskripsikan, sedangkan dalam evaluasi program, pelaksana ingin mengetahui seberapa tinggi mutu atau kualitas sesuatu sebagai hasil pelaksanaan program, setelah data terkumpul dibandingkan dengan kriteria atau standar tertentu.
2. Dalam kegiatan penelitian peneliti dituntun oleh rumusan masalah karena ingin mengetahui jawaban dari penelitiannya, sedangkan dalam evaluasi program pelaksana ingin mengetahui tingkat ketercapaian tujuan program, dan apabila tujuan belum tercapai, pelaksana ingin mengetahui dimana letak kekurangannya dan apa sebabnya.¹⁷

Walaupun banyak perbedaan antara penelitian dengan evaluasi tetapi tetap saja ada juga persamaannya. Ketika kita lihat dari sisi keilmuan maka keduanya jelas memiliki pijakan yang sama-sama ilmiah dan dalam mencapainya membutuhkan tahapan-tahapan yang sistematis dan terstruktur serta terukur sesuai standar perolehan ilmu pengetahuan. Mertens menguatkan dengan pernyataan bahwa *“Both make use of systematic inquiry methods to collect, analyze, interpret, and use data to understand, describe, predict, control, or empower. Evaluation is more typically associated with the need for information for*

¹⁷ Donna M. Mertens and Amy T. Wilson, *Program Evaluation Theory and Practice Second Edition: A Comprehensive Guide*, (New York: The Guilford Press A Division of Guilford Publications, Inc, 2019), P. 7.

decision making in a specific setting, and research is more typically associated with generating new knowledge that can be transformed to other settings".¹⁸ Kedua-duanya menggunakan metode penelitian yang sistematis untuk mengumpulkan, menganalisis, menginterpretasikan, dan menggunakan data-data yang terkumpul tersebut untuk memahami, menggambarkan, memprediksi, mengontrol, atau memberdayakan. Evaluasi lebih cenderung diperuntukan untuk kebutuhan mengumpulkan informasi guna membuat keputusan di tempat tertentu, sementara penelitian lebih menekankan pada perolehan pengetahuan yang dapat diterapkan di berbagai tempat lain selain dilokasi penelitian

Menurut Wirawan (2016) penelitain evaluasi memiliki karakteristik sebagai berikut.

Tabel 1: Karakteristik Penelitain Evaluasi. Sumber: Wirawan

Tujuan	Mengumpulkan informasi untuk menentukan nilai dan manfaat objek evaluasi, mengontrol, memperbaiki dan mengambil keputusan mengenai objek tersebut
Motivasi Peneliti	Dilakukan untuk memberikan kontribusi kepada solusi suatu problem tertentu.
Hukum-hukum versus deskripsi	Bersifat ideografik atau mendeskripsikan aktifitas khusus
Peran dari penjelasan	Dapat dilakukan tanpa menjelaskan mengapa program atau kegiatan yang dievaluasi baik atau buruk, bagaimana program menghasilkan pengaruhnya
Otonomi penelitian	Evaluator merumuskan problem penelitiannya berdasarkan pesanan klien atau penyandang dana
Property dari fenomena yang diakses	Evaluasi berupaya mengakses nilai intrinsik sesuatu (merit) dan manfaat (worth)
Metode penelitian	Pada prinsipnya semua metode penelitian dapat digunakan dalam evaluasi. Tetapi evaluasi pada umumnya dilakukan dalam waktu tertentu karena berkaitan dengan pengambilan keputusan
Model-model evaluasi	Evaluasi menggunakan model-model di samping metode penelitian
Penilaian	Tidak bebas nilai (penilaian berdasarkan kriteria yang telah ditentukan sebelumnya)
Generalisasi fenomena yang diteliti	Evaluasi umumnya tidak berupaya untuk menggeneralisasikan dari sampel ke populasi
Kompetensi peneliti	Evaluator harus mempunyai latar belakang multi disiplin ilmu pengetahuan

¹⁸ Donna M. Mertens, *Research and Evaluation in Educational and Psychology: Integrating Diversity with Quantitative, Qualitative, And Mixed Methods*, 2nd Edition (California: Sage Publications, Inc. 2005), p. 2.

RINGKASAN

Kriteria kebenaran terbagi dalam tiga kategori, yaitu; kebenaran koherensi, kebenaran korespondensi, dan kebenaran pragmatism.

Koherensi adalah konsistensi dalam keseluruhan suatu system pemikiran sedang konsistensi berlaku hanya dalam satu jalan pikir atau subsistem pemikiran. Kebenaran pragmatisme menganggap bahwa kebenaran suatu pernyataan diukur dengan kriteri apakah pernyataan tersebut bersifat fungsional dan berguna dalam kehidupan praktis. Kebenaran korespondensi menyatakan bahwa suatu pernyataan adalah benar bila materi yang terkandung dalam pernyataan itu berkorespondensi (berhubungan) dengan objek yang dituju oleh pernyataan tersebut.

Riset murni adalah riset yang tujuannya hanya untuk mengembanangkan ilmu pengetahuan dengan melakukan penelitian untuk menciptakan teori-teori ilmu pengetahuan baru.

Evaluasi adalah proses pengumpulan data dan informasi secara teruktur dan memaknainya sebagai bahan pertimbangan untuk membuat keputusan mengenai objek evaluasi.

Latihan 1

1. Jelaskan definisi evaluasi menurut saudara!
2. Jelaskan definisi riset menurut saudara!
3. Sebutkan langkah-langkah penelitian secara umum!
4. Jelaskan perbedaan mendasar antara penelitian dan evaluasi!
5. Jelaskan kriterian kebenaran korespodensi!

Pembahasan Soal Latihan

1. Evaluasi adalah proses pengumpulan data dan informasi secara terstruktur dan memaknainya sebagai bahan pertimbangan untuk membuat keputusan mengenai objek evaluasi.
2. Riset murni adalah riset yang tujuannya hanya untuk mengembangkan ilmu pengetahuan dengan melakukan penelitian untuk menciptakan teori-teori ilmu pengetahuan baru.
3. Penelitian ilmiah secara sederhana terdiri dari empat langkah yaitu dimulai dari sebuah pertanyaan yang harus dijawab, dilanjutkan dengan pengumpulan informasi atau data menggunakan metode-metode tertentu sesuai standar dan menganalisisnya. Berdasarkan hasil analisis maka diperoleh hasil penelitian yang dapat digunakan untuk membuat kesimpulan. Hasil itulah yang kemudian menjadi pengetahuan baru yang disebut ilmiah karena telah melalui proses dan tahapan-tahapan yang sistematis dan terukur untuk mendapatkan pengetahuan.
4. Perbedaan antara penelitian dengan evaluasi program yaitu; Dalam kegiatan penelitian, peneliti ingin mengetahui gambaran tentang sesuatu kemudian hasilnya dideskripsikan, sedangkan dalam evaluasi program, pelaksana ingin mengetahui seberapa tinggi mutu atau kualitas sesuatu sebagai hasil pelaksanaan program, setelah data terkumpul dibandingkan dengan kriteria atau standar tertentu.

Jika dalam kegiatan penelitian peneliti dituntun oleh rumusan masalah karena ingin mengetahui jawaban dari penelitiannya, maka dalam evaluasi program, peneliti ingin mengetahui tingkat ketercapaian tujuan program, dan apabila tujuan belum tercapai, pelaksana ingin mengetahui dimana letak kekurangannya dan apa sebabnya.

5. Kebenaran korespondensi menyatakan bahwa suatu pernyataan adalah benar bila materi yang terkandung dalam pernyataan itu berkorespondensi (berhubungan) dengan objek yang dituju oleh pernyataan tersebut. Sekiranya ada pernyataan bahwa gula itu manis dan ketika coba dikecap ternyata manis, maka pernyataan itu dianggap benar karena ada kesesuaian antara pernyataan dengan fakta yang dimaksud.

Tes Formatif Pilihan Berganda

1. Suatu pengetahuan yang diperoleh melalui proses yang sistematis dan logis dengan menggunakan metode tertentu yang terstruktur disebut?
 - a. Teori
 - b. Evaluasi
 - c. Penelitian
 - d. Pengetahuan ilmiah
2. Penilitain evaluasi memiliki karakteristik sebagai berikut, kecuali...?
 - a. Penilaian evaluasi tidak berdasarkan kriteria yang telah ditentukan.
 - b. Rumusan problem penelitiannya berdasarkan pesanan klien.
 - c. Menggunakan model-model di samping metode penelitian.
 - d. Memberikan kontribusi kepada solusi suatu problem tertentu.
3. Evaluasi dikategorikan sebagai riset terapan karena...?
 - a. Berupaya untuk menggeneralisasikan dari sampel ke populasi.
 - b. Bertujuan untuk menyelesaikan masalah yang dihadapi oleh masyarakat atau organisasi.
 - c. Dilakukan untuk membuat kesimpulan berdasarkan data untuk dapat diterapkan di wilayah lain.
 - d. Menggunakan data yang terkumpul untuk memahami, mendeskripsikan, memperkirakan, atau mengontrol fenomena.
4. Berikut ini adalah ciri-ciri pernyataan tentang kriteria kebenaran koherensi, kecuali...!
 - a. Seorang anak kecil menceritakan kebingungannya setelah mendengar cerita temannya bahwa $3 + 2 = 5$, $1 + 4 = 5$, $8 - 3 = 5$, $7 - 2 = 5$, $6 - 1 = 5$.
 - b. Ketika pernyataan dan kesimpulan yang ditariknya adalah konsisten dengan pernyataan dan kesimpulan ter-dahulu yang dianggap benar.
 - c. Teknologi informasi memiliki peran penting dalam menunjang keberlangsungan proses pembelajaran di masa pandemi.
 - d. “semua manusia akan mati”, “si Fulan adalah manusia”, maka “si Fulan akan mati”.

5. Ciri khas penelitian murni dapat diidentifikasi melalui karakter sebagai berikut!
- Dapat direplikasi dan bersifat general.
 - Otonomi penelitian tergantung penyandang dana.
 - Data yang terkumpul dibandingkan dengan kriteria atau standar tertentu.
 - Peneliti ingin mengetahui seberapa tinggi mutu atau kualitas sesuatu sebagai hasil pelaksanaan program.

Kunci jawaban dan pedoman penskoran

- D
- A
- B
- C
- A

Bobot

Latihan soal essay 75 % dengan bobot per soal 15.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{50}{75} \times 100 = \mathbf{66,67}$$

Tes Formatif Pilihan Berganda 25% dengan bobot per soal 5.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{20}{25} \times 100 = \mathbf{80}$$

$$66,67 \times 75\% = \mathbf{50}$$

$$80 \times 25\% = \mathbf{20}$$

$$50 + 20 = \mathbf{70}$$

$$\text{Nilai akhir} = \mathbf{70}$$

MODUL 2

TAHAPAN EVALUASI, JENIS RISET DAN EVALUASI VS RISET

Metode Pembelajaran	Estimasi waktu	Capaian Pembelajaran
-Presentasi virtual -Diskusi Virtual interaktif -Tanya jawab	120 menit X 1	-Mahasiswa memahami Tahapan evaluasi, Jenis riset dan Evaluasi vs riset

Materi 2: Tahapan Evaluasi, Jenis Riset dan Evaluasi vs Riset.

A. Tahapan Evaluasi

Sama halnya dengan penelitian pada umumnya, penelitian evaluasi juga memiliki tahapan-tahapan yang harus dicermati oleh evaluator sehingga hasil yang diharapkan dapat tercapai. Suharsimi Arikunto dan Cipi Safruddin Abdul jabar (2009) memberikan deskripsi langkah-langkah dalam evaluasi program kedalam 3 tahap meliputi: persiapan evaluasi, pelaksanaan evaluasi, dan tahap monitoring pelaksanaan program. Pada setiap tahapan evaluasi membutuhkan perhatian dan penanganan yang berbeda-beda sesuai tingkatannya. Tahap persiapan evaluasi program merupakan langkah awal dimana seorang evaluator harus menyiapkan penyusunan evaluasi, penyusunan instrument evaluasi, validasi instrument evaluasi, menentukan jumlah sampel yang diperlukan dalam kegiatan evaluasi, dan penyamaan persepsi antar evaluator sebelum pengambilan data.¹⁹ Penyusunan evaluasi adalah proses menentukan model evaluasi apa yang akan digunakan dalam kegiatan evaluasi yang akan dilaksanakan. Apakah menggunakan model CIPP yang dikembangkan Stufflebeam, model

¹⁹ Suharsimi Arikunto, Cipi Safruddin Abdul Jabar, *Evaluasi Program Pendidikan* (Jakarta: Bumi Aksara, 2009), hlm. 108.

Goal Free, model kesenjangan, atau model lain yang sesuai dengan tujuan evaluasi itu sendiri.

Kegiatan penyusunan instrument evaluasi menjadi salah satu tahap yang sangat krusial karena berkenaan dengan alat yang akan digunakan untuk mengukur sebuah program apakah sudah sesuai target atau belum, sesuai standar atau tidak, sudah mencapai sasaran yang diharapkan atau masih dibawah sasaran. Seorang evaluator harus cermat dalam menyusun instrumen evaluasi, maka diperlukan proses pengukuran validasi dan reliabilitas instrument untuk memastikan bahwa instrument layak dan dapat digunakan. Instrument yang valid berarti dapat digunakan untuk mengukur objek yang akan diukur. Jika kita ingin mengukur kecepatan orang berlari tentu kita harus menggunakan speedometer, untuk mengukur berat badan menggunakan timbangan, dan untuk mengukur suhu tubuh tentu menggunakan thermometer. Sehingga alat yang digunakan untuk mengukur objek yang dimaksud sesuai dengan peruntukannya. Reliabilitas berarti instrument yang akan digunakan memiliki kendalan untuk mengukur objek dimaksud. Instrument harus dapat digunakan untuk mengukur objek yang sama dalam beberapa periode yang berbeda tetapi menghasilkan data yang relative sama. Ketika kita ingin mengukur Panjang tentu alat yang digunakan adalah meteran, tetapi meteran yang terbuat dari bahan fleksibel seperti karet tidak memiliki reliabilitas karena sifatnya yang dapat berubah-ubah. Meteran akan dapat dianggap reliabel jika memiliki konsistensi, misalnya yang terbuat dari bahan yang stabil seperti kayu, besi atau bahan lain yang memiliki sifat tetap. Masih dalam tahap persiapan, yang tidak kalah pentingnya adalah penentuan jumlah sampel yang akan menjadi responden dalam evaluasi. Langkah terakhir di tahap persiapan adalah penyamaan persepsi antar evaluator yang terlibat sebelum pengambilan data. Hal ini penting untuk menjaga ritme agar tetap terarah dan proses pengambilan data berjalan lancar dan sesuai rencana yang telah disusun.

Tahap kedua, adalah pelaksanaan evaluasi program. Evaluasi program terdiri dari evaluasi reflektif, evaluasi rencana, evaluasi proses, dan

evaluasi hasil. Evaluasi reflektif digunakan untuk mengevaluasi kurikulum sebagai suatu ide.²⁰ Proses reflektif merujuk pada kegiatan melihat ulang sebuah ide dan menelaah lebih dalam hal-hal yang telah direncanakan dan akan dilaksanakan. Hal ini menjadi penting dalam rangka memastikan segala ide dapat dituangkan dalam setiap kegiatan yang telah direncanakan. Reflektif sekaligus menjadi ajang kontemplasi secara komprehensif sebelum gagasan diimplementasikan.

Evaluasi rencana berkenaan dengan persyaratan-persyaratan program sebagai rencana seperti format, keterbacaan, hubungan antar komponen, organisasi vertical dan horizontal dari pengalaman belajar, biasanya merupakan hal yang menuntut perhatian evaluator pada waktu melakukan evaluasi program sebagai suatu rencana.²¹ Guna memastikan segala perencanaan telah dipersiapkan dengan matang dan tidak ada yang tertinggal, maka rencana itupun perlu dievaluasi. Jika evaluasi reflektif lebih menekankan pada ide atau gagasan yang dievaluasi, pada evaluasi rencana ingin mengetahui apakah rencana yang sudah disusun terkait dengan program yang akan dilaksanakan sudah memenuhi standar. Disini juga akan dilihat apakah perencanaan yang dibuat masuk akal dan dapat menunjukkan prediksi tingkat keberhasilan program jika dilaksanakan. Evaluasi rencana memungkinkan evaluator memperhitungkan segala kemungkinan yang bakal terjadi dalam implementasi program. Sehingga evaluator dapat menyiapkan alternatif solusi jika skenario awal meleset atau tidak berjalan sesuai rencana.

Evaluasi proses adalah ketika sebuah ide atau gagasan yang telah dirumuskan dan disiapkan perencanaan untuk dilaksanakan lalu dieksekusi dilapangan sebagai program atau kegiatan yang berjalan. Evaluator memiliki tugas dan tanggungjawab untuk mengecek apakah program berjalan sesuai rencana atau tidak. Pada tahap ini pula akan dilihat kontribusi positif dari implementasi program yang dimaksud. Evaluasi proses memungkinkan evaluator memberikan *treatmen*

²⁰ Suharsimi Arikunto, Cepi Safruddin Abdul Jabar. *Evaluasi Program Pendidikan: Pedoman Teoritis Praktis bagi Mahasiswa dan Praktisi Pendidikan*. (Jakarta: Bumi Aksara, 2009), h. 111.

²¹ Ibid. h. 112.

terhadap program yang sedang berjalan jika terindikasi adanya distorsi atau ketidaksesuaian kegiatan dengan rencana atau standar yang telah ditentukan sebelumnya.

Berikutnya adalah evaluasi hasil, secara umum sebagian kita memahami bahwa kegiatan evaluasi lebih ditekankan pada hasil yang didapat dari pelaksanaan program. Hal ini bisa dimaklumi karena menjadi salah satu indikator kesuksesan atau kegagalan implementasi program. Ketika hasil kegiatan atau program menunjukkan kontribusi positif bagi sasaran program maka dapat dikatakan bahwa program tersebut berhasil. Atau jika program telah berjalan sesuai rencana dan hasilnya sesuai yang diharapkan, maka program secara relatif telah berhasil.

Setelah melalui tahap persiapan, maka langkah berikutnya adalah melaksanakan kegiatan evaluasi. Dalam melaksanakan kegiatan evaluasi seorang evaluator akan melakukan pengumpulan data sebagai bahan untuk dianalisis. Proses pengumpulan data ini membutuhkan kecermatan agar informasi yang terkumpul memiliki validitas yang memadai. Sehingga hasil analisisnya dapat dipertanggungjawabkan dan bisa menjadi dasar untuk membuat keputusan yang tepat. Beberapa hal yang perlu mendapat perhatian dalam proses pengumpulan data dan informasi menurut Suharsimi Arikunto dan cepi Safruddin adalah; 1) pengambilan data dengan tes, 2) pengambilan data dengan Observasi, 3) pengambilan data dengan angket, 4) pengambilan data dengan wawancara, 5) pengambilan data dengan metode analisis dokumen dan artifak, 6) pengambilan data dengan Teknik lainnya.

Pengumpulan data evaluasi dengan tes dapat dilakukan dengan cara;

1. Buka buku versus tutup buku,
2. Tes dumumkan versus tes dirahasiakan,
3. Tes lisan atau tes tertulis, dan
4. Tes tindakan atau praktik.²²

²² Suharsimi Arikunto, Cepi Safruddin Abdul Jabar. *Evaluasi Program Pendidikan* (Jakarta: Bumi Aksara, 2009), h. 113.

Pengumpulan data dengan observasi dapat dikatakan sebuah proses yang cukup rumit karena seorang evaluator perlu mengerahkan segala kemampuan dalam melihat, mendengar, dan mengingat segala hal yang diamati pada saat program berjalan. Evaluator masuk dan terlibat langsung baik pasif maupun aktif dalam kegiatan berlangsung. Terlibat secara Pasif artinya ketika evaluator melakukan observasi kegiatan belajar siswa didalam kelas maka hanya mengamati kegiatan berjalan dan membuat catatan-catatan serta mengingat hal-hal yang terjadi dalam kelas selama kegiatan belajar berlangsung. Terlibat secara aktif menunjukkan bahwa evaluator bisa saja terlibat langsung dalam pelaksanaan program. Sebagai contoh seorang guru yang sedang mengajar didalam kelas bisa saja sekaligus menjadi evaluator proses pembelajaran yang sedang dilaksanakan.

Proses pengambilan data dengan angket mengharuskan evaluator mengembangkan daftar pertanyaan atau instrument yang akan digunakan untuk mengukur keterlaksanaan program yang sedang berjalan. Dengan angket seorang evaluator tidak harus hadir langsung ke tempat kegiatan. Sementara pengumpulan data dengan wawancara dapat dilakukan oleh evaluator dengan mendatangi para informan atau responden secara langsung. Dalam kegiatan wawancara seorang evaluator dapat merekam hasil wawancara atau juga bisa mencatat hasil wawancara dengan responden. Hal yang perlu diperhatikan dalam wawancara adalah agar evaluator memosisikan diri secara objektif, tidak boleh memengaruhi responden dalam memberikan jawaban.

Pengumpulan data berikutnya dapat dilakukan dengan studi dokumen dan artifak. Studi dokumen dapat dilakukan dengan mempelajari catatan harian, notulen rapat, surat-surat, dokumen digital, rekaman digital, dll.

Tahap ketiga, monitoring pelaksanaan evaluasi. Pada tahap akhir kegiatan evaluasi ini seorang evaluator akan melakukan pemantauan terhadap aktifitas evaluasi yang berguna untuk mengetahui kesesuaian pelaksanaan program dengan rencana program dan untuk mengetahui

seberapa pelaksanaan program yang sedang berlangsung dapat diharapkan akan menghasilkan perubahan yang diinginkan.²³

Dalam proses pemantauan pelaksanaan evaluasi, evaluator berusaha mendeteksi secara lebih awal bahwa program telah berjalan sesuai dengan rencana. Jika terjadi hal-hal yang melenceng dari rencana program maka akan dapat diidentifikasi dan segera dilakukan tindakan atau treatment yang tepat agar berjalan pada jalur yang benar. Maka pada kegiatan inilah perlu memastikan standar operasional prosedur yang sudah dibuat benar-benar diterapkan. Selain mendeteksi pelaksanaan program sudah sesuai rencana, seorang evaluator juga ingin mendapat informasi tentang program yang sedang dijalankan benar-benar memberi dampak positif terhadap sasaran program. Ambil contoh program bantuan operasional sekolah (BOS) yang digulirkan pemerintah. Salah satu tujuannya adalah untuk memutus rantai angka putus sekolah bagi siswa yang kurang mampu. Maka, pada tahap pemantauan ini akan dilihat pengaruh positif program BOS tersebut dalam mereduksi angka putus sekolah. Jika sebelum ada BOS angka putus sekolah misalnya 20% dari total peserta didik yang terdaftar, ternyata dengan adanya BOS angka putus sekolah turun menjadi 5% maka dapat disimpulkan bahwa program BOS memiliki dampak positif terhadap perubahan yang diinginkan dari program tersebut.

Kegiatan pemantauan tentu saja membutuhkan teknik dan alat pemantauan agar aktifitas memantau dapat berjalan dengan lancar dan mendapatkan hasil optimal. Menurut Suharsimi Arikunto dan Cepi Safruddin Abdul Jabar menyebutkan bahwa teknik dan alat yang dibutuhkan dalam pemantauan tersebut adalah;

1. Teknik pengamatan partisipatif dengan menggunakan lembar pengamatan, catatan lapangan, dan alat perekam elektronik.
2. Teknik wawancara, secara bebas atau terstruktur dengan alat pedoman wawancara dan perekam wawancara.

²³ Suharsimi Arikunto, Cepi Safruddin Abdul Jabar. *Evaluasi Program Pendidikan* (Jakarta: Bumi Aksara, 2009), hlm. 123.

3. Teknik pemanfaatan dan analisis data dokumentasi seperti daftar hadir, satuan pelajaran, hasil karya siswa, hasil karya guru, dan sebagainya.²⁴

Seluruh rangkaian kegiatan pemantauan tentu membutuhkan sebuah perencanaan yang matang untuk mendapatkan hasil maksimal. Dalam perencanaan pemantauan perlu mempertimbangkan hal-hal sebagai berikut.

1. Perumusan tujuan pemantauan.
2. Penetapan sasaran pemantauan, apa yang menjadi objek pemantauan.
3. Penjabaran data yang dibutuhkan pemantauan, penjabaran dari sasaran.
4. Penyiapan metode/alat pemantauan sesuai dengan sifat objek dan sumber atau jenis datanya.
5. Perancangan analisis data pemantauan dan pemaknaannya dengan berorientasi pada tujuan pemantauan.²⁵

B. Jenis Riset

Secara umum jenis-jenis riset dapat dikategorisasi sebagai berikut;

Tabel 2: Jenis-jenis Penelitian, source: James McMilan, 2008.

Type	Purpose	Example
Quantitative	To describe phenomena numerically to answer specific questions or hypotheses	The relationship between amount of homework and student achievement
Qualitative	To provide rich narrative descriptions of phenomena that enhance understanding	Observation of school renewal teams to understand the role of parents

²⁴ Suharsimi Arikunto, Cepi Safruddin Abdul Jabar, *Evaluasi Program Pendidikan* (Jakarta: Bumi Aksara, 2009), hlm. 124.

²⁵ *Ibid.* h. 125.

Type	Purpose	Example
Nonexperimental	To describe, compare, and predict phenomena without actively manipulating factors that influence the phenomena	Determine the relationship between socioeconomic status and students attitude
Experimental	To determine the casual relationship between two or more phenomena by direct manipulation of factors that influence the phenomena	Determine which of two approaches to teaching science results in the highest student achievement
Basic	To increase knowledge and understanding of phenomena	Understand how feedback affects motivation or learning style of adolescent
Applied	To solve practical educational problem	Determine the best approach for training teacher
Action	To improve practice in a school or classroom	Determine which grouping procedure results in the highest achievement
Evaluation	To make a decision about a program or activity	Decide whether to keep or phases out a prekindergarten program

C. Evaluasi Vs Riset

Menurut Wirawan (2016) Riset dan Evaluasi dapat dibedakan dari beberapa aspek sebagai berikut.

Tabel 3: Perbedaan Riset dengan Evaluasi

	Penelitian	Evaluasi
Tujuan	Membuktikan adanya atau kebenaran saintifik dan menciptakan teori mengenai kebenaran fenomena ilmu pengetahuan	Mengumpulkan informasi untuk menentukan nilai dan manfaat objek evaluasi, mengontrol, memperbaiki dan mengambil keputusan mengenai objek tersebut
Motivasi Peneliti	Dilakukan untuk mencapai kepuasan keingintahuan peneliti	Dilakukan untuk memberikan kontribusi kepada solusi suatu problem tertentu.
Hukum-hukum versus deskripsi	Berorientasi pada kesimpulan bersifat nomotetik atau memberikan hukum	Bersifat ideografik atau mendeskripsikan aktifitas khusus.

	Penelitian	Evaluasi
Peran dari penjelasan	Penjelasan riset memerlukan mengaitkannya dengan hokum, teori atau dalil	Dapat dilakukan tanpa menjelaskan mengapa program atau kegiatan yang dievaluasi baik atau buruk, bagaimana program menghasilkan pengaruhnya
Otonomi penelitian	Penelitian memformulasikan problem secara independent atau bebas dari pengaruh luar	Evaluator merumuskan problem penelitiannya berdasarkan pesanan clien atau penyandang dana
Property dari fenomena yang diakses	Penelitian berupaya mengakses kebenaran saintifik	Evaluasi berupaya mengakses nilai intrinsil sesuatu (merit) dan manfaat (worth)
Metode penelitian	Survey, ex post facto, ekperimen, naturalistic, action reseach, dll	Pada prinsipnya semua metode penelitian dapat digunakan dalam evaluasi. Tetapi evaluasi pada umumnya dilakukan dalam waktu tertentu karena berkaitan dengan pengambilan keputusan
Model-model evaluasi	Tidak menggunakan model-model seperti Goal Based, Goal free, CIPP dst	Evaluasi menggunakan model-model di samping metode penelitian
Penilaian	Bebas nilai	Tidak bebas nilai (penilaian berdasarkan kriteria yang telah ditentukan sebelumnya)
Generalisasi fenomena yang diteliti	Penarikan kesimpulan dari hasil penelitian yang berdasarkan sampel ke populasi penelitian	Evaluasi umumnya tidak berupaya untuk menggeneralisasikan dari sampel ke populasi
Kompetensi peneliti	Peneliti murni harus memiliki latar belakang Pendidikan bidang ilmunya	Evaluator harus mempunyai latar belakang multi disiplin ilmu pengetahuan

RINGKASAN

Langkah-langkah dalam evaluasi program meliputi: persiapan evaluasi, pelaksanaan evaluasi, dan tahap monitoring pelaksanaan program.

Tahap persiapan evaluasi program merupakan langkah awal dimana seorang evaluator harus menyiapkan penyusunan evaluasi, penyusunan instrument, validasi instrument, menentukan jumlah sampel, dan penyamaan persepsi antar evaluator sebelum pengambilan data.

Dalam melaksanakan evaluasi, seorang evaluator akan melakukan pengumpulan data sebagai bahan untuk dianalisis. Beberapa hal yang perlu mendapat perhatian dalam proses pengumpulan data dan informasi adalah; 1) pengambilan data dengan tes, 2) pengambilan data dengan Observasi, 3) pengambilan data dengan angket, 4) pengambilan data dengan wawancara, 5) pengambilan data dengan metode analisis dokumen dan artifak, 6) pengambilan data dengan Teknik lainnya.

Pada tahap monitoring, seorang evaluator akan melakukan pemantauan terhadap aktifitas evaluasi untuk mengetahui kesesuaian pelaksanaan program dengan rencana program dan untuk mengetahui pelaksanaan program yang sedang berlangsung dapat menghasilkan perubahan yang diinginkan.

Latihan 2

1. Jelaskan hal-hal yang perlu dilakukan dalam tahap perencanaan evaluasi?
2. Jelaskan apa yang harus dilakukan evaluator dalam tahap pelaksanaan evaluasi?
3. Sebutkan tujuan monitoring pada kegiatan evaluasi!
4. Sebutkan jenis-jenis penelitian?

5. Mengapa instrument yang akan digunakan harus diuji validitas dan reliabilitasnya?

Pembahasan Latihan Soal 2

1. Tahap persiapan evaluasi program merupakan langkah awal dimana seorang evaluator harus menyiapkan penyusunan evaluasi, penyusunan instrument, validasi instrument, menentukan jumlah sampel, dan penyamaan persepsi antar evaluator sebelum pengambilan data.
2. Tahap pelaksanaan evaluasi, seorang evaluator akan melakukan pengumpulan data sebagai bahan untuk dianalisis. Beberapa hal yang perlu mendapat perhatian dalam proses pengumpulan data dan informasi adalah; 1) pengambilan data dengan tes, 2) pengambilan data dengan Observasi, 3) pengambilan data dengan angket, 4) pengambilan data dengan wawancara, 5) pengambilan data dengan metode analisis dokumen dan artifak, 6) pengambilan data dengan Teknik lainnya.
3. Seorang evaluator akan melakukan pemantauan terhadap aktifitas evaluasi yang berguna untuk mengetahui kesesuaian pelaksanaan program dengan rencana program dan untuk mengetahui seberapa pelaksanaan program yang sedang berlangsung dapat diharapkan akan menghasilkan perubahan yang diinginkan. evaluator berusaha mendeteksi secara lebih awal bahwa program telah berjalan sesuai dengan rencana. Jika terjadi hal-hal yang melenceng dari rencana program maka akan dapat diidentifikasi dan segera dilakukan tindakan atau treatment yang tepat agar berjalan pada jalur yang benar. Maka pada kegiatan inilah perlu memastikan standar operasional prosedur yang sudah dibuat benar-benar diterapkan. Selain mendeteksi pelaksanaan program sudah sesuai rencana, seorang evaluator juga ingin mendapat informasi tentang program yang sedang dijalankan benar-benar memberi dampak positif terhadap sasaran program.
4. *Quantitative, Qualitative, Nonexperimental, Experimental, Basic, Applied, Action, Evaluation.*
5. Uji validitas diperlukan untuk memastikan bahwa instrument yang akan digunakan dapat digunakan untuk mengukur objek yang akan diukur. Jika kita ingin mengukur kecepatan orang berlari tentu kita

harus menggunakan speedometer, untuk mengukur berat badan menggunakan timbangan, dan untuk mengukur suhu tubuh tentu menggunakan thermometer. Sehingga alat yang digunakan untuk mengukur objek yang dimaksud sesuai dengan peruntukannya. Reliabilitas berarti instrument yang akan digunakan memiliki kendali untuk mengukur objek dimaksud. Instrument harus dapat digunakan untuk mengukur objek yang sama dalam beberapa periode yang berbeda tetapi menghasilkan data yang relative sama. Ketika kita ingin mengukur Panjang tentu alat yang digunakan adalah meteran, tetapi meteran yang terbuat dari bahan fleksibel seperti karet tidak memiliki reliabilitas karena sifatnya yang dapat berubah-ubah. Meteran akan dapat dianggap reliabel jika memiliki konsistensi, misalnya yang terbuat dari bahan yang stabil seperti kayu, besi atau bahan lain yang memiliki sifat tetap.

Tes Formatif Pilihan Berganda

1. Dalam tahap persiapan evaluasi program seorang evaluator harus menyiapkan hal-hal berikut, kecuali...!
 - a. Penyusunan instrument evaluasi.
 - b. Validasi instrumen evaluasi.
 - c. Menentukan jumlah sampel.
 - d. Melakukan penggalan data.
2. Tahapan evaluasi program yang benar terdiri dari...!
 - a. Persiapan evaluasi, pelaksanaan evaluasi, dan tahap monitoring pelaksanaan program.
 - b. persiapan evaluasi, pelaksanaan evaluasi, tahap monitoring pelaksanaan program, dan observasi.
 - c. persiapan evaluasi, pelaksanaan evaluasi, tahap monitoring pelaksanaan program, dan pengambilan data dengan tes.
 - d. persiapan evaluasi, pelaksanaan evaluasi, dan tahap monitoring pelaksanaan program, dan pengambilan data dengan wawancara.
3. Mengapa evaluasi rencana dibutuhkan?
 - a. Mengetahui apakah rencana yang disusun sudah standar, logis dan dapat memprediksi tingkat keberhasilan program.

- b. Mengecek apakah program berjalan sesuai rencana dan apa kontribusi positif dari implementasi program.
 - c. Memberikan *treatmen* terhadap program yang terindikasi menyimpang dari rencana atau standar.
 - d. Mengecek kegiatan atau program telah berkontribusi positif terhadap sasaran program.
4. Pernyataan yang menunjukkan karakteristik penelitian adalah...!
- a. Dilakukan dalam waktu tertentu dan terkait dengan pengambilan keputusan.
 - b. Berupaya mengakses nilai intrinsil sesuatu (*merit*) dan manfaat (*worth*).
 - c. Dilakukan untuk mencapai kepuasan keingintahuan penelitiannya.
 - d. Bersifat ideografik atau mendeskripsikan aktifitas khusus.
5. Pemantauan pelaksanaan evaluasi dapat dilakukan dengan berbagai Teknik sebagai berikut, kecuali...!
- a. Observasi/partisipasi
 - b. Analisis dokumen
 - c. Wawancara
 - d. Motivasi

Kunci jawaban dan pedoman penskoran

1. D
2. A
3. A
4. C
5. D

Bobot

Latihan soal essay 75 % dengan bobot per soal 15.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{50}{75} \times 100 = \mathbf{66,67}$$

Tes Formatif Pilihan Berganda 25% dengan bobot per soal 5.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{20}{25} \times 100 = \mathbf{80}$$

$$66,67 \times 75\% = \mathbf{50}$$

$$80 \times 25\% = \mathbf{20}$$

$$50 + 20 = \mathbf{70}$$

$$\text{Nilai akhir} = \mathbf{70}$$

MODUL 3

PERAN DAN TUJUAN EVALUASI

Metode Pembelajaran	Estimasi waktu	Capaian Pembelajaran
-Presentasi virtual -Diskusi Virtual interaktif -Tanya jawab	120 menit X 1	- Mahasiswa memahami Peran dan tujuan evaluasi, Batasan dalam evaluasi: 1. Peran dan Tujuan Evaluasi 2. Arti evaluasi 3. Fungsi evaluasi

Materi 3. Peran dan tujuan evaluasi, Fungsi evaluasi.

A. Peran Dan Tujuan Evaluasi

Evaluasi menjadi bagian yang sangat penting dalam setiap kegiatan atau program. Evaluasi dapat dikatakan sebagai bentuk control terhadap pelaksanaan program atau kegiatan yang dilaksanakan oleh Lembaga ataupun oleh orang-perorangan. Dalam hal ini evaluasi memiliki peran strategis yang dapat memberi sumbangan terhadap keberhasilan suatu program. Tentu jika evaluasi dilaksanakan secara benar mengikuti prosedur evaluasi yang telah ditetapkan.

Dikatakan memiliki peran strategis karena dalam evaluasi akan diperoleh informasi dan data mengenai sisi-sisi program yang telah berjalan sesuai rencana dan standar. Serta memberi informasi tentang hal-hal yang melenceng dari rencana atau tidak sesuai harapan. Disini juga akan diperoleh informasi mengenai seberapa tinggi pencapaian program jika dibandingkan dengan target yang telah ditentukan. Bagaimana kontribusi positif program terhadap lingkungan masyarakat yang menjadi sasaran program. Semua informasi dan data yang dimaksud merupakan modal bagi evaluator dan pelaksana kegiatan atau program dan pihak-pihak yang berkepentingan untuk

menjadi dasar dalam membuat keputusan terkait program yang sedang dijalankan.

Menurut Wirawan dalam bukunya evaluasi teori, model, metodologi, standar, aplikasi dan profesi menjelaskan beberapa tujuan evaluasi diantaranya adalah:

1. Mengukur pengaruh program terhadap masyarakat. Karena program sesungguhnya dilaksanakan untuk memberi dampak yang lebih baik bagi masyarakat yang menjadi sasaran program, maka evaluasi berperan untuk mengukur apakah program benar-benar berkontribusi positif terhadap masyarakat yang menjadi sasarannya. Dalam konteks social kita bisa melihat adanya program bantuan langsung tunai. Evaluasi akan melihat secara lebih dalam apakah program bantuan langsung tunai efektif dalam mengurangi beban ekonomi masyarakat. Jika iya seberapa tingkat efektivitasnya dan pada bagian mana dari program tersebut yang memiliki potensi penyelewengan dan lain-lain.
2. Menilai apakah program telah dilaksanakan sesuai dengan rencana. Sudah menjadi standar bahwa setiap program atau kegiatan formal tentu memiliki rencana pelaksanaan. Dalam kegiatan evaluasi salah satu tujuannya adalah memastikan bahwa program dijalankan sesuai dengan rencana yang telah disiapkan sebelumnya. Hal ini penting untuk menjaga seluruh tahapan dan kegiatan berjalan tanpa kendala, dan jika terjadi kendalapun akan dapat segera dikoreksi untuk diperbaiki sehingga tetap berada di jalan yang benar.
3. Mengukur apakah pelaksanaan program sesuai standar. Hal ini berkaitan dengan standar operasional prosedur (SOP). Biasanya suatu program atau kegiatan akan disiapkan dengan SOP yang menjadi pedoman bagi para eksekutor program dalam menjalankan kegiatan. Pada tahap ini seorang evaluator akan mengecek bahwa program telah dijalankan sesuai dengan standar yang telah ditentukan. Sebagai contoh program ujian nasional yang sekarang sudah ditiadakan. Dalam setiap tahap kegiatan Ujian Nasional ada standar yang harus dilalui dan dilaksanakan

oleh pelaksana kegiatan. Proses penyiapan soal ujian misalnya untuk sampai pada sekolah penyelenggara dari sekolah Rayon harus dikawal oleh polisi, kemudian saat ujian berlangsung harus dijaga oleh dua orang pengawas ujian dalam satu ruangan, satu ruangan hanya boleh diisi 20 peserta ujian dan standar-standar lain yang telah diatur.

4. Evaluasi program diharapkan dapat mengidentifikasi dan menemukan dimensi program yang berjalan dan tidak berjalan. Sebagai sebuah rencana tentu tidak semua hal yang disiapkan dapat berjalan sesuai harapan. Evaluasi dilakukan dengan tujuan untuk menemukan bagian-bagian dari sebuah program yang belum berjalan semestinya. Sebaliknya, dalam evaluasi juga akan dapat menemukan hal-hal yang sudah berjalan sesuai harapan dan perlu untuk dijaga agar jangan keluar dari perencanaan sebelumnya.
5. Pengembangan staf program. Sesuai dengan arahnya bahwa evaluasi di proyeksikan untuk mengembangkan staf program, bahwa dalam proses evaluasi akan memberi informasi bagi penanggungjawab kegiatan tentang kinerja staf yang terlibat dalam program tersebut. Dengan data dan informasi yang ada, pengelola program dapat segera menindaklanjuti langkah-langkah apa yang akan diambil. Jika terindikasi kinerja staf belum optimal maka pelatihan apa yang tepat diberikan kepada staf tersebut sehingga dapat mendukung kinerja program secara umum.
6. Memenuhi ketentuan undang-undang. Jika suatu program merupakan amanat undang-undang maka kegiatan evaluasi perlu dilakukan untuk memastikan bahwa program dijalankan sesuai dengan peraturan perundangan yang menjadi dasar pelaksanaan program tersebut.
7. Akreditasi program. Berkaitan dengan tujuan akreditasi, evaluasi dilakukan untuk mengukur tingkat layanan program kepada masyarakat. Apakah program dijalankan benar-benar sudah memberikan pelayanan yang memadai bagi masyarakat atau belum. Jika sudah memberikan pelayanan kepada masyarakat, seberapa derajat pelayanan yang diberikan. Disinilah fungsi

evaluasi akreditasi untuk mengukur masing-masing pelaksana program dalam menjalankan amanatnya. Seperti yang berlaku dalam dunia Pendidikan maka ada proses akreditasi bagi satuan Pendidikan yang telah memberikan pelayanan Pendidikan kepada masyarakat. Melalui akreditasi akan ditentukan peringkat akreditasi masing-masing satuan Pendidikan, apakah peringkat unggul (A), Baik (B), atau Cukup (C), atau malah tidak terakreditasi.

8. Mengukur *cost effectiveness* dan *cost-efficiency*. *Cost effectiveness* berkaitan dengan kesesuaian biaya yang dikeluarkan dengan program yang dilaksanakan. *Cost-efficiency* adalah perbandingan jumlah biaya yang dikeluarkan terhadap manfaat program yang dijalankan. Jika biaya yang dikeluarkan sepadan dengan manfaat yang diterima atas pelaksanaan program tersebut maka dapat dinyatakan memiliki *cost-efficiency* yang memadai. Semakin kecil biaya yang dikeluarkan tetapi capaian programnya memenuhi bahkan melampaui target, maka semakin efisien pembiayaan yang dikeluarkan.
9. Mengambil keputusan mengenai program. Evaluasi yang dilakukan pada akhirnya untuk mengambil keputusan apakah program akan dilanjutkan, distop, atau diperbaiki.
10. *Accountabilitas*. Suatu program dijalankan biasanya berpijak pada kebijakan, orang yang melaksanakan program memiliki tanggungjawab untuk menyampaikan tingkat keberhasilan atas program yang dijalankan. Dalam hal ini penanggungjawab program harus menyampaikan capaian-capaian program yang telah dan sedang dilaksanakan sebagai bentuk pertanggungjawaban. Tentu untuk mengetahui perbandingan antara sasaran/capaian program dengan standar atau rencana yang telah ditentukan.
11. Memberi balikan kepada pimpinan dan staf program. Dengan data dan informasi yang terkumpul melalui evaluasi baik pimpinan program maupun staf program dapat umpan balik sebagai bahan refleksi. Apa yang sudah dilaksanakan dan apa yang belum. Pada bagian mana yang sudah mencapai standar dan pada dimensi mana yang belum standar. Secara keseluruhan, informasi dan

data baik positif maupun negative menjadi sangat penting artinya bagi pimpinan program untuk menentukan langkah selanjutnya terhadap program yang telah dijalankan.

B. Arti Evaluasi

Stufflebeam mengatakan bahwa “*evaluation means a study designed and conducted to assist some audience to assess an object’s merit and worth*”.²⁶ Evaluasi diartikan sebagai penelitian yang didesain dan dilakukan untuk membantu audien dalam mengukur nilai dan manfaat objek yang dievaluasi. Definisi ini terlihat sederhana tetapi seharusnya dapat diterima secara luas karena logis dan mengarahkan kita untuk memahami pada dua hal yang sangat mendasar yaitu nilai suatu objek dan manfaatnya. Evaluasi dilakukan untuk mendapatkan informasi secara detail mengenai objek tertentu yang kemudian diberi makna berdasarkan data yang terkumpul dan akan digunakan sebagai bahan pertimbangan dalam membuat keputusan mengenai objek tersebut.

Scriven mendefinisikan evaluasi sebagai berikut “*What is evaluation? Synthesizing what the dictionaries and common usage tell us, it is the process of determining the merit, worth, or significance of things (near-synonyms are quality/value/importance). Reports on the results of this process are called evaluations if complex, evaluative claims if simple sentences, and we here use the term evaluand for whatever it is that is evaluated (optionally, we use evaluatee to indicate that an evaluand is a person)*”.²⁷

Apa yang dimaksud dengan evaluasi? Menyintesis apa yang disebutkan dalam kamus dan istilah yang digunakan secara umum yang sampai ke kita, evaluasi adalah proses memaknai suatu objek berdasarkan nilai dan manfaat, atau derajat signifikansi objek tersebut (sinonimnya adalah mutu/nilai/sesuatu yang penting). Laporan yang dibuat berdasarkan proses tersebut, dinamakan evaluasi, kami

²⁶DANIEL L. STUFFLEBEAM. *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, Edited by Daniel L. Stufflebeam, George F. Madaus, Thomas Kellaghan, (New York: Kluwer Academic Publishers, 2002), p. 35.

²⁷ MICHAEL SCRIVEN, *International Handbook of Educational Evaluation*, 15-30 T. Kellaghan, D.L. Stufflebeam (eds.) (Dordrecht: Kluwer Academic Publishers, 2003), p. 15.

menggunakan istilah “*evaluand*” untuk menyebut sesuatu yang dievaluasi (pilihan lain juga dapat digunakan istilah “*valuee*” untuk menyebut jika objek evaluasinya orang). Ada dua Kata kunci yang sama seperti definisi evaluasi menurut Stufflebeam yaitu “*merit*” dan “*worth*”.

Menurut Michael Quinn Patton evaluasi program adalah “*Program evaluation is the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness and/or inform decisions about future programming*”.²⁸ Evaluasi adalah proses pengumpulan informasi mengenai suatu aktifitas, karakteristik, dan hasil program yang dilakukan secara sistematis dan digunakan sebagai bahan pertimbangan menentukan program tersebut atau memperbaiki efektivitas program dan membuat keputusan mengenai kelanjutan program. Meskipun tidak spesifik menyebutkan nilai dan manfaat dari objek yang dievaluasi, definisi evaluasi menurut Patton ini tetap memiliki makna yang setara jika diperhatikan berdasarkan tujuan pengumpulan informasi adalah untuk membuat keputusan tentang program. Artinya, terdapat pertimbangan yang didasarkan pada nilai suatu program apakah berdampak positif terhadap penerima program, dan apakah ada manfaat yang diterima oleh sasaran program.

MARVIN C. ALKIN mengatakan bahwa “*Evaluation, after all, is concerned with providing information about identifiable programs - and, hopefully their use*”.²⁹ Evaluasi berkenaan dengan proses pengumpulan informasi tentang program yang dapat diidentifikasi dan berharap manfaat dari program tersebut. Upaya yang dilakukan dalam mengumpulkan informasi secara detail untuk mendapatkan data-data yang valid dan akurat sehingga akan diperoleh deskripsi program secara menyeluruh dan utuh. Dengan deskripsi yang komprehensif tersebut diharapkan ada nilai guna yang bisa disumbangkan oleh data tersebut dalam membuat kebijakan selanjutnya.

²⁸MICHAEL QUINN PATTON, *International Handbook of Educational Evaluation*, 15-30 T. Kellaghan, D.L. Stufflebeam (eds.) (Dordrecht: Kluwer Academic Publishers, 2003), P. 224

²⁹ MARVIN C. ALKIN, *International Handbook of Educational Evaluation*, 15-30 T. Kellaghan, D.L. Stufflebeam (eds.) (Dordrecht: Kluwer Academic Publishers, 2003), P. 189.

Berdasarkan beberapa penjelasan mengenai definisi evaluasi menurut para ahli di atas, dapat ditarik benang merah yang merepresentasikan makna evaluasi, diantaranya adalah nilai dan manfaat. Nilai merujuk pada kualitas program dan implementasinya. Sementara manfaat berkaitan dengan pengaruh positif yang diakibatkan oleh program dan implementasinya. Apakah masyarakat yang menjadi sasaran program menerima manfaatnya atau tidak. Program yang memiliki nilai baik belum tentu bermanfaat bagi sasaran program jika dijalankan dengan asal-asalan atau tidak mengikuti prosedur standar. Sebagai contoh, program bantuan langsung tunai bagi masyarakat kurang mampu, secara nilai baik tetapi jika dilaksanakan secara sembarangan, tanpa control ketat dan evaluasi berkelanjutan yang terjadi justru menumbuhkan potensi korupsi diberbagai lapisan.

C. Fungsi Evaluasi

Fungsi utama evaluasi adalah menelaah suatu objek atau keadaan untuk mendapatkan informasi yang tepat sebagai dasar untuk pengambilan keputusan. Fungsi evaluasi dapat dilihat berdasarkan jenis evaluasinya: (a) formatif, yaitu memberikan *feed back* bagi guru/instruktur sebagai dasar untuk memperbaiki proses pembelajaran dan mengadakan program remedial bagi peserta didik yang belum menguasai sepenuhnya materi yang dipelajari, (b) sumatif, yaitu mengetahui tingkat penguasaan peserta didik terhadap materi pelajaran, menentukan angka (nilai) sebagai bahan keputusan kenaikan kelas dan laporan perkembangan belajar, serta dapat meningkatkan motivasi belajar, (c) diagnostik, yaitu dapat mengetahui latar belakang peserta didik (psikologis, fisik, dan lingkungan) yang mengalami kesulitan belajar, (d) seleksi dan penempatan, yaitu hasil evaluasi dapat dijadikan dasar untuk menyeleksi dan menempatkan peserta didik sesuai dengan minat dan kemampuannya.

RINGKASAN

Evaluasi dapat dikatakan sebagai bentuk control terhadap pelaksanaan program atau kegiatan yang dilaksanakan oleh Lembaga ataupun oleh orang-perorangan. Dalam hal ini evaluasi memiliki peran strategis yang dapat memberi sumbangan terhadap keberhasilan suatu program.

Makna evaluasi dapat dilihat dari sisi nilai dan manfaat. Nilai merujuk pada kualitas program dan implementasinya. Sementara manfaat berkaitan dengan pengaruh positif yang diakibatkan oleh program dan implementasinya. Apakah masyarakat yang menjadi sasaran program menerima manfaatnya atau tidak. Program yang memiliki nilai baik belum tentu bermanfaat bagi sasaran program jika dijalankan dengan asal-asalan atau tidak mengikuti prosedur standar.

Fungsi evaluasi dapat dilihat berdasarkan jenis evaluasinya: (a) formatif, yaitu memberikan *feed back* bagi guru/instruktur sebagai dasar untuk memperbaiki proses pembelajaran dan mengadakan program remedial bagi peserta didik yang belum menguasai materi, (b) sumatif, yaitu mengetahui tingkat penguasaan peserta didik terhadap materi pelajaran, (c) diagnostik, yaitu mengetahui latar belakang peserta didik, (d) seleksi dan penempatan.

Latihan 3

1. Jelaskan tujuan-tujuan evaluasi!
2. Jelaskan evaluasi menurut fungsinya!
3. Jelaskan apa yang dimaksud *cost effectiveness* dan *cost-efficiency*!
4. Salah satu tujuan evaluasi adalah mengukur pengaruh program terhadap masyarakat. Jelaskan makna Pernyataan tersebut!
5. Jelaskan peran evaluasi!

Pembahasan Latihan Soal 3

1. Mengukur pengaruh program terhadap masyarakat, Menilai apakah program telah dilaksanakan sesuai dengan rencana, Mengukur apakah pelaksanaan program sesuai standar, Evaluasi program diharapkan dapat mengidentifikasi dan menemukan dimensi program yang berjalan dan tidak berjalan, Pengembangan staf program, Memenuhi ketentuan undang-undang, Akreditasi program, Mengukur *cost effectiveness* dan *cost-efficiency*, mengambil keputusan mengenai program, Accountabilitas, memberi balikan kepada pimpinan dan staf program.
2. Fungsi evaluasi dapat dilihat berdasarkan jenis evaluasinya: (a) formatif, yaitu memberikan feed back bagi guru/instruktur sebagai dasar untuk memperbaiki proses pembelajaran dan mengadakan program remedial bagi peserta didik yang belum menguasai materi, (b) sumatif, yaitu mengetahui tingkat penguasaan peserta didik terhadap materi pelajaran, (c) diagnostik, yaitu mengetahui latar belakang peserta didik, (d) seleksi dan penempatan.
3. Cost effectiveness berkaitan dengan kesesuaian biaya yang dikeluarkan dengan program yang dilaksanakan. Cost-efficiency adalah perbandingan jumlah biaya yang dikeluarkan terhadap manfaat program yang dijalankan. Jika biaya yang dikeluarkan sepadan dengan manfaat yang diterima atas pelaksanaan program tersebut maka dapat dinyatakan memiliki *cost-efficiency* yang memadai. Semakin kecil biaya yang dikeluarkan tetapi capaian programnya memenuhi bahkan melampaui target, maka semakin efisien pembiayaan yang dikeluarkan.
4. Mengukur pengaruh program terhadap masyarakat. Karena program sesungguhnya dilaksanakan untuk memberi dampak yang lebih baik bagi masyarakat yang menjadi sasaran program, maka evaluasi berperan untuk mengukur apakah program benar-benar berkontribusi positif terhadap masyarakat yang menjadi sasarannya. Dalam konteks social kita bisa melihat adanya program bantuan langsung tunai. Evaluasi akan melihat secara lebih dalam apakah program bantuan langsung tunai efektif dalam mengurangi beban ekonomi masyarakat.

Jika iya seberapa tingkat efektifitasnya dan pada bagian mana dari program tersebut yang memiliki potensi penyelewengan dan lain-lain.

5. Evaluasi dapat dikatakan sebagai bentuk kontrol terhadap pelaksanaan program atau kegiatan yang dilaksanakan oleh Lembaga ataupun oleh orang-perorangan. Dalam evaluasi akan diperoleh informasi dan data mengenai sisi-sisi program yang telah berjalan sesuai rencana dan standar. Serta memberi informasi tentang hal-hal yang melenceng dari rencana.

Test Formatif Pilihan Berganda

1. Sumber informasi yang berinteraksi dengan evaluator disebut...!
 - a. Audience
 - b. Sponsor
 - c. Participant
 - d. Client
 2. Objek evaluasi berikut disebut sebagai objek statis kecuali...!
 - a. Sarana dan prasarana
 - b. Instruktur
 - c. Kurikulum
 - d. Proyek
 3. Evaluasi Formatif adalah proses evaluasi yang dilakukan pada...!
 - a. Awal kegiatan
 - b. Akhir kegiatan
 - c. Saat kegiatan berlangsung
 - d. Saat kegiatan bermasalah
 4. Seseorang yang melakukan kegiatan evaluasi tetapi bukan pegawai dari Lembaga yang dievaluasinya disebut...!
 - a. Internal evaluator
 - b. Amatir evaluator
 - c. External evaluator
 - d. Visiting evaluator
 5. Salah satu fungsi evaluasi adalah diagnostik, yang berguna untuk...!
 - a. Mengetahui latar belakang peserta didik (psikologis, fisik, dan lingkungan) yang mengalami kesulitan belajar.
-

- b. Dijadikan dasar dalam menyeleksi dan menempatkan peserta didik sesuai dengan minat dan kemampuannya.
- c. Memberikan *feed back* bagi guru sebagai dasar untuk memperbaiki proses pembelajaran dan mengadakan program remedial bagi peserta didik yang belum menguasai sepenuhnya materi.
- d. Mengetahui tingkat penguasaan peserta didik terhadap materi pelajaran, menentukan angka (nilai) sebagai bahan keputusan kenaikan kelas dan laporan perkembangan belajar, serta dapat meningkatkan motivasi belajar.

Kunci jawaban dan pedoman penskoran

1. C
2. B
3. C
4. C
5. A

Bobot

Latihan soal essay 75 % dengan bobot per soal 15.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{50}{75} \times 100 = \mathbf{66,67}$$

Tes Formatif Pilihan Berganda 25% dengan bobot per soal 5.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{20}{25} \times 100 = \mathbf{80}$$

$$66,67 \times 75\% = \mathbf{50}$$

$$80 \times 25\% = \mathbf{20}$$

$$50 + 20 = \mathbf{70}$$

$$\text{Nilai akhir} = \mathbf{70}$$

MODUL 4

OBJEK EVALUASI, DIMENSI OBJEK EVALUASI, KRITERIA DALAM EVALUASI DAN BEBERAPA ISTILAH DALAM EVALUASI

Metode Pembelajaran	Estimasi waktu	Capaian Pembelajaran
-Presentasi virtual -Diskusi Virtual interaktif -Tanya jawab	120 menit X 1	Mahasiswa memahami: 1. Obyek evaluasi 2. Dimensi obyek evaluasi 3. Menentukan kriteri dalam evaluasi, 4. Istilah dalam evaluasi.

Materi 4. Objek evaluasi, Dimensi objek evaluasi, Kriteria dalam evaluasi dan istilah dalam evaluasi.

A. Obyek Evaluasi

Cakupan atau ruang lingkup evaluasi program pada umumnya lebih luas daripada evaluasi pembelajaran. Ruang lingkup pembahasan evaluasi program bisa bergerak dari kurikulum dan system instruksional yang digunakan untuk mendukung tercapainya visi, misi, dan strategi Lembaga diklat. Evaluasi program juga mencakup pembahasan sebagai bagian dari lima pilar manajemen, yaitu pilar pengawasan (*monitoring*), evaluasi (*evaluation*), dan pengendalian (*controlling*).³⁰ Sesungguhnya objek evaluasi bisa apa saja selama objek itu dapat diamati. Menurut Sukardi, objek evaluasi dapat dikelompokkan menjadi dua, yaitu objek yang berkarakteristik statis dan dinamis. Objek evaluasi statis adalah jika komponen yang dimaksud posisinya tetap tidak berubah saat dinilai, sedangkan objek dimanis adanya kemungkinan berubah terhadap objek evalauasi tersebut saat di nilai.

Objek evaluasi statis diantaranya, sarana dan prasarana, kurikulum atau paket program instruksional Lembaga diklat, proyek, dan materi

³⁰ Sukardi, evaluasi program Pendidikan dan pelatihan (Jakarta: PT Bumi Aksara, 2015), h. 10.

atau bahan ajar Lembaga diklat. Sementara objek evaluasi dinamis diantaranya adalah guru, instruktur, pengelola, staff administrasi, dan siswa.

Dalam Kawasan administrasi Pendidikan objek evaluasi dapat saja berupa;

- EVALUASI IMPLEMENTASI PROGRAM PENDIDIKAN INKLUSI
- EVALUASI IMPLEMENTASI PROGRAM *LEADERSHIP*
- EVALUASI IMPLEMENTASI PROGRAM NON-FORMAL BAGI REMAJA PUTUS SEKOLAH
- EVALUASI IMPLEMENTASI PROGRAM HAFALAN AL-QUR'AN
- EVALUASI IMPLEMENTASI PROGRAM DIKLAT CALON WIDYAISWARA
- EVALUASI IMPLEMENTASI KURIKULUM
- EVALUASI IMPLEMENTASI KURIKULUM JARINGAN SEKOLAH ISLAM TERPADU (JSIT)
- EVALUASI STANDAR PENGELOLAAN DI TINGKAT SATUAN PENDIDIKAN
- EVALUASI STANDAR SARANA DAN PRASARANA DI TINGKAT SATUAN PENDIDIKAN
- EVALUASI STANDAR PENDIDIK DAN TENAGA KEPENDIDIKAN DI SATUAN PENDIDIKAN
- EVALUASI PEMBELAJARAN AGAMA BERBASIS KARAKTER
- EVALUASI IMPLEMENTASI PROGRAM PENDIDIKAN DAN PELATIHAN
- EVALUASI IMPLEMENTASI PROGRAM INOBEL
- EVALUASI MANAJEMEN BERBASIS SEKOLAH (MBS)
- EVALUASI IMPLEMENTASI PROGRAM *LIFE SKILL*
- EVALUASI PERAN KOMITE SEKOLAH
- EVALUASI STANDAR PROSES PADA TINGKAT SATUAN PENDIDIKAN
- EVALUASI IMPLEMENTASI PENDIKAR DALA INTEGRASI MAPEL
- EVALUASI IMPLEMENTASI PENDIKAR DALAM BUDAYA SEKOLAH

- EVALUASI IMPLEMENTASI PENDIKAR DALAM KEGIATAN EKSTRA-KURIKULER
- EVALUASI IMPLEMENTASI PROGRAM PRAKTIK KERJA INDUSTRI
- EVALUASI PELAKSANAAN PROGRAM KANTIN KEJUJURAN DI SATUAN PENDIDIKAN
- EVALUASI PENYELENGARAAN OLIMPIADE SAINS TINGKAT...
- EVALUASI IMPLEMENTASI PROGRAM SEKOLAH BERBASIS LINGKUNGAN (SBL)/ADIWIYATA
- EVALUASI IMPLEMENTASI PROGRAM *BOARDNG SCHOOL*
- EVALUASI IMPLEMENTASI PROGRAM BOS /BOP
- EVALUASI IMPLEMENTASI PROGRAM LOMBA BUDAYA MUTU SEKOLAH
- EVALUASI IMPLEMENTASI PROGRAM LEMBAGA PENJAMINAN MUTU DI SEKLOAH ATAU DI PERGURUAN TINGGI
- Dst....Dsb...

B. Dimensi Obyek Evaluasi

Langkah pertama dalam proses evaluasi adalah mengidentifikasi objek evaluasi. Untuk tujuan ini, deskripsi singkat mengenai objek yang akan dievaluasi harus dituliskan. Deskripsi ini harus dirumuskan dengan cara yang menyeluruh sehingga dapat dipahami oleh pembaca luar. Item-item yang relevan juga perlu dimasukkan dalam deskripsi ini, sebagai contoh, nama objek evaluasi, manajer yang bertanggungjawab, pegawai yang terlibat dalam program tersebut, tujuan, referensi ilmiah, isi atau topiknya, elemen-elemen yang ada didalamnya, metode, bidang, sumber daya yang ada, figur kunci, perkembangan terdahulu, dll.

Pada saat mengidentifikasi objek evaluasi, klarifikasi mengenai tujuan sangat penting, karena pencapaian tujuan ini harus diuji dalam evaluasi. Sistem tiga dalam satu tujuan akan mempermudah tujuan-tujuan tersebut diklarifikasi dengan menempatkan mereka dalam kategori *macro*, *meso* dan kategori *micro*. Tujuan pada umumnya mencerminkan orientasi dasar dan fokus jangka panjang. Tujuan makro (*macro-goal*) ini kemudian di pecah-pecah ke dalam serangkaian

tujuan menengah (*meso-aims*) yang selanjutnya diturunkan lagi menjadi tujuan jangka pendek yang didefinisikan dengan baik (*micro-objectives*). Setiap tujuan *micro* harus konkret, dapat diverifikasi, waktunya tertentu dan realistis. Tujuan *micro* juga dimaksudkan untuk mengklarifikasi tujuan itu sendiri, target dan sasaran, serta untuk mengantisipasi potensi konsekuensi-konsekuensi yang tidak diinginkan atau tidak diharapkan.

Bagi objek evaluasi yang rumit, representasi visual-tektual dalam bentuk model yang logis bisa sangat membantu, menunjukkan elemen-elemen objek evaluasi dan bagaimana mereka menghubungkan antara satu dengan yang lainnya. Secara umum, penting juga untuk merekam situasi terkini, intervensi yang dilakukan dan tujuan atau hasil yang diinginkan. Pohon program menggunakan model yang jelas dan logis yang memudahkan dalam membedakan antara kondisi yang ada (*context, resources*), rencana (konsep, dan tujuan-tujuan yang sesuai), implementasi aktifitas tertentu dan hasil, baik yang diinginkan dan tidak diinginkan atas objek evaluasi yang diberikan. Objek evaluasi biasanya berbentuk proyek spesifik, tetapi bisa juga paket pengukuran, program dan seluruh kegiatan organisasi secara keseluruhan.

Pada umumnya evaluasi dilakukan terhadap dimensi hasil dari suatu program atau kegiatan. Model CIPP Stufflebeam mengemukakan 4 dimensi, antara lain: Konteks (*Context*), Input (*Input*), Proses implementasi (*Process*), Produk (*Product*). Maka, evaluasi pendidikan lengkap dilakukan mencakup dimensi: Manfaat tujuan, Mutu rencana, Sejauh mana tujuan dijalankan, dan Mutu hasil.

C. Menentukan Kriteria Dalam Evaluasi

Kriteria dalam evaluasi merujuk pada standar yang digunakan untuk menentukan ukuran atau derajat nilai suatu objek yang dievaluasi. Sebagai contoh, jika kita ingin mengetahui tingkat kecerdasan anak-anak maka kita akan buat rentang nilai antara 0 – 100 sebagai dasar untuk menentukan kategori-kategori tingkat kecerdasan. Siswa yang mendapatkan nilai 90 – 100 dianggap anak yang cerdas, nilai 80 – 89

kecerdasan sedang, nilai 60 -79 dengan kecerdasan cukup, dan nilai 0 -59 kecerdasan dibawah rata-rata. Dapat disimpulkan bahwa anak yang cerdas adalah jika yang bersangkutan mendapatkan nilai antara 90 – 100, tentu nilai yang diperoleh setelah melalui proses yang benar. Kriteria dibuat untuk menjadi standar minimal yang harus dicapai tetapi bukan berarti tidak boleh melebihi standar yang ditentukan. Dalam konteks evaluasi, kriteria digunakan sebagai pedoman dasar dalam melakukan evaluasi suatu program sehingga evaluator memiliki panduan yang jelas dalam menjalankan tugasnya. Dengan standar, evaluator memiliki pegangan dalam menentukan hasil evaluasi. Apakah program yang dievaluasi sesuai standar, melebihi standar, atau malah dibawah standar.

Menurut Suharsimi Arikunto dan Cepi Safruddin Abdul Jabar (2009) kriteria terdiri dari dua macam yaitu kuantitatif dan kualitatif. Kriteria kuantitatif itu sendiri terbagi menjadi (1) kriteria tanpa pertimbangan dan (2) kriteria dengan pertimbangan.

1. Kriteria Kuantitatif

a. Kriteria Kuantitatif Tanpa Pertimbangan

Kriteria yang disusun hanya dengan memperhatikan rentangan bilangan. Contoh: kondisi maksimal yang diharapkan untuk prestasi belajar diperhitungkan 100% jika penyusun menggunakan lima kategori nilai maka antara 1% dengan 100% dibagi rata sehingga menghasilkan kategori sebagai berikut.

- Nilai 5 (Baik Sekali), jika mencapai 81 – 100%
- Nilai 4 (Baik), jika mencapai 61 – 80%
- Nilai 3 (Cukup), jika mencapai 41 – 60%
- Nilai 2 (Kurang), jika mencapai 21 – 40%
- Nilai 1 (Kurang Sekali), jika mencapai <21%

b. Kriteria Kuantitatif dengan pertimbangan

Di beberapa perguruan tinggi misalnya, dalam menentukan nilai untuk masing-masing huruf mengacu pada peraturan akademik berdasarkan besarnya persentase pencapaian tujuan belajar. Misalnya, pengkategorian nilai tersebut memiliki rentangan yang tidak sama dalam setiap kategori, demikian juga jarak antara kategori yang satu dengan lainnya. Hal tersebut dibuat dengan adanya pertimbangan berdasarkan sudut pandang dan pertimbangan evaluator.

- Nilai A: rentangan nilai 80 – 100%
- Nilai B: rentangan nilai 66 – 79%
- Nilai C: rentangan nilai 56 – 65%
- Nilai D: rentangan nilai 40 – 55%
- Nilai E: kurang dari 40%

2. Kriteria kualitatif

Kriteria kualitatif adalah kriteria yang dikembangkan tidak menggunakan angka-angka, tetapi dengan mempertimbangkan indikator dan yang dikenai kriteria adalah komponen. Dalam kriteria kualitatif juga dibagi menjadi dua, yaitu (a) kriteria kualitatif tanpa pertimbangan dan (b) kriteria kualitatif dengan pertimbangan.

a. Kriteria Kualitatif Tanpa Pertimbangan

Evaluator yang akan mengembangkan kriteria kualitatif tanpa pertimbangan, melakukannya dengan menghitung banyaknya indikator dalam komponen yang dapat memenuhi persyaratan. Artinya, komponen adalah unsur pembentuk kriteria program, dan indikator adalah unsur pembentuk kriteria komponen.

b. Kriteria Kualitatif dengan Pertimbangan

Kriteria kualitatif dengan pertimbangan disusun melalui dua acara, yaitu (1) dengan mengurutkan indikator, dan (2) menggunakan pembobotan.

1) Kriteria Kualitatif dengan Pertimbangan Mengurutkan Indikator

Sebagai contoh;

- Nilai 5, jika memenuhi semua indikator.
- Nilai 4, jika memenuhi (a), (b), (c), atau (d).
- Nilai 3, jika memenuhi salah satu dari (a) atau (b) saja, dan salah satu dari (c) atau (d).
- Nilai 2, jika memenuhi salah satu dari empat indikator.
- Nilai 1, jika tidak satupun indikator yang dipenuhi.

2) Kriteria Kualitatif dengan Pertimbangan Pembobotan

Dalam pertimbangan pembobotan indikator-indikator yang akan diberi nilai dengan bobot yang berbeda. Penentuan peranan subindikator dalam mendukung nilai indikator harus disertai dengan alasan-alasan yang tepat. Jika sudah ditentukan pembobotannya, para penilai memilih skala yang akan digunakan dalam menilai objek. Bisa skala 1 -3, 1 - 4, atau 1 - 5, atau bahkan seperti yang biasa digunakan disekolah-sekolah, yaitu skala 1 - 10. Cara memperoleh nilai akhir indikator dapat dilakukan dengan;

- a) Mengalikan nilai masing-masing subindikator dengan bobotnya;
- b) Membagi jumlah nilai subindikator dengan jumlah bobot.

Rumus nilai akhir yang dapat digunakan adalah sebagai berikut.

$$\text{Nilai indikator} = \frac{\text{jumlah bobot subindikator} \times \text{nilai subindikator}}{\text{Jumlah bobot}}$$

Selanjutnya menentukan nilai komponen dengan dasar nilai indikator, dan nilai program dengan dasar nilai

komponen. Secara umum rumus yang digunakan sama dengan penghitungan nilai indikator di atas yang membedakan adalah aspek yang menjadi dasar penghitungan.

$$\text{Nilai komponen} = \frac{\text{jumlah bobot indikator} \times \text{nilai indikator}}{\text{Jumlah bobot}}$$

D. Beberapa Istilah Dalam Evaluasi

Kegiatan evaluasi seperti dijelaskan sebelumnya merupakan aktivitas yang tidak bebas nilai. Evaluasi sangat dipengaruhi oleh beberapa aspek eksternal yang tidak bisa dihindari. Salah satu alasannya adalah bahwa evaluasi pada akhirnya untuk membuat keputusan terhadap objek yang dievaluasi. Keputusan yang diambil berdasarkan hasil evaluasi tersebut akan memiliki efek bagi pihak-pihak yang berkepentingan diantaranya; masyarakat yang menjadi sasaran program, pihak sponsor, audien, klien, partisipan, dan para stakeholder. Berkaitan dengan para pihak yang memiliki kepentingan terhadap hasil evaluasi, terdapat beberapa istilah yang lazim digunakan dalam kegiatan evaluasi. Istilah-istilah yang digunakan merupakan ciri khas dalam penelitian evaluasi. Berikut ini beberapa istilah yang umum digunakan dalam kegiatan evaluasi.

Program: Ialah segala sesuatu yang dilakukan oleh seseorang, baik yang nyata (*tangible*) maupun yang abstrak (*intangibile*), dengan harapan akan mendatangkan hasil atau pengaruh.

Sponsor: Ialah orang atau lembaga yang meminta evaluasi dan membiayai proses evaluasi tersebut.

Audiens: Ialah pihak yang secara langsung atau tidak langsung berurusan dengan evaluasi. Mereka yang mempunyai perhatian dan minat terhadap evaluasi dan memperoleh hasilnya (sponsor, klien, partisipan, stakeholder).

Instrument: Ialah alat ukur yang digunakan dalam evaluasi, seperti: tes, kuesioner, observasi, wawancara, laporan, checklist, dan lain-lain.

Client: Ialah orang atau lembaga yang memesan atau meminta kepada evaluator untuk melakukan evaluasi.

Stakeholder: Pihak yang berkepentingan dan langsung dipengaruhi oleh hasil evaluasi.

Partisipan: Sumber informasi yang berinteraksi dengan evaluator.

Data kualitatif: Ialah data dalam keterangan tentang kejadian, transkrip wawancara, dan dokumen tertulis.

Data kuantitatif: Ialah data berupa angka atau berupa jumlah yang dapat diukur.

Selain istilah tersebut di atas, berikut ini adalah Istilah-istilah dalam evaluasi yang penulis ambil dari sumber lain, dapat dilihat pada tabel berikut. Baik istilah dan definisinya menggunakan Bahasa aslinya sebagai upaya menambah wawasan.

Tabel 4: Istilah-istilah dalam Evaluasi. Sumber: Donna M. Mertens & Amy T. Wilson

Term	Definition	Example
Internal Evaluator	Someone who conducts an evaluation and is an employee of the organization that houses what is to be evaluated.	Lace Up Shoes' Research and Development Department polled their employees on their satisfaction in their work environment.
External Evaluator	Someone who conducts an evaluation but is not an employee of the organization that houses what is to be evaluated.	Susan Nguyen, of Urban Planners East Coast, was hired to evaluate the effectiveness of the city's use of social media to attract new residents to its downtown.
Formative Evaluation	An evaluation done during the planning or implementation of a program or product that is designed to provide feedback to improve the evaluand.	Teachers of English as a second language planned and participated in a project to connect adult mentors with their students. The formative evaluation showed schools, community centers, and nonprofits how best to reach out, connect with, and nurture mentoring relationships.
Summative Evaluation	An evaluation done at the end of or on completion of a program.	The creators of the videogame "Full Life" announced that their new version of this game was more believable and

Term	Definition	Example
		engaging, based on the feedback from 200 graduate students who participated in focus groups after playing the game over a period of 1 month.
Gender Analysis	A variety of methods used to recognize the difference in participation, access, and rights of women compared to men in social, economic, and legal systems.	In conducting a gender analysis of the data concerning education advocacy led by administrators of higher education, it was learned that although women of color were twice as likely to engage in advocacy than any other group, they were the least likely to be accepted to represent higher education at legislative meetings

RINGKASAN

Objek evaluasi dapat dikelompokkan menjadi dua, yaitu objek yang berkarakteristik statis dan dinamis. Objek statis adalah jika komponen yang dimaksud posisinya tetap tidak berubah saat dinilai, sedangkan objek dinamis adanya kemungkinan berubah terhadap objek evaluasi tersebut saat di nilai.

Fungsi evaluasi dapat dilihat berdasarkan jenis evaluasinya: (a) formatif, yaitu memberikan *feed back* bagi guru/instruktur sebagai dasar untuk memperbaiki proses pembelajaran dan mengadakan program remedial bagi peserta didik yang belum menguasai materi, (b) sumatif, yaitu mengetahui tingkat penguasaan peserta didik terhadap materi pelajaran, (c) diagnostik, yaitu mengetahui latar belakang peserta didik, (d) seleksi dan penempatan.

Kriteria dalam evaluasi merujuk pada standar yang digunakan untuk menentukan ukuran atau derajat nilai suatu objek yang dievaluasi. kriteria terdiri dari dua macam yaitu kuantitatif dan kualitatif. Baik Kriteria kuantitatif maupun Kriteria kualitatif, masing-masing terbagi menjadi 2 bentuk, yaitu; (1) kriteria tanpa pertimbangan dan (2) kriteria dengan pertimbangan.

Latihan 4

1. Jelaskan tujuan-tujuan evaluasi!
2. Jelaskan evaluasi menurut fungsinya!

3. Jelaskan apa yang dimaksud dengan kriteria dalam evaluasi dan sebutkan jenisnya!
4. Jelaskan perbedaan antara objek evaluasi dengan dimensi evaluasi!
5. Jelaskan peran evaluasi!

Pembahasan Latihan Soal 4

1. Mengukur pengaruh program terhadap masyarakat, Menilai apakah program telah dilaksanakan sesuai dengan rencana, Mengukur apakah pelaksanaan program sesuai standar, Evaluasi program diharapkan dapat mengidentifikasi dan menemukan dimensi program yang berjalan dan tidak berjalan, Pengembangan staf program, Memenuhi ketentuan undang-undang, Akreditasi program, Mengukur *cost effectiveness* dan *cost-efficiency*, mengambil keputusan mengenai program, Accountabilitas, memberi balikan kepada pimpinan dan staf program.
2. Fungsi evaluasi dapat dilihat berdasarkan jenis evaluasinya: (a) formatif, yaitu memberikan feed back bagi guru/instruktur sebagai dasar untuk memperbaiki proses pembelajaran dan mengadakan program remedial bagi peserta didik yang belum menguasai materi, (b) sumatif, yaitu mengetahui tingkat penguasaan peserta didik terhadap materi pelajaran, (c) diagnostik, yaitu mengetahui latar belakang peserta didik, (d) seleksi dan penempatan.
3. Kriteria dalam evaluasi merujuk pada standar yang digunakan untuk menentukan ukuran atau derajat nilai suatu objek yang dievaluasi. kriteria terdiri dari dua macam yaitu kuantitatif dan kualitatif. Baik Kriteria kuantitatif maupun Kriteria kualitatif, masing-masing terbagi menjadi 2 bentuk, yaitu; (1) kriteria tanpa pertimbangan dan (2) kriteria dengan pertimbangan.
4. Objek evaluasi bisa apa saja selama objek itu dapat diamati. Objek evaluasi dapat dikelompokkan menjadi dua, yaitu objek yang berkarakteristik statis dan dinamis. Objek evaluasi statis adalah jika komponen yang dimaksud posisinya tetap tidak berubah saat dinilai, sedangkan objek dinamis adanya kemungkinan berubah terhadap objek evaluasi tersebut saat di nilai. Pada umumnya evaluasi dilakukan terhadap dimensi hasil dari suatu program atau kegiatan. Model CIPP

Stufflebeam mengemukakan 4 dimensi, antara lain: Konteks (*Context*), Input (*Input*), Proses implementasi (*Process*), Produk (*Product*). Maka, evaluasi pendidikan lengkap dilakukan mencakup dimensi: Manfaat tujuan, Mutu rencana, Sejauhmana tujuan dijalankan, dan Mutu hasil.

5. Evaluasi dapat dikatakan sebagai bentuk kontrol terhadap pelaksanaan program atau kegiatan yang dilaksanakan oleh Lembaga ataupun oleh orang-perorangan. Dalam evaluasi akan diperoleh informasi dan data mengenai sisi-sisi program yang telah berjalan sesuai rencana dan standar. Serta memberi informasi tentang hal-hal yang melenceng dari rencana.

Test Formatif Pilihan Berganda

1. Sumber informasi yang berinteraksi dengan evaluator disebut...!
 - a. Audience
 - b. Sponsor
 - c. Participant
 - d. Client
2. Objek evaluasi berikut disebut sebagai objek statis kecuali...!
 - a. Sarana dan prasarana
 - b. Instruktur
 - c. Kurikulum
 - d. Proyek
3. Evaluasi Formatif adalah proses evaluasi yang dilakukan pada...!
 - a. Awal kegiatan
 - b. Akhir kegiatan
 - c. Saat kegiatan berlangsung
 - d. Saat kegiatan bermasalah
4. Seseorang yang melakukan kegiatan evaluasi tetapi bukan pegawai dari Lembaga yang dievaluasinya disebut...!
 - a. Internal evaluator
 - b. Amatir evaluator
 - c. External evaluator
 - d. Visiting evaluator

5. Salah satu fungsi evaluasi adalah diagnostik, yang berguna untuk...!
 - a. Mengetahui latar belakang peserta didik (psikologis, fisik, dan lingkungan) yang mengalami kesulitan belajar.
 - b. Dijadikan dasar dalam menyeleksi dan menempatkan peserta didik sesuai dengan minat dan kemampuannya.
 - c. Memberikan *feed back* bagi guru sebagai dasar untuk memperbaiki proses pembelajaran dan mengadakan program remedial bagi peserta didik yang belum menguasai sepenuhnya materi.
 - d. Mengetahui tingkat penguasaan peserta didik terhadap materi pelajaran, menentukan angka (nilai) sebagai bahan keputusan kenaikan kelas dan laporan perkembangan belajar, serta dapat meningkatkan motivasi belajar.

Kunci Jawaban dan Pedoman Penskoran

1. C
2. B
3. C
4. C
5. A

Bobot

Latihan soal essay 75 % dengan bobot per soal 15.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{50}{75} \times 100 = \mathbf{66,67}$$

Tes Formatif Pilihan Berganda 25% dengan bobot per soal 5.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{20}{25} \times 100 = \mathbf{80}$$

$$66,67 \times 75\% = \mathbf{50}$$

$$80 \times 25\% = \mathbf{20}$$

$$50 + 20 = \mathbf{70}$$

$$\text{Nilai akhir} = \mathbf{70}$$

MODUL 5

PENERIMA MANFAAT EVALUASI, PROSEDUR EVALUASI, METODE DAN PENDEKATAN EVALUASI, PIHAK YANG MELAKUKAN EVALUASI

Metode Pembelajaran	Estimasi waktu	Capaian Pembelajaran
-Presentasi virtual -Diskusi Virtual interaktif -Tanya jawab	120 menit X 1	- Mahasiswa mampu memahami : 1. Penerima manfaat evaluasi 2. Prosedur evaluasi 3. Metode dan pendekatan evaluasi 4. Pihak yang melakukan evaluasi

Materi 5. Penerima manfaat evaluasi, prosedur evaluasi, Metode dan pendekatan evaluasi, Pihak yang melakukan evaluasi.

A. Penerima Manfaat Evaluasi

Evaluasi dilakukan untuk tujuan yang spesifik sesuai dengan lingkup objek yang akan dievaluasi. Alasannya adalah bahwa evaluasi dilakukan terhadap objek yang memiliki sifat unik dan khusus sehingga hasil evaluasi tidak bisa digeneralisir seperti penelitian murni. Selain itu, evaluasi juga memiliki tujuan untuk membuat keputusan terhadap program atau objek evaluasi. Dengan spesifikasi khusus itu, maka pihak-pihak yang berkepentingan terkait hasil evaluasi tentu akan berbeda-beda. Dalam hal ini, menjadi penting untuk mengidentifikasi para pemangku kepentingan dalam kegiatan evaluasi. Menurut Wirawan (2016) para pemangku kepentingan yaitu individu atau kelompok masyarakat yang memiliki kepentingan untuk;

1. Mendapatkan layanan atau perlakuan/klien program.
2. Memperoleh penghasilan dari program.
3. Mensponsori atau penyandang dana program.
4. Mensponsori evaluasi program.
5. Masyarakat yang terpengaruh oleh pelaksanaan program.

6. Masyarakat yang dirugikan oleh pelaksanaan program.
7. Masyarakat yang memengaruhi pelaksanaan program.
8. Mereka yang akan mempergunakan hasil evaluasi.

B. Prosedur Evaluasi

Tyler menganggap bahwa evaluasi harus menjadi tahap yang esensial dalam pengembangan program. Kegiatan evaluasi harus dilakukan berdasarkan langkah-langkah yang terstruktur. Menurut Tyler, prosedur dalam evaluasi program terdiri atas delapan langkah sebagai berikut:

1. Menetapkan sasaran atau tujuan.
2. Menempatkan tujuan dalam klasifikasi-klasifikasi.
3. Mendefinisikan tujuan dalam tindakan.
4. Menciptakan situasi dan kondisi yang konjussif sehingga memungkinkan untuk menunjukkan pencapaian tujuan.
5. Berusaha menjelaskan tujuan strategi bagi personil terkait dalam memilih situasi.
6. Untuk memilih dan mengembangkan teknik-teknik pengukuran yang tepat.
7. Sebagai cara untuk mengumpulkan data kinerja (dalam hal program pendidikan dapat berupa kinerja siswa).
8. Untuk membandingkan data dengan tujuan perilaku.³¹

Hal yang perlu dicatat adalah bahwa definisi evaluasi menurut Tyler sangat sesuai dengan langkah-langkah prosedur ini. Jika tujuan program kurikulum dan instruksi adalah untuk membuat perubahan pola perilaku bagi yang diuntungkan oleh program tersebut, maka metode pengukuran tentang perubahan tersebut menjadi penting untuk disiapkan. Tujuan dan sasaran (*Goal dan Objectives*) harus ditentukan dan didefinisikan dengan jelas. Tujuan program (*Goal*) merupakan cita-cita yang dapat diusahakan untuk dicapai, sementara

³¹ Daniel L. Stufflebeam Anthony J. Shinkfield, *Systematic Evaluation A Self-Instructional Guide to Theory and Practice*, (Boston: Kluwer-NIJhoff Publishing a member of the Kluwer Academic Publishers Group, 1988), p.71

Objective adalah pernyataan *subgoal* yang dapat diekspresikan dalam bentuk entitas yang terukur.

Metfessel dan Michael mengembangkan delapan langkah evaluasi untuk membantu sekolah mengevaluasi pencapaian tujuan program sekolah. Delapan langkah proses evaluasi yang dimaksud akan penulis coba terjemahkan dari bahasa aslinya secara bebas.

1. Melibatkan seluruh anggota komunitas sekolah sebagai peserta, atau fasilitator evaluasi.
2. Mengonstruksi kepaduan antara tujuan besar dengan tujuan spesifik yang diatur secara hirarki dari yang bersifat umum ke yang bersifat khusus sesuai yang diinginkan melalui (a) menetapkan tujuan yang luas meliputi tujuan teoritis program; (b) menyatakan tujuan spesifik kedalam istilah yang operasional sehingga memungkinkan bagi pengukuran tujuan kapanpun dibutuhkan; dan (c) mengembangkan kriteria penilaian yang memungkinkan pendefinisian signifikansi dan hasil-hasil yang relevan, menentukan prioritas yang realistis terkait kebutuhan sosial, kesiapan siswa untuk belajar, balikan siswa-guru, dan ketersediaan staf serta sumber daya material.
3. Menerjemahkan tujuan perilaku spesifik kedalam bentuk yang dapat di komunikasikan, dapat diterapkan untuk memfasilitasi pembelajaran di sekolah.
4. Menyeleksi atau mengonstruksi berbagai instrumen yang akan melengkapi pengukuran dimana kesimpulan dapat digambarkan melalui efektifitas program untuk memenuhi tujuan yang telah direncanakan.
5. Melakukan observasi secara periodik menggunakan instrumen yang variatif untuk mengukur perubahan perilaku yang sesuai dengan tujuan-tujuan tertentu yang diinginkan.
6. Menganalisa data yang tersedia dengan mengukur perubahan menggunakan metode statistik yang tepat.
7. Menerjemahkan data relatif dengan tujuan spesifik dalam bentuk standar penilaian tertentu dan nilai-nilai yang sesuai dengan level kinerja yang diinginkan; cara ini digunakan untuk

menggambarkan kesimpulan yang menyediakan informasi mengenai arah pertumbuhan, perkembangan siswa, dan efektifitas program secara keseluruhan.

8. Membuat rekomendasi yang dapat digunakan sebagai dasar bagi implementasi selanjutnya, modifikasi, dan revisi tujuan umum dan tujuan khusus dengan tujuan perkembangan program; menyediakan umpan balik, berdasarkan rekomendasi, bagi seluruh individu yang terlibat dalam program; dan untuk membuat keputusan bagi putaran proses evaluasi sampai rekomendasi ketika hasilnya telah disampaikan ke pihak yang berhak untuk diketahui oleh para audiens.³²

C. Metode dan Pendekatan Evaluasi

Wirawan menyebutkan bahwa metode penelitian yang dapat digunakan dalam evaluasi adalah Metode Kuantitatif (*quantitative methods*), metode kualitatif (*qualitative methods*), dan metode campuran (*mixed-methods*). Artinya, semua metode yang biasa digunakan dalam penelitian murni sesungguhnya dapat diadopsi untuk digunakan dalam penelitian evaluasi. Penggunaan metode ditentukan oleh jenis data yang perlu dijaring; sumber informasi; waktu yang diperlukan untuk melakukan evaluasi; sumber daya yang diperlukan (tenaga, biaya, dan alat), tersebarnya wilayah layanan program dan pihak yang mendapat layanan program.

³² Daniel L. Stufflebeam Anthony J. Shinkfield, *Systematic Evaluation A Self-Instructional Guide to Theory and Practice*, (Boston: Kluwer-Nijhoff Publishing a member of the Kluwer Academic Publishers Group, 1988), p.77.

Tabel 5: Metode Kuantitatif, Metode Kualitatif, Metode Campuran. Sumber: John W. Creswell

Metode Kuantitatif	Metode Campuran	Metode Kualitatif
<ul style="list-style-type: none"> • Bersifat <i>pre-determined</i> (sudah ditentukan sebelumnya). • Pertanyaan-pertanyaan yang didasarkan pada instrumen penelitian. • Data performa, data sikap, data observasi, dan data sensus. • Analisis statistik. • Interpretasi statistik. 	<ul style="list-style-type: none"> • Bersifat <i>pre-determined</i> dan berkembang dinamis. • Pertanyaan-pertanyaan terbuka dan pertanyaan-pertanyaan tertutup. • Bentuk-bentuk data berganda yang terbuka pada kemungkinan-kemungkinan lain. • Analisis statistik dan analisis tekstual. • Lintas-interpretasi data base. 	<ul style="list-style-type: none"> • Berkembang dinamis. • Pertanyaan-pertanyaan terbuka • Data wawancara, data observasi, data dokumentasi, dan data audio-visual. • Analisis tekstual dan gambar. • Interpretasi tema-tema, pola-pola.

D. Pihak yang Melakukan Evaluasi

Sesungguhnya siapapun dapat menjadi evaluator selama yang bersangkutan memiliki pengetahuan dan kemampuan yang cukup memadai dalam melakukan evaluasi. Para pihak yang melakukan evaluasi dapat dibedakan menjadi dua yaitu; evaluator internal dan evaluator eksternal. Evaluator internal merupakan seseorang atau tim yang berasal dari dalam lembaga yang akan dievaluasi. Evaluator eksternal adalah seseorang atau tim yang berasal dari luar lembaga atau institusi yang akan di evaluasi. Jika evaluator berasal dari luar institusi atau lembaga biasanya profesional yang memang memiliki keahlian tertentu yang dibutuhkan untuk melakukan evaluasi. Sehingga secara kompetensi lebih bisa dipercaya karena pengetahuan, kompetensi dan juga pengalaman dalam melaksanakan evaluasi. Kelebihan lain dari evaluator eksternal akan lebih fokus dalam melakukan kegiatan evaluasi tanpa dibatasi perasaan tendensius terhadap lembaga yang dievaluasi yang pada akhirnya berpotensi atau dapat menyebabkan bias dalam pengumpulan dan pengolahan data serta informasi yang diperoleh.

Seorang evaluator sebaiknya memiliki kompetensi dalam hal penelitian yang relevan dengan bidang keahlian. Evaluator juga dituntut

menguasai bidang pengukuran dan metodologi penelitian serta memahami konteks sosial yang berhubungan dengan sasaran evaluasi dan juga pelaku yang akan dievaluasi.

Menurut Suharsimi Arikunto dan Cipi Safruddin Abdul Jabar syarat seseorang yang dapat menjadi evaluator adalah sebagai berikut.

1. Mampu melaksanakan, artinya seorang evaluator harus memiliki kemampuan untuk melaksanakan evaluasi yang didukung oleh teori dan keterampilan praktik.
2. Cermat, dapat melihat celah-celah dan detail dari program serta bagian program yang akan dievaluasi.
3. Objektif, tidak mudah dipengaruhi oleh kepentingan pribadi, supaya proses pengumpulan data dan informasi sesuai apa adanya, dan proses analisis serta pengambilan kesimpulannyapun dilakukan sesuai prosedur dan standar yang harus dilakukan.
4. Sabar dan tekun, karena kegiatan evaluasi membutuhkan ketelitian dan waktu yang cukup, terutama saat perencanaan, proses pengumpulan data dan analisis data sangat dibutuhkan konsentrasi yang maksimal sehingga hasil evaluasi menjadi tepat.
5. Hati-hati dan bertanggungjawab, yaitu melakukan pekerjaan evaluasi dengan penuh pertimbangan, namun apabila masih ada kekeliruan yang diperbuat, berani menanggung resiko atas segala kesalahannya.³³

Dalam konteks satuan pendidikan dapat dilakukan evaluasi yang evaluatornya berasal dari internal dan juga dari eksternal. Evaluator internal bisa seorang guru yang dengan kewenangannya kemudian mengembangkan instrumen atau soal-soal ujian untuk mengukur kinerja siswa. Pengukuran ini dapat dilakukan sebelum ataupun sesudah proses pembelajaran. Jika evaluasi dilakukan sebelum siswa diberikan treatment atau perlakuan dalam bentuk pembelajaran maka

³³ Suharsimi Arikunto dan cipi Safruddin Abdul Jabar. *Evaluasi Program Pendidikan: Pedoman Teoritis Praktis bagi Mahasiswa dan Praktisi Pendidikan*. (Jakarta: PT Bumi Aksara, 2009), hh. 22-23.

disebut *pre-test*, dan jika soal-soal ujian diberikan setelah program kegiatan pembelajaran selesai disebut *post-test*.

Menurut Suharsimi dan Cepi safruddin Abdul Jabar evaluator Internal maupun eksternal memiliki kelebihan dan kekurangannya masing-masing diantaranya adalah.

Tabel 6: kelebihan dan kekurangan Evaluator Internal dan Eksternal. Sumber: Suharsimi dan Cepi Safruddin Abdul Jabar, 2009, hh. 23-24.

Evaluator Internal	Evaluator Eksternal
<p>Kelebihan</p> <ol style="list-style-type: none"> 1. Memahami betul program yang akan dievaluasi sehingga kekhawatiran untuk tidak atau kurang tepat sasaran tidak perlu ada. 2. Karena evaluastor dari dalam, pengambil keputusan tidak perlu banyak mengeluarkan biaya untuk membayar evaluator profesional 	<p>Kelebihan</p> <ol style="list-style-type: none"> 1. Tidak memiliki kepentingan atas keberhasilan program, maka evaluator luar 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 sesuai dengan keadaan dan kenyataan. 2. Seorang ahli yang dibayar, biasanya akan mempertahankan kredibilitas kemampuannya. Dengan begitu, evaluator akan bekerja secara serius dan hati-hati.
<p>Kekurangan</p> <ol style="list-style-type: none"> 1. Adanya unsur subjektivitas dari evaluator, sehingga berusaha menyampaikan aspek positif dari program yang dievaluasi dan menginginkan agar kebijakan tersebut dapat diimplementasikan dengan baik pula. 2. Karena evaluator sudah memahami seluk-beluk program, jika evaluator kurang sabar, kegiatan evaluasi akan dilaksanakan dengan tergesa-gesa sehingga kurang cermat. 	<p>Kekurangan</p> <ol style="list-style-type: none"> 1. Evaluator luar adalah orang baru, yang sebelumnya tidak mengenal kebijakan program yang akan dievaluasi. Mereka berusaha mempelajari seluk-beluk program tersebut setelah mendapat permintaan untuk mengevaluasi. Hal ini berpotensi adanya ketidakjelasan informasi meskipun sudah dijelaskan dan mempelajari isi kebijakannya. Dan itu wajar karena evaluator dari luar tidak terlibat dalam pembuatan kebijakan. Dampak ketidakjelasan pemahaman tersebut memungkinkan kesimpulan yang diambil kurang tepat. 2. Pemborosan, pengambil keputusan harus mengeluarkan

Evaluator Internal	Evaluator Eksternal
	biaya yang cukup banyak untuk membayar evaluator luar.

Evaluator eksternal dapat diambil contoh ketika sekolah akan melaksanakan akreditasi. Biasanya evaluator berasal dari luar satuan pendidikan yang akan dievaluasi yang disebut asesor. Asesor adalah seseorang yang ditugaskan oleh lembaga resmi pemerintah Badan Akreditasi Nasional Sekolah/Madrasah untuk melakukan evaluasi dengan cara mengumpulkan data dan informasi di sekolah sasaran. Asesor akreditasi biasanya berasal dari berbagai institusi diantaranya dosen, praktisi pendidikan, widyaiswara, dan profesi lain yang relevan. Karena dalam kegiatan evaluasi pendidikan yang disebut akreditasi dibutuhkan kompetensi dan pengetahuan khusus yang berkaitan dengan pendidikan. Untuk memenuhi kebutuhan akan kompetensi yang dimaksud, seorang calon asesor biasanya akan diberikan pendidikan dan pelatihan sehingga mereka memiliki kecakapan standar untuk melakukan tugas.

Masing-masing evaluator yang berasal dari internal maupun dari eksternal institusi memiliki kelebihan dan kelemahan. Jika ingin mendapatkan hasil evaluasi yang optimal pilihan alternatif dapat diambil dengan model penggabungan evaluator internal dan eksternal. Lembaga pelaksana program atau staf program dapat mengambil beberapa personil internal untuk dilibatkan dalam kegiatan evaluasi yang dilakukan oleh pihak profesional. Dengan keterlibatan personil internal diharapkan dapat membantu menggali informasi dan data yang lebih detail dan akurat, yang pada akhirnya ketika keputusan harus diambil benar-benar berdasarkan pertimbangan dan proses yang tepat.

RANGKUMAN

Prosedur dalam evaluasi program terdiri atas delapan langkah sebagai berikut;

- a. Menetapkan sasaran atau tujuan.
- b. Menempatkan tujuan dalam klasifikasi-klasifikasi.
- c. Mendefinisikan tujuan dalam tindakan.
- d. Menciptakan situasi dan kondisi yang konjussif sehingga memungkinkan untuk menunjukkan pencapaian tujuan.
- e. Berusaha menjelaskan tujuan strategi bagi personil terkait dalam memilih situasi.
- f. Untuk memilih dan mengembangkan teknik-teknik pengukuran yang tepat.
- g. Sebagai cara untuk mengumpulkan data kinerja (dalam hal program pendidikan dapat berupa kinerja siswa).
- h. Untuk membandingkan data dengan tujuan perilaku.

Penggunaan metode evaluasi ditentukan oleh jenis data yang perlu dijaring, sumber informasi, waktu yang diperlukan untuk melakukan evaluasi, sumber daya yang diperlukan (tenaga, biaya, dan alat), tersebarnya wilayah layanan program dan pihak yang mendapat layanan program. Metode yang dapat digunakan dalam evaluasi adalah Metode Kuantitatif (*quantitative methods*), metode kualitatif (*qualitative methods*), dan metode campuran (*mixed-methods*).

Syarat menjadi evaluator; Mampu melaksanakan, Cermat, Objektif, Sabar dan tekun, hati-hati dan bertanggungjawab. Evaluator dapat berasal dari Internal maupun eksternal. Evaluator internal adalah evaluator yang berasal dari/sebagai bagian/pegawai dari organisasi atau lembaga yang dievaluasi. Evaluator eksternal merupakan evaluator profesional yang berasal dari luar organisasi atau lembaga yang dievaluasi.

Latihan 5

1. Sebutkan pihak-pihak yang berkepentingan terhadap evaluasi!
2. Jelaskan prosedur evaluasi!
3. Jelaskan metode apa saja yang dapat digunakan dalam evaluasi!
4. Sebutkan syarat-syarat menjadi evaluator dan sebutkan jenis evaluator!
5. Jelaskan kelebihan evaluator eksternal!

Pembahasan Latihan Soal 5

1. Klien program, Memperoleh penghasilan dari program, penyandang dana program, sponsor evaluasi program, Masyarakat yang terpengaruh oleh pelaksanaan program, Masyarakat yang dirugikan oleh pelaksanaan program, Masyarakat yang memengaruhi pelaksanaan program, pihak-pihak yang akan mempergunakan hasil evaluasi.
2. Prosedur dalam evaluasi program terdiri atas delapan langkah sebagai berikut;
 - a. Menetapkan sasaran atau tujuan.
 - b. Menempatkan tujuan dalam klasifikasi-klasifikasi.
 - c. Mendefinisikan tujuan dalam tindakan.
 - d. Menciptakan situasi dan kondisi yang konjussif sehingga memungkinkan untuk menunjukkan pencapaian tujuan.
 - e. Berusaha menjelaskan tujuan strategi bagi personil terkait dalam memilih situasi.
 - f. Untuk memilih dan mengembangkan teknik-teknik pengukuran yang tepat.
 - g. Sebagai cara untuk mengumpulkan data kinerja (dalam hal program pendidikan dapat berupa kinerja siswa).
 - h. Untuk membandingkan data dengan tujuan perilaku.
3. Metode yang dapat digunakan dalam evaluasi adalah Metode Kuantitatif (*quantitative methods*), metode kualitatif (*qualitative methods*), dan metode campuran (*mixed-methods*).
4. Syarat menjadi evaluator; Mampu melaksanakan, Cermat, Objektif, Sabar dan tekun, hati-hati dan bertanggungjawab. Evaluator dapat berasal dari Internal maupun eksternal. Evaluator internal adalah

evaluator yang berasal dari/sebagai bagian/pegawai dari organisasi atau lembaga yang dievaluasi. Evaluator eksternal merupakan evaluator profesional yang berasal dari luar organisasi atau lembaga yang dievaluasi.

5. Kelebihan evaluator eksternal adalah 1) tidak memiliki kepentingan atas keberhasilan program, maka evaluator luar 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 sesuai dengan keadaan dan kenyataan. 2) Seorang ahli yang dibayar, biasanya akan mempertahankan kredibilitas kemampuannya. Dengan begitu, evaluator akan bekerja secara serius dan hati-hati.

Test Formatif Pilihan Berganda

1. Salah satu syarat untuk menjadi evaluator adalah objektif yang memiliki makna...!
 - a. Dapat melihat celah-celah dan detail dari program serta bagian program yang akan dievaluasi.
 - b. Memiliki kemampuan untuk melaksanakan evaluasi yang didukung oleh teori dan keterampilan praktik.
 - c. Melakukan pekerjaan evaluasi dengan penuh pertimbangan, namun apabila masih ada kekeliruan yang diperbuat, berani menanggung resiko atas segala kesalahannya.
 - d. Tidak mudah dipengaruhi oleh kepentingan pribadi, supaya proses pengumpulan data dan informasi sesuai apa adanya, dan proses analisis serta pengambilan kesimpulannyapun dilakukan sesuai prosedur dan standar yang harus dilakukan.
2. Adanya unsur subjektivitas dari evaluator, sehingga berusaha menyampaikan aspek positif dari program yang dievaluasi dan menginginkan agar kebijakan tersebut dapat diimplementasikan dengan baik, merupakan karakteristik dari...!
 - a. Kelebihan evaluator internal.
 - b. Kelebihan evaluator eksternal.
 - c. Kekurangan/kelemahan evaluator internal.

- d. Kekurangan/kelemahan evaluator eksternal.
3. Salah satu contoh evaluator eksternal dalam satuan Pendidikan adalah...!
 - a. Guru
 - b. Kepala sekolah
 - c. Asesor akreditasi
 - d. Tenaga kependidikan
4. Ciri khas yang ada dalam metode penelitian kualitatif adalah...!
 - a. Data wawancara, data observasi, data dokumentasi, dan data audio-visual.
 - b. Pertanyaan-pertanyaan yang didasarkan pada instrumen penelitian.
 - c. Data performa, data sikap, data observasi, dan data sensus.
 - d. Bersifat pre-determined (sudah ditentukan sebelumnya).
5. Interpretasi statistik dalam penelitian menunjukkan bahwa riset tersebut dilakukan dengan metode...!
 - a. Kualitatif
 - b. Kuantitatif
 - c. Evaluasi
 - d. *Mixed method*

Kunci jawaban dan pedoman penskoran

1. D
2. C
3. C
4. A
5. B

Bobot

Latihan soal essay 75 % dengan bobot per soal 15.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{50}{75} \times 100 = \mathbf{66,67}$$

Tes Formatif Pilihan Berganda 25% dengan bobot per soal 5.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{20}{25} \times 100 = \mathbf{80}$$

$$66,67 \times 75\% = \mathbf{50}$$

$$80 \times 25\% = \mathbf{20}$$

$$50 + 20 = \mathbf{70}$$

$$\text{Nilai akhir} = \mathbf{70}$$

MODUL 6

TEORI EVALUASI DAN PERAN TEORI DALAM EVALUASI

Metode Pembelajaran	Estimasi waktu	Capaian Pembelajaran
-Presentasi virtual -Diskusi Virtual interaktif -Tanya jawab	120 menit X 1	- Mahasiswa memahami teori evaluasi dan peran teori dalam evaluasi

Materi 6. Teori Evaluasi dan Peran Teori dalam Evaluasi.

A. Konsep Teori Evaluasi

Dalam penelitian salah satu bagian penting adalah teori. Teori menjadi pijakan bagi peneliti untuk melakukan kajian lebih dalam tentang objek penelitian yang akan dilakukan. Sebagai pijakan tentu teori akan memberi warna terhadap hasil penelitian. Seorang peneliti akan lebih mudah dalam kegiatan penelitian dengan adanya teori sebagai pemandu dalam meneliti. Hal ini dapat dilihat karena teori menurut McMilan *“A theory can be defined as a set of propositions that explain the relationships among observed phenomena.”*³⁴ Teori adalah dalil-dalil yang dapat menjelaskan hubungan diantara fenomena yang diteliti. Sebagai peneliti kita membutuhkan dasar pijakan yang dapat mendukung dan menguatkan argumentasi kita dalam penelitian, sebagai contoh mengapa variabel A berpengaruh terhadap variable B. Dibutuhkan teori yang cukup untuk mengarahkan peneliti memahami lebih dalam hubungan antar variable tersebut. Pada saat peneliti akan meneliti tentang pengaruh Kepemimpinan terhadap terhadap disiplin pegawai misalnya, peneliti membutuhkan teori-teori yang ada atau

³⁴ James H. McMilan. *Educational Research: Fundamentals for the Consumer*. (USA: Pearson Education, Inc. 2008), hlm. 7.

berdasarkan hasil penelitian terdahulu yang menyatakan bahwa kepemimpinan memiliki pengaruh terhadap kedisiplinan pegawai. Tanpa adanya teori dan hasil penelitian terdahulu yang menyatakan hal tersebut, penelitian dan hasilnya akan menjadi kurang memiliki kredibilitas, karena ketiadaan teori yang mendukung.

Sugiyono mendefinisikan teori sebagai berikut; 1) teori itu berkenaan dengan konsep, asumsi dan generalisasi yang logis, 2) berfungsi untuk mengungkapkan, menjelaskan dan memprediksi perilaku yang memiliki keteraturan, 3) sebagai stimulant dan panduan untuk mengembangkan pengetahuan.³⁵ Definisi yang disampaikan Sugiyono sangat lengkap dan jelas bahwa teori memiliki peran yang sangat krusial dalam penelitian. Dengan teori, peneliti akan dibimbing dalam kegiatan penelitian supaya tidak sekedar coba-coba dalam proses penggalian data dan informasi. Teori juga akan memberi arah yang jelas kemana penelitian akan dibawa.

Creswell memberi definisi teori sebagai “seperangkat konstruk (atau variabel) yang saling berhubungan, yang berasosiasi dengan proposisi atau hipotesis yang memerinci hubungan antar variabel (biasanya dalam konteks *magnitude* atau *direction*). Suatu teori dalam penelitian bisa saja berfungsi sebagai argumentasi, pembahasan, atau alasan. Teori biasanya membantu menjelaskan (atau memprediksi) fenomena yang muncul di dunia.”³⁶

Berdasarkan beberapa pendapat mengenai teori yang telah disampaikan di atas, kita dapat menarik benang merah bahwa teori merupakan kerangka berpikir beserta dalil-dalil yang sah untuk digunakan sebagai rujukan dalam memprediksi dan argumentasi mengenai hubungan atau pengaruh antar variabel. Kerangka berpikir merupakan konsep yang menjelaskan suatu fenomena, dengan didukung oleh dalil-dalil yang kuat mengenai keterkaitan antar variabel. Dalil atau proposisi menjadi bagian yang penting dan mendasar dalam teori karena melalui proposisi-proposisi tersebut akan

³⁵ Sugiyono. *Metode Penelitian Administrasi*, edisi ke-12, (Bandung: Alfabeta, 2005), hh. 55-56.

³⁶ John W. Creswell. *Research Design; Pendekatan Kualitatif, Kuantitatif, dan Mixed*, (Yogyakarta: Pustaka Pelajar, 2010), h. 79.

terbangun kredibilitas hasil penelitian yang sesungguhnya. Proposisi yang masih berbentuk konsep, membutuhkan data-data pendukung yang dapat memperkuat argumentasi, sehingga konsep yang terbangun tidak hanya logis tetapi secara faktual memiliki bukti yang dapat dipertanggungjawabkan secara ilmiah.

Scriven menyebutkan *“What is evaluation? Synthesizing what the dictionaries and common usage tell us, it is the process of determining the merit, worth, or significance of things (near-synonyms are quality/value/importance).”*³⁷ Apa itu evaluasi? Sintesis apa yang kamus dan masyarakat sampaikan kepada kita, yaitu proses menentukan nilai, manfaat atau derajat signifikansi tentang sesuatu (mendekati sinonim mutu/nilai/hal yang penting). Kata kunci yang dapat diambil dari definisi ini adalah proses menentukan nilai, manfaat dan derajat signifikansi barang atau sesuatu. Tentu dibutuhkan langkah-langkah yang tepat dan sesuai tahapan dalam mendapatkan pengetahuan secara ilmiah. Hasil penelitian harus memiliki nilai tambah baik dari aspek keilmuan maupun dari sisi manfaat bagi objek penelitian dan juga subjek penelitian. Terutama bagi pengguna manfaat penelitian tersebut (para *stakeholder*).

MICHAEL SCRIVEN	
 	
<p>Michael Scriven is a philosopher of science who has contributed extensively to the growth of the evaluation profession. He has sharply criticized both classical and more recent conceptualizations of evaluation. He has grounded his consumerist view of evaluation in a basic philosophical position and has evolved concepts and methods to help articulate and apply his approach. He has also been one of the foremost leaders in the effort to professionalize evaluation work</p> <p>He was born and raised in Victoria, Australia. He took his master of arts degree in applied mathematics and symbolic logic at the University of Melbourne and completed his Ph. D. in philosophy of science at Oxford University. He served in professional roles at the University of Minnesota from 1952 to 1956, at Swarthmore College from 1956 to 1960, Indiana University from 1960 to 1965, University of California at Berkeley from 1965 to 1975, and University of San Francisco from 1975 to 1982. At this writing, he was Distinguished University Professor at the University of Western Australia in Perth.</p>	

Profil 1: Michael Scriven

³⁷ MICHAEL SCRIVEN. *International Handbook of Educational Evaluation*, T. Kellaghan, D.L. Stufflebeam (eds.), (Dordrecht: Kluwer Academic Publishers, 2003), P. 15.

Scriven juga menjelaskan bahwa *“An evaluation must, by definition, lead to a particular type of conclusion - one about merit, worth, or significance - usually expressed in the language of good/bad, better/worse, well/ill, elegantly/poorly etc”*.³⁸ Evaluasi harus di artikan sebagai suatu proses yang dapat mengantarkan kepada kesimpulan khusus-satu tentang nilai, manfaat, atau hal yang penting-biasanya diungkapkan dalam Bahasa yang baik/buruk, lebih baik/lebih buruk, elegan/jelek dll. Scriven ingin menegaskan bahwa evaluasi bukan sekedar kegiatan mengumpulkan data dan informasi yang tidak bermakna, tetapi data dan informasi yang dihasilkan harus benar-benar dapat diolah dan menjadi dasar untuk mengambil keputusan yang tepat dan solutif bagi objek evaluasi. Data dan informasi akan bermakna jika proses pengumpulan data tersebut dilakukan dengan cara yang benar sesuai standar yang ada dalam kegiatan evaluasi.

Masih menurut Scriven *“An evaluation theory (or theory of evaluation) can be of one or the other of two types. Normative theories are about what evaluation should do or be, or how it should be conceived or defined. Descriptive theories are about what evaluations there are, or what evaluations types there are (classificatory theories), and what they in fact do, or have done, or why or how they did or do that (explanatory theories)”*.³⁹ Teori evaluasi bisa jadi satu atau dua tipe teori yang ada. Teori normative adalah tentang apa yang seharusnya dilakukan oleh evaluasi, atau bagaimana evaluasi definisikan atau dipahami. Teori deskriptif berbicara mengenai evaluasi apa yang digunakan, atau tipe evaluasi apa (berkaitan dengan teori penggolongan/klasifikasi), dan apa yang sesungguhnya dapat dilakukan dilakukan dengan teori itu, atau yang telah dilakukan, atau mengapa dan bagaimana mereka melakukannya (tentang teori penjelasan). Scriven mengemukakan bahwa teori evaluasi ada dua macam yaitu normative dan deskriptif. Jika teori normative berbicara tentang apa yang dapat dilakukan oleh

³⁸ Scriven. *International Handbook of Educational Evaluation*, 15-30, T. Kellaghan, D.L. Stufflebeam (eds.), (Dordrecht: Kluwer Academic Publishers, 2003). P.16.

³⁹ *Ibid.*, P. 15

evaluasi, maka teori deskriptif berbicara tentang jenis teori yang dapat digunakan untuk mengevaluasi objek evaluasi yang dimaksud.

Stufflebeam mendefinisikan evaluasi *“Evaluation is the process of delineating, obtaining, providing, and applying descriptive and judgmental information about the merit and worth of some object's goals, design, implementation, and outcomes to guide improvement decisions, provide accountability reports, inform institutionalization/ dissemination decisions, and improve understanding of the involved phenomena.”*⁴⁰

Evaluasi adalah proses mendeskripsikan, mendapatkan, menyediakan, dan menerapkan dan menyatakan pendapat tentang informasi mengenai nilai dan manfaat tujuan sesuatu, design, implementasi, dan hasil untuk membimbing kepada perbaikan keputusan, menyediakan laporan pertanggungjawaban, menyampaikan keputusan institusi, dan memperbaiki pemahaman fenomena yang meliputinya. Pada halaman yang sama Stufflebeam juga menyampaikan penjelasan

*This definition summarizes the key ideas in the CIPP Model. The definition posits four purposes for evaluation: guiding decisions; providing records for accountability; informing decisions about installing and/or disseminating developed products, programs, and services; and promoting understanding of the dynamics of the examined phenomena. It says the process of evaluation includes four main tasks: delineating, obtaining, providing, and applying information.”*⁴¹

Bahwa definisi evaluasi di atas adalah rangkuman inti dari model CIPP. Definisi yang dimaksud memposisikan empat tujuan evaluasi: petunjuk pengambilan keputusan, bentuk pertanggungjawaban; menyampaikan keputusan tentang pengenalan produk yang dikembangkan, program, dan jasa; dan mempromosikan pemahaman yang dinamis mengenai fenomena yang di uji. Dapat dikatakan bahwa proses evaluasi meliputi empat tugas pokok: mendeskripsikan, memperoleh, menyediakan, dan menerapkan informasi. Mendeskripsikan merupakan usaha seorang

⁴⁰ Daniel Stufflebeam. *International Handbook of Educational Evaluation*, 15-30, T. Kellaghan, D.L. Stufflebeam (eds.), (Dordrecht: Kluwer Academic Publishers, 2003). p. 34.

⁴¹*Ibid.*, P. 34.

evaluator menyampaikan hasil evaluasinya berdasarkan informasi dan data yang diperoleh melalui proses pengumpulan data. Dari data dan informasi yang terkumpul, kemudian diolah menjadi temuan penelitian yang kemudian akan digunakan sebagai bahan pertimbangan untuk membuat keputusan.

Shadish (1998) dikutip oleh Merten dan Wilson mengatakan bahwa *“has described evaluation theory as “who we are”, in the sense that it gives us the language we use in this transdiscipline to describe what we do uniquely as evaluators.”*⁴² Shadish mendeskripsikan teori evaluasi sebagai “siapa kita”, dalam arti bahwa teori evaluasi memberi kita pemahaman yang dapat kita gunakan dalam berbagai jenis disiplin ilmu untuk menggambarkan apa yang kita lakukan secara unik sebagai evaluator.

Shadish (1998) “Evaluation Theory Is Who We Are”

What we say about what we do (our theory) is just as important as—and deeply informs--*what we do* (our practice) and *who we are* (our profession) as evaluators



Profil 2: Shadish

Stufflebeam & Coryn mengemukakan *“A program evaluation theory is a coherent set of conceptual, hypothetical, pragmatic, and ethical principles forming a general framework to guide the study and practice of program*

⁴² Donna M. Mertens dan Amy T. Wilson. *Program Evaluation Theory and Practice: Second Edition A Comprehensive Guide*, (New York: The Guilford Press, 2019), P. 39.

evaluation.⁴³ Teori evaluasi program adalah serangkaian konsep yang utuh, hipotetical, pragmatis, dan prinsip-prinsip etika yang membentuk kerangka umum untuk membimbing penelitian dan praktik evaluasi program. Konsep yang utuh menunjukkan bahwa dibutuhkan berbagai bidang ilmu untuk mengembangkan landasan penelitian evaluasi. Hal ini karena evaluasi merupakan bidang ilmu yang berkembang dari berbagai latar belakang keilmuan. Hipotesis dapat ditarik kearah suatu evaluasi merupakan bagian dari penelitian yang harus dilakukan dengan prosedur dan langkah-langkah sistematis dalam mengumpulkan data dan informasi serta dalam pengolahan datanya. Pragmatis tentu mengarah pada evaluasi sesungguhnya ditujukan untuk membuat keputusan tentang objek evaluasi. Dan dalam kegiatan evaluasi tentu harus mengikuti pedoman etika yang berlaku secara umum, termasuk diantaranya adalah objektif, jujur, tanggungjawab, dan transparan.

Sementara Smith (2008) dikutip oleh Mertens dan Wilson mengatakan *“Evaluation theory is that aspect that reflects “our thinking about how and why we engage in evaluation. Is the purpose of evaluation validation, accountability, monitoring, or improvement and development?”*.⁴⁴ Teori evaluasi adalah bahwa aspek yang merefleksikan “pikiran kita tentang bagaimana dan mengapa kita terlibat dalam evaluasi. Apakah tujuan evaluasi adalah melakukan validasi, pertanggungjawaban, monitoring, atau perbaikan dan pengembangan?

Mertens dan Wilson mengatakan *“There are theories of evaluation, program theories, and social science theories that inform our work.”*⁴⁵ Terdapat beberapa teori evaluasi, teori program dan teori social yang dapat mendukung penelitian kita.

Social science theories are inclusive of such areas as development, learning, motivation, and social change; and identity-based theories such as feminist, queer, and critical race theories. These are used

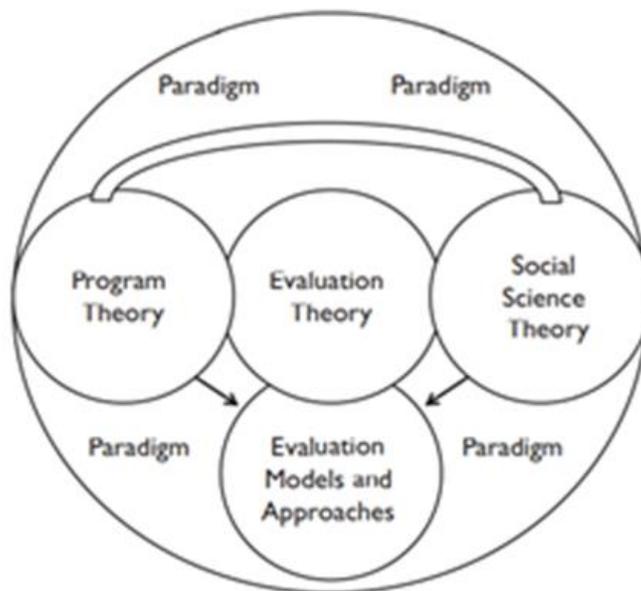
⁴³ Daniel L. Stufflebeam dan Chris L. S. Coryn. *Evaluation Theory, Models, and Applications*, Second Edition. (San Francisco: Jossey-Bass A Wiley Brand), p. 50.

⁴⁴ Donna M. Mertens & Amy T. Wilson. *Program Evaluation Theory and Practice, second edition: A Comprehensive Guide*, (New York: The Guilford Press, 2019), P. 39.

⁴⁵ *Ibid.*, P. 36.

*both to inform decisions about evaluation practice and to inform programmatic decisions. Program theories help explain the mechanisms believed to influence the achievement of the desired program outcomes. This is sometimes referred to as the theory of change—in other words, what needs to be done with which resources by whom in order to achieve the desired change.*⁴⁶

Teori sosial merupakan bidang inklusif, wilayahnya seperti pengembangan, pembelajaran, motivasi, dan perubahan social; dan teori berbasis indentitas seperti aliran feminis, homo, dan teori tentang ras/suku bangsa. Semuanya digunakan untuk memberikan informasi keputusan terkait praktik evaluasi dan untuk menginformasikan keputusan pragmatis. Teori program membantu menjelaskan mekanisme yang diyakini bisa mempengaruhi prestasi hasil program yang diinginkan. Kadangkala merujuk pada teori perubahan-dengan kata lain, apa yang perlu dilakukan dengan sumber daya yang seperti apa agar supaya dapat mencapai perubahan yang diinginkan.



Gambar 2: Paradigms; Program, Evaluation, and Social Science Theories; and Evaluation Models and approaches. Sumber: Donna M. Mertens & Amy T. Wilson, 2019.

⁴⁶ Donna M. Mertens & Amy T. Wilson. *Program Evaluation Theory and Practice, second edition: A Comprehensive Guide*, (New York: The Guilford Press, 2019), pp. 36-38.

Teori program

Merten dan Wilson mengutip pendapat Stufflebeam and Shinkfield (2007) yang mengatakan *“state that a program evaluation theory should have six components: “overall coherence, core concepts, tested hypotheses on how evaluation procedures produce desired outcomes, workable procedures, ethical requirements, and a general framework for guiding program evaluation practice and conducting research on program evaluation”*.⁴⁷ Pernyataan bahwa teori evaluasi program seharusnya memiliki enam komponen: keseluruhan yang padu, konsep dasar, hipotesis yang teruji terhadap bagaimana prosedur-prosedur evaluasi dapat menghasilkan produk yang diinginkan, prosedur dapat berjalan sesuai standar, etika, dan kerangka umum untuk membimbing pelaksanaan evaluasi program dan melakukan penelitian evaluasi program. Sebuah teori harus sudah teruji dapat menghasilkan produk yang diinginkan sesuai dengan prosedur yang ada. Karena teori akan digunakan oleh masyarakat secara umum, maka perlu mengikuti standar etika yang berlaku dalam penelitian dan memiliki kerangka dasar yang dapat digunakan atau sebagai rujukan peneliti lain dalam melakukan penelitian.

Teori social science

Donaldson dan Lipsey (2006) dikutip oleh Mertens dan Wilson *“suggest that the use of social science theories can contribute to the development of interventions that reflect collective wisdom, and hence can reduce wasted time on treatments already known to be ineffective”*.⁴⁸ Disarankan bahwa penggunaan teori ilmu social dapat berkontribusi terhadap intervensi perkembangan yang menrefleksikan kebijaksanaan kolektif, dan oleh sebab itu dapat mengurangi pemborosan waktu pada perlakuan yang sudah diketahui keefektifannya.

Mertens dan Wilson menegaskan bahwa *“A caveat that must be mentioned is the difficulty of transferring social science theories across cultures. As obvious as it seems, evaluators must be cognizant of unique*

⁴⁷ Donna M. Mertens & Amy T. Wilson. *Program Evaluation Theory and Practice, second edition: A Comprehensive Guide*, (New York: The Guilford Press, 2019), P. 40.

⁴⁸ *Ibid.*, p. 43.

cultural factors that might limit their ability to generalize from a theory developed on a white, middle-class group of people from the United States to other populations and contexts".⁴⁹ Keberatan yang harus disebutkan adalah kesulitan mentransfer teori social secara lintas budaya. Sangat jelas terlihat, bahwa evaluator harus mengetahui keunikan factor budaya yang mungkin membatasi kemampuan mereka melakukan generalisasi dari teori yang dikembangkan di masyarakat kulit putih, masyarakat kelas menengah di US kepada populasi lain dan dengan konteks yang berbeda.

Peran evaluator

Menurut Ryan and Schwandt (2002) dikutip oleh Mertens dan Wilson "*connect evaluation theory with evaluators' roles quite directly: "The concept of the 'role of the evaluator' is central to the theory and practice of evaluation*".⁵⁰ Mengaitkan teori evaluasi dengan peran evaluator secara langsung: konsep tentang peran evaluator adalah pusatnya teori dan praktik evaluasi. Dalam hal ini tidak dapat dipisahkan bahwa teori yang berkembang merupakan peran evaluator sebagai pelaku dalam kegiatan evaluasi.

Mertens dan Wilson menegaskan bahwa "*.....evaluators play many different roles, depending on the stage of the evaluation and the demands of the situation. They examine what they consider to be a generic list of activities that most evaluators are called upon to pursue in an external evaluation, and then describe dominant and supportive roles for an evaluator.*"⁵¹ Evaluator memiliki peran yang berbeda-beda, tergantung pada tahapan evaluasi dan situasinya. Mereka mengecek apakah yang mereka anggap sebagai daftar aktifitas umum yang sebagian besar evaluator mempersilakan eksternal evaluator untuk terlibat, kemudian mendeskripsikan peran dominan dan peran pendukung bagi evaluator.

⁴⁹ Donna M. Mertens & Amy T. Wilson. *Program Evaluation Theory and Practice, second edition: A Comprehensive Guide*, (New York: The Guilford Press, 2019), p. 43.

⁵⁰ *Ibid*, p. 44.

⁵¹ *Ibid*, p. 44.

Kriteria teori evaluasi yang baik menurut Shadish et al. (1991) yang dikutip oleh Mertens & Wilson harus memenuhi kriteria sebagai berikut.

- *Knowledge: What do we need to do to produce credible knowledge?*
- *Use: How can we use the knowledge we gain from an evaluation?*
- *Valuing: How do we construct our value judgments?*
- *Practice: What do we evaluators actually do in practice?*
- *Social programming: What is the nature of social programs and their roles in solving societal problems?*⁵²

Knowledge adalah tentang apa yang harus kita lakukan untuk menghasilkan pengetahuan yang kredibel. *Use* memiliki pesan mengenai bagaimana kita bisa menggunakan pengetahuan yang kita peroleh dari kegiatan evaluasi. *Valuing* berkenaan dengan bagaimana kita membangun nilai suatu keputusan. Sementara *practice* adalah tentang apa yang sesungguhnya dilakukan evaluator dalam praktiknya. *Social programming* berbicara tentang apa asal mula program sosial dan perannya dalam menyelesaikan masalah-masalah sosial.

B. Peran Teori dalam Evaluasi

Dalam kegiatan penelitian murni maupun penelitian evaluasi diperlukan teori sebagai dasar rujukan dalam proses pengumpulan data dan informasi sehingga memiliki arah yang sesuai standar ilmiah. Karena hasil penelitaian harus dapat dipertanggungjawabkan secara ilmiah. Meskipun masih terdapat perdebatan apakah penelitian evaluasi perlu menggunakan landasan teori atau tidak. Sebagian mengatakan bahwa teori dalam evaluasi diperlukan, sebagian yang lain mengatakan tidak perlu teori dalam kegiatan evaluasi. Jika merujuk pada kegiatan akademik, maka kita yakin bahwa teori sangat dibutuhkan dalam evaluasi.

Pendapat tentang pentingnya teori dalam evaluasi disampaikan oleh para pakar evaluasi diantaranya; Alkin, Chen, Donaldson, Fetterman,

⁵² Donna M. Mertens & Amy T. Wilson. *Program Evaluation Theory and Practice, second edition: A Comprehensive Guide*, (New York: The Guilford Press, 2019), pp. 39-40.

Lipsey, Mark, Rossi, Freeman, Shadish, Cooc & Campbell dan Weiss (Wiarwan, 2016). Wirawan sendiri dalam bukunya evaluasi teori, model, metodologi, standar, aplikasi dan profesi (2016) berpendapat bahwa sebagai suatu cabang ilmu riset, evaluasi memerlukan dukungan teori evaluasi dan teori-teori ilmu social yang berkaitan.

Mertens dan Wilson mengatakan bahwa *“Theories provide guidance in determining the purposes for evaluations, as well as in defining what we consider to be acceptable evidence for making decisions in an evaluation”*.⁵³ Teori memberi petunjuk dalam mengetahui tujuan evaluasi, serta dalam mendefinisikan apa yang kita anggap dapat diterima sebagai dasar pengambilan keputusan dalam evaluasi. Paling tidak ada dua hal yang penting dalam teori menurut Mertens dan Wilson yaitu menentukan tujuan evaluasi dan menentukan apakah data dan informasi yang terkumpul dapat diterima atau kredibel.

Mertens dan Wilson menyebutkan *“Theories provide frameworks for thinking about the interrelationships of constructs and are more limited in scope than paradigms. Hence a variety of theoretical perspectives can be associated with a particular paradigm”*.⁵⁴ Teori memberikan kerangka berpikir tentang hubungan antara konstruk dan lebih membatasi bidang penelitian daripada paradigma. Oleh karenanya, berbagai perspektif teori yang digunakan dapat dikaitkan dengan paradigma tertentu. Mertens ingin menyampaikan bahwa dengan teori, peneliti akan lebih terarah dalam penelitiannya. Bidang yang diteliti menjadi jelas dan tidak melebar sehingga hasil penelitian akan lebih tegas objeknya. Teori juga memberikan penjelasan mengenai adanya hubungan antar variable yang akan diteliti. Dengan ini, seorang peneliti akan diarahkan dan dibimbing untuk melakukan penggalian data yang sesuai dan menentukan responden penelitian dengan lebih mudah. Dengan teori penelitian dapat lebih focus pada objek yang akan diteliti. Peneliti dapat lebih terarah mengenai lingkup penelitiannya. Karena

⁵³ Donna M. Mertens & Amy T. Wilson. *Program Evaluation Theory and Practice, second edition: A Comprehensive Guide*, (New York: The Guilford Press, 2019), P. 39.

⁵⁴ *Ibid.*, p.36.

sangat mungkin dalam suatu penelitian akan menjadi bias dan melebar ke area yang sangat luas jika tidak dipandu dengan teori.

Lebih lanjut Stufflebeam dan Coryn mengatakan *“a program evaluation theory to be useful (but not sufficient) for considering the scope and rigor required by sound theories of program evaluation”*.⁵⁵ Teori evaluasi program akan menjadi sangat berguna (tetapi belum cukup) sebagai pertimbangan terkait jangkauan dan ketepatan yang dibutuhkan oleh teori-teori program evaluasi. Bahwa teori evaluasi program hanyalah salah satu teori yang perlu dukungan teori lain untuk menguatkan dasar penelitian yang akan dilakukan. Dengan variasi teori yang menjadi rujukan, penelitian evaluasi yang dilakukan menjadi kuat argumentasinya.

Menurut Stufflebeam dan Coryn *“A program evaluation theory is a coherent set of conceptual, hypothetical, pragmatic, and ethical principles forming a general framework to guide the study and practice of program evaluation.”*⁵⁶ Teori evaluasi program adalah serangkaian konsep yang utuh, hipotetis, pragmatis, dan prinsip-prinsip etika yang membentuk kerangka umum untuk memandu penelitian dan praktik evaluasi program. Stufflebeam menegaskan bahwa pentingnya teori dalam evaluasi, hal mana teori dapat menjadi petunjuk bagi evaluator dalam melakukan evaluasi.

Dalam teori ada konsep yang dapat dijadikan dasar bagi peneliti untuk mengidentifikasi indikator objek penelitian. Peneliti juga akan dipandu dalam membuat hipotesis serta memahami prinsip-prinsip etika dalam penelitian. Dan tentu saja teori akan menjadi panduan pragmatis dalam penelitian evaluasi pada khususnya.

Dengan indikator yang ada pada penjelasan konsep di atas tentu suatu teori memiliki kekhasan. Berikut ciri-ciri teori evaluasi program menurut Stufflebeam dan Coryn *“According to this definition, a sound program evaluation theory has six main features: overall coherence, core concepts, tested hypotheses concerning how evaluation procedures*

⁵⁵Daniel L. Stufflebeam & Chris L. S. Coryn. *Evaluation Theory, Models, and Applications: Second Edition*, (San Fransisco: Published by Jossey-Bass A Wiley Brand), P.50.

⁵⁶ *Ibid.*, P.50.

produce desired outcomes, workable procedures, ethical requirements, and a general framework for guiding program evaluation practice and conducting research on program evaluation".⁵⁷ Menurut definisi teori tersebut di atas, teori evaluasi program memiliki enam ciri yaitu: hubungan yang menyeluruh, konsep inti, hipotesis telah teruji terhadap bagaimana prosedur evaluasi menghasilkan sesuatu yang diharapkan, prosedur bekerja sesuai standar, menggunakan etika penelitian, dan kerangka umum sebagai petunjuk pelaksanaan evaluasi program dan melaksanakan penelitian evaluasi program.

Wirawan (2016) mengutip pendapat Shadish bahwa paling tidak ada enam peran teori evaluasi dalam kegiatan evaluasi.

1. Teori evaluasi memiliki bahasa yang dapat digunakan oleh para evaluator untuk berkomunikasi satu dengan yang lain mengenai evaluasi.
2. Teori evaluasi meliputi banyak hal dalam bidang evaluasi yang menjadi perhatian mendalam para evaluator.
3. Teori evaluasi membahas tema-tema pokok dalam konferensi evaluasi profesional.
4. Teori evaluasi menunjukkan identitas yang berbeda bagi evaluator dengan identitas profesional lainnya.
5. Teori evaluasi menjadi ciri khas evaluator yang dapat disampaikan kepada dunia luar.
6. Teori evaluasi merupakan dasar pengetahuan yang mendefinisikan profesi evaluator.⁵⁸

Mertens dan Wilson mengatakan "*Program theories help explain the mechanisms believed to influence the achievement of the desired program outcomes. This is sometimes referred to as the theory of change—in other words, what needs to be done with which resources by whom in order to achieve the desired change.*"⁵⁹ Teori-teori program membantu menjelaskan mekanisme yang diyakini dapat mempengaruhi

⁵⁷Daniel L. Stufflebeam & Chris L. S. Coryn. *Evaluation Theory, Models, and Applications*: Second Edition, (San Fransisco: Published by Jossey-Bass A Wiley Brand), p. 50.

⁵⁸ Wirawan, *Evaluasi, teori, Model, Metodologi, Standar, Aplikasi dan profesi*, (Jakarta: Rajawali Pers, 2016), h. 40.

⁵⁹Donna M. Mertens & Amy T. Wilson. *Program Evaluation Theory and Practice, second edition: A Comprehensive Guide*, (New York: The Guilford Press, 2019), P. 38.

pencapaian hasil yang diinginkan dari suatu program. Kadang-kadang hal seperti ini merujuk pada teori perubahan-dengan kata lain, apa yang perlu dilakukan dengan sumber daya yang ada dan siapa yang terlibat dalam upaya mencapai perubahan yang diinginkan. Kata kuncinya adalah bahwa teori memiliki peran penting dalam evaluasi untuk membantu peneliti menentukan hal-hal apa saja yang perlu dilakukan sehingga perubahan yang diinginkan terhadap objek evaluasi dapat di capai. Karena tujuan akhir suatu evaluasi adalah membuat keputusan terkait objek evaluasi.

Pada bagian lain terdapat beberapa pakar evaluasi yang menyebutkan bahwa evaluasi tidak terlalu membutuhkan teori. Mengutip tulisan Wirawan dalam bukunya *Evaluasi: teori, Model, Metodologi, Standar, Aplikasi dan Profesi* Salah satu tokoh evaluasi yang berpendapat bahwa evaluasi tidak terlalu membutuhkan teori adalah Scriven. Michael Scriven menyatakan bahwa evaluator mungkin melakukan evaluasi program dengan baik tanpa mempergunakan teori evaluasi atau teori program. Pikiran evaluator yang salah menurut Scriven adalah bahwa dalam melaksanakan evaluasi ia harus mempunyai logika teori evaluasi dan teori program.

RANGKUMAN

Teori merupakan kerangka berpikir beserta dalil-dalil yang sah untuk digunakan sebagai rujukan dalam memrediksi dan argumentasi mengenai hubungan atau pengaruh antar variabel.

Penggunaan teori dalam penelitian evaluasi masih menyisakan perbedaan pendapat ada yang mengatakan teori bagian penting dalam evaluasi, disisi lain ada yang berpendapat bahwa evaluasi tidak terlalu membutuhkan teori.

Teori memiliki peran yang sangat mendasar dalam evaluasi, sebagai landasan berpikir mengenai hubungan atau pengaruh antar variable dan sebagai pedoman bagi evaluator dalam menentukan tujuan evaluasi, karena teori menyediakan konsep, petunjuk, dan standar yang harus diikuti oleh peneliti dalam melakukan penelitian.

Latihan 6

1. Jelaskan apa yang dimaksud dengan teori!
2. Jelaskan peran evaluator dalam penelitian evaluasi!
3. Jelaskan metode apa saja yang dapat digunakan dalam evaluasi!
4. Sebutkan peran teori dalam penelitian evaluasi!
5. Sebutkan standar teori evaluasi yang baik!

Pembahasan Latihan Soal 6

1. Teori merupakan kerangka berpikir beserta dalil-dalil yang sah untuk digunakan sebagai rujukan dalam memrediksi dan argumentasi mengenai hubungan atau pengaruh antar variabel.
2. Evaluator sebagai subjek penelitian dalam kegiatan evaluasi memiliki peran sentral yang tidak tergantikan. Evaluator akan terlibat dalam keseluruhan proses evaluasi mulai dari tahap persiapan, pelaksanaan, dan pelaporan.
3. Metode kuantitatif, Kualitatif, dan Mixed Method
4. Peran teori evaluasi dalam kegiatan evaluasi;
 - a. Teori evaluasi memiliki bahasa yang dapat digunakan oleh para evaluator untuk berkomunikasi satu dengan yang lain mengenai evaluasi.
 - b. Teori evaluasi meliputi banyak hal dalam bidang evaluasi yang menjadi perhatian mendalam para evaluator.
 - c. Teori evaluasi membahas tema-tema pokok dalam konferensi evaluasi profesional.
 - d. Teori evaluasi menunjukkan identitas yang berbeda bagi evaluator dengan identitas professional lainnya.
 - e. Teori evaluasi menjadi ciri khas evaluator yang dapat disampaikan kepada dunia luar.
 - f. Teori evaluasi merupakan dasar pengetahuan yang mendefinisikan profesi evaluator.
5. Kriteria teori yang baik harus memiliki unsur sebagai berikut; *Knowledge* adalah tentang apa yang harus kita lakukan untuk menghasilkan pengetahuan yang kredibel. *Use* memiliki pesan mengenai bagaimana kita bisa menggunakan pengetahuan yang kita peroleh dari kegiatan evaluasi. *Valuing* berkenaan dengan bagaimana

kita membangun nilai suatu keputusan. Sementara *practice* adalah tentang apa yang sesungguhnya dilakukan evaluator dalam praktiknya. *Social programming* berbicara tentang apa asal mula program sosial dan perannya dalam menyelesaikan masalah-masalah sosial.

Test Formatif Pilihan Berganda

1. Dari beberapa tokoh evaluasi berikut yang menganggap teori tidak terlalu penting dalam penelitian evaluasi adalah...?
 - a. Shadish
 - b. Michael Scriven
 - c. Donna M. Mertens
 - d. Daniel L. Stufflebeam
2. Salah satu alasan pentingnya teori dalam evaluasi adalah....?
 - a. Teori evaluasi program adalah serangkaian konsep yang utuh, hipotetis, pragmatis, dan prinsip-prinsip etika yang membentuk kerangka umum untuk memandu penelitian dan praktik evaluasi program.
 - b. Evaluator mungkin melakukan evaluasi program dengan baik tanpa mempergunakan teori evaluasi atau teori program.
 - c. Teori tidak dapat menjadi rujukan dalam penelitian evaluasi yang dilakukan sehingga argumentasinya lemah.
 - d. Teori hanya konsep yang dapat mendistorsi pemahaman peneliti terhadap objek penelitian.
3. Evaluator yang setuju dengan peran teori dalam evaluasi memiliki argument berikut, kecuali...!
 - a. Teori program membantu menjelaskan mekanisme yang diyakini dapat mempengaruhi pencapaian hasil yang diinginkan dari suatu program.
 - b. Teori evaluasi program mendeskripsikan pertentangan yang menyeluruh, konsep inti, dan hipotesis sehingga dapat mengaburkan makna.
 - c. Teori evaluasi program sangat berguna sebagai pertimbangan jangkauan dan ketepatan dalam evaluasi.
 - d. Teori sebagai petunjuk pelaksanaan evaluasi program.

4. Salah satu tujuan dilakukannya evaluasi adalah...!
 - a. Membuat keputusan terhadap program atau objek evaluasi.
 - b. Membuat generalisasi hasil evaluasi.
 - c. Memrediksi hubungan sebab akibat.
 - d. Mengembangkan teori.
5. Komponen-komponen yang harus ada dalam teori evaluasi program adalah sebagai berikut, kecuali...!
 - a. Konsep dasar.
 - b. Hipotesis belum teruji.
 - c. Keseluruhan yang padu.
 - d. Prosedur dapat berjalan sesuai standar.

Kunci jawaban dan pedoman penskoran

1. B
2. A
3. B
4. A
5. B

Bobot

Latihan soal essay 75 % dengan bobot per soal 15.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{50}{75} \times 100 = \mathbf{66,67}$$

Tes Formatif Pilihan Berganda 25% dengan bobot per soal 5.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{20}{25} \times 100 = \mathbf{80}$$

$$66,67 \times 75\% = \mathbf{50}$$

$$80 \times 25\% = \mathbf{20}$$

$$50 + 20 = \mathbf{70}$$

$$\text{Nilai akhir} = \mathbf{70}$$

MODUL 7

PENGERTIAN KONSEP TES DAN PENGUKURAN

Metode Pembelajaran	Estimasi waktu	Capaian Pembelajaran
-Presentasi virtual -Diskusi Virtual interaktif -Tanya jawab	120 menit X 1	- Mahasiswa memahami Pengertian konsep tes, pengukuran, pengujian.

Materi 7. Pengertian konsep tes dan pengukuran.

A. Pengertian Konsep Tes

Dalam Pendidikan kita cukup familiar dengan istilah tes. Biasanya dilakukan oleh seorang guru kepada siswanya setelah materi pelajaran diberikan. Guru akan mencoba mengetahui tingkat pemahaman siswa setelah mengikuti kegiatan belajar. Kegiatan tes ini sesungguhnya salah satu bentuk evaluasi terutama evaluasi pembelajaran. Karena sifat tes ini untuk mengetahui pemahaman siswa atau daya ingat siswa setelah menempuh pembelajaran. Hingga saat ini, tes menjadi satu-satunya alat ukur kemajuan belajar siswa setelah menempuh pembelajaran. Tes atau testing merupakan prosedur sistematis yang direncanakan oleh evaluator guna membandingkan antar perilaku yang dievaluasi. Dalam evaluasi program, tes bisa dilakukan di awal atau tengah kegiatan program atau proyek yang disebut sebagai *pre-test* atau akhir kegiatan program atau proyek yang disebut dengan *post-test*.⁶⁰ Tes yang dimaksud dapat dilakukan dalam berbagai bentuk, tertulis, lisan, atau juga dalam bentuk praktik.

⁶⁰ Sukardi. Evaluasi Program Pendidikan dan Kepelatihan, (Jakarta: PT Bumi Aksara, 2015), h. 92.

Menurut McMilan “A test is an instrument that requires subjects to complete a cognitive task by responding to a standar set of question. The answer to the questions are summarized to obtain a numerical value that represents a cognitive characteristic of the subjects”.⁶¹ Tes adalah alat yang harus diisi oleh responden untuk mengukur kemampuan kognitifnya dengan cara menjawab serangkaian pertanyaan. Jawaban atas pertanyaan merupakan rangkuman capaian nilai dalam bentuk angka yang mencerminkan karakteristik kognitif peserta tes.



Gambar 3. Ilustrasi Tes

Test adalah alat untuk mengukur pengetahuan seseorang mengenai sesuatu. Setelah siswa belajar tentang cara mengoperasikan computer, maka guru akan melakukan tes untuk mengetahui tingkat pengetahuan siswa tentang bagaimana menyalakan computer, menonaktifkan computer, cara mengetik yang benar, cara mengatur margin, dll.

Suharsimi Arikunto menyebutkan bahwa sebuah tes dapat dikatakan baik sebagai alat pengukur, jika memenuhi persyaratan tes yaitu:

1. Validitas, alat ukur atau alat tes dikatakan valid jika instrument yang dikembangkan untuk melakukan tes dapat digunakan untuk mengukur apa yang seharusnya diukur. Sebagai contoh

⁶¹ James H. McMilan. *Educational research fundamentals for the consumer*, (United States: Pearson Education Inc., 2008), p. 160.

alat ukur berat badan tentu timbangan, untuk mengukur suhu tubuh thermometer, untuk mengukur kecepatan speedometer.

2. Reliabilitas, adalah bahwa alat ukur yang akan digunakan harus memiliki sifat keajegan/tetap tidak lentur. Misalnya alat untuk mengukur Panjang adalah meteran tetapi jika meteran tersebut dibuat dari bahan karet yang bersifat elastis maka alat ukur Panjang itu atau meteran yang dimaksud tidak memiliki sifat reliabilitas memadai.
3. Objektifitas, sebuah tes dikatakan objektif jika dalam melaksanakan tes itu tidak ada factor subjektif yang memengaruhi. Artinya pelaksanaan tes berorientasi pada sasaran tes, termasuk proses pembuatan instrumennya pun harus berdasar pada materi yang akan diujikan.
4. Praktikabilitas, tes yang praktis apabila memenuhi syarat; a) mudah dilaksanakan misalnya tidak menuntut peralatan yang banyak dan memberi kebebasan kepada siswa untuk mengerjakan terlebih dahulu bagian yang dianggap mudah, b) mudah pemeriksaannya, bahwa tes itu dilengkapi dengan kunci jawaban maupun pedoman skoringnya, c) dilengkapi dengan petunjuk-petunjuk yang jelas sehingga dapat diberikan/diawali oleh orang lain.
5. Ekonomis berarti bahwa pelaksanaan tes tidak membutuhkan biaya yang mahal, tenaga yang banyak, dan waktu yang lama.⁶²

B. Fungsi Tes

Suharsimi Arikunto (2013) menyebutkan bahwa fungsi tes dapat ditinjau dari tiga hal:

1. Fungsi kelas,
2. Fungsi bimbingan, dan
3. Fungsi untuk administrasi.

Fungsi kelas meliputi;

1. Mengadakan diagnosis terhadap kesulitan belajar siswa.

⁶² Suharsimi Arikunto. Dasar-dasar Evaluasi Pendidikan, (Jakarta: PT Bumi Aksara, 2013), hh. 72-77.

2. Mengevaluasi celah antara bakat dengan pencapaian.
3. Menaikan tingkat prestasi.
4. Mengelompokkan siswa dalam kelas pada waktu metode kelompok.
5. Merencanakan kegiatan proses belajar mengajar untuk siswa secara perseorangan.
6. Menentukan siswa mana yang memerlukan bimbingan khusus.
7. Menentukan tingkat pencapaian untuk setiap anak.

Fungsi Bimbingan meliputi;

1. Menentukan arah pembicaraan dengan orang tua tentang anak-anak mereka.
2. Membantu siswa dalam menentukan pilihan.
3. Membantu siswa mencapai tujuan Pendidikan dan jurusan.
4. Memberi kesempatan kepada pembimbing, guru, dan orang tua dalam memahami kesulitan anak.

Fungsi Administrasi terdiri dari;

1. Memberi petunjuk dalam pengelompokkan siswa.
2. Penempatan siswa baru.
3. Membantu siswa memilih kelompok.
4. Menilai kurikulum.
5. Memperluas hubungan masyarakat.
6. Menyediakan informasi untuk badan-badan lain di luar sekolah.

C. Bentuk Tes

Secara umum dikenal dua bentuk tes yaitu tes standar dan tes buatan guru. Antara Tes standar dengan tes non-standar sebenarnya tidak jauh berbeda karena tes disusun dengan tipe soal yang sama dan meliputi bahan atau pengetahuan yang sama banyak dengan bahan atau pengetahuan yang dicakup oleh tes buatan guru.⁶³ Mertens mengatakan bahwa *“the distinguishing characteristic of standardized tests is that they have uniformity in directions for administering and scoring the instrument, as well as having been through rigorous*

⁶³ Suharsimi Arikunto. Dasar-dasar Evaluasi Pendidikan, (Jakarta: PT Bumi Aksara, 2013), h. 159.

developmental cycles. The nonstandardized tests is usually not developed through a rigorous process and has a limited purpose and application, such as a teacher-made test or that developed by a researcher for a specific study".⁶⁴ Perbedaan karakter antara tes standar adalah bahwa mereka memiliki keseragaman pada petunjuk pengadministrasian dan penskoran instrument, serta dikembangkan dengan pola yang teliti atau dengan cara yang setepat-tepatnya. Sementara tes nonstandard biasanya tidak dikembangkan melalui proses yang teliti dan memiliki tujuan dan penggunaan terbatas, seperti tes yang dibuat oleh guru atau yang dikembangkan oleh peneliti untuk kebutuhan penelitian tertentu.

Perbedaan diantara tes standar dan tes non-standard menurut Suharsimi Arikunto adalah sebagai berikut.

Tabel 7. Tes Standar dan Tes Non-Standard

No	Tes Standar	Tes Buatan
1	Didasarkan atas bahan dan tujuan umum dari sekolah-sekolah diseluruh negara.	Didasarkan atas bahan dan tujuan khusus yang dirumuskan oleh guru untuk kelasnya sendiri.
2	Mencakup aspek yang luas dan pengetahuan atau keterampilan dengan hanya sedikit butir tes untuk setiap keterampilan atau topik.	Dapat terjadi hanya mencakup pengetahuan atau keterampilan yang sempit.
3	Disusun dengan kelengkapan staf professor, pembahas, editor, butir tes.	Biasanya disusun sendiri oleh guru dengan sedikit atau tanpa bantuan orang lain/tenaga ahli.
4	Menggunakan butir-butir tes yang sudah diujicobakan, dianalisis dan direvisi sebelum menjadi sebuah tes.	Jarang-jarang menggunakan butir-butir tes yang sudah diujicobakan, dianalisis, dan direvisi.
5	Mempunyai reliabilitas yang tinggi.	Mempunyai reliabilitas sedang atau rendah.
6	Dimungkinkan menggunakan norma untuk seluruh negara.	Norma kelompok terbatas kelas tertentu

D. Pengukuran

McMilan mengatakan *"Measurement can be defined as the assignment of numbers to indicate different values of a variable; some researchers may also use it to refer to quantitative data collection. Measurement is used to determine how much of a trait, attribute, or characteristic an*

⁶⁴ Donna M. Mertens. *Research and Evaluation in Educational and Psychology: Integrating Diversity with Quantitative, Qualitative, And Mixed Methods*, (California: Sage Publications, Inc. 2005), p. 368.

individual possesses.⁶⁵ Pengukuran dapat didefinisikan sebagai angka-angka yang mengindikasikan perbedaan nilai diantara variable; beberapa peneliti biasanya menggunakan pengukuran untuk mengumpulkan data kuantitatif. Pengukuran digunakan untuk menentukan sifat, jenis atau karakteristik yang dimiliki individu.



Gambar 4: Ilustrasi Pengukuran (*Measurement*)

Skala Pengukuran

“The purpose of measurement is to obtain information about the variables that are being studied. Measurement is a critical component of quantitative research because it provides a systematic procedure for recording observations, performance, or other responses of subjects”.⁶⁶

Tujuan pengukuran adalah memperoleh informasi tentang variable yang sedang diteliti. Pengukuran merupakan komponen penting terkait penelitian kuantitatif karena memberikan prosedur yang sistematis dalam perekaman saat observasi, kinerja, dan respon subjek yang lain.

⁶⁵ James H. McMilan. *Educational research fundamentals for the consumer*, (United States: Pearson Education Inc., 2008), p. 131.

⁶⁶ James H. McMilan. *Educational research fundamentals for the consumer*, (United States: Pearson Education Inc., 2008), p. 1

Dalam hal pengukuran terdapat empat skala yang dapat digunakan terutama dalam kegiatan penelitian. Empat jenis skala pengukuran yang dimaksud adalah nominal, ordinal, interval, dan rasio.

Data Skala nominal

Data nominal adalah “*numbers assigned to category. Example of nominal data in research are gender, race, type of school, and nature of community (e.g., rural, suburban, urban)*”.⁶⁷ Skala nominal adalah angka-angka yang menunjukkan kategorisasi. Contoh data nominal dalam penelitian adalah jenis kelamin, ras, tipe sekolah, dan asal komunitas (kota, pinggiran, dll). Laki-laki dapat di beri kode nominal “1” dan Perempuan “2”.

Data Ordinal

Data ordinal adalah bagian lain dari data kontinum. Selain memiliki atribut, juga memiliki peringkat atau urutan. Ia digunakan untuk mengurutkan objek dari yang paling rendah sampai yang paling tinggi, atau sebaliknya. Skala ordinal merupakan skala pengukuran yang sudah menyatakan peringkat antar tingkatan. Jarak atau interval antar tingkatan juga tidak harus sama. Skala ordinal ini memiliki tingkatan yang lebih tinggi daripada skala nominal, karena skala ini tidak hanya menunjukkan kategori saja tetapi juga menunjukkan peringkat.

Contoh data ordinal adalah tingkat Pendidikan yang disusun sebagai berikut:

- Taman Kanak-kanak (TK)
- Sekolah Dasar (SD)
- Sekolah Menengah Pertama (SMP)
- Sekolah Menengah Atas (SMA)
- Diploma
- Sarjana
- Magister
- Doktor

⁶⁷ *Ibid.* p. 133.

Analisis terhadap urutan data diatas menunjukkan bahwa SD memiliki tingkatan lebih tinggi daripada TK dan lebih rendah dibandingkan SMP.

Data interval

Data interval adalah data hasil pengukuran yang dapat diurutkan atas dasar kriteria tertentu serta menunjukkan semua sifat yang dimiliki oleh data ordinal. Kelebihan sifat data interval dibandingkan dengan data ordinal adalah memiliki sifat kesamaan jarak atau memiliki rentang yang sama antara data yang telah diurutkan. Karena sifat tersebut, terhadap data interval dapat dilakukan operasi matematika penjumlahan dan pengurangan (+,-). Namun demikian masih terdapat satu sifat yang belum dimiliki yaitu tidak adanya angka Nol mutlak pada data interval.

Berikut ciri-cirinya:

1. Kategori data memiliki sifat saling memisah.
2. Kategori data memiliki aturan yang logis.
3. Kategori data ditentukan skalanya berdasarkan jumlah karakteristik khusus yang dimilikinya.
4. Perbedaan karakteristik yang sama tergambar dalam perbedaan yang sama dalam jumlah yang dikenakan pada kategori.
5. Angka nol hanya menggambarkan satu titik dalam skala (tidak memiliki nilai nol absolut)

Adapun contoh data interval adalah sebagai berikut: Rata-rata tinggi badan berdasarkan usia, untuk anak-anak yang berusia 6-12 thn memiliki rata-rata tinggi badan 130-145 cm. Untuk remaja yang berusia 13-18 thn memiliki rata-rata tinggi badan 146-160 cm.

Data Rasio

Data rasio adalah data dengan skala rasio yang lebih menekankan pada serangkaian skala pengukuran variabel yang tidak hanya menghasilkan urutan variabel penelitian, tetapi juga membuat perbedaan antar variabel beserta informasi nilai *true zero* (nol sesungguhnya).

Skala rasio ini setidaknya memberikan informasi yang paling detail karena peneliti dan ahli statistik dan statistika dapat menghitung tendensi sentral menggunakan teknik statistik seperti mean, median, mode, dan metode seperti rata-rata geometrik, koefisien variasi, atau mean harmonic. Contoh paling umum dari data rasio adalah pengukuran ketinggian. Tinggi dapat diukur dalam sentimeter, meter, inci, atau kaki. Tidak mungkin memiliki tinggi negatif. Saat membandingkan dengan data interval, misalnya, suhu bisa -10 derajat Celcius, tetapi ketinggian tidak boleh negatif, seperti yang disebutkan di atas

Contoh pengukuran, dari 100 butir soal yang diajukan dalam tes, Amir menjawab betul sebanyak 80 butir soal. Maka dapat dipahami bahwa pengukuran itu sifatnya kuantitatif. Suatu proses pemberian angka (skor) terhadap proses dan hasil pembelajaran berdasarkan kriteria atau ukuran tertentu yang jelas dan sesuai dengan tujuan yang telah ditentukan dalam rangka memberikan judgement, yakni berupa keputusan terhadap proses dan hasil pembelajaran.

RANGKUMAN

Tes atau testing (*test*) merupakan prosedur sistematis yang direncanakan oleh evaluator guna membandingkan antar perilaku yang dievaluasi. Dalam evaluasi program, tes bisa dilakukan di awal atau tengah kegiatan program atau proyek yang disebut sebagai *pre-test* atau akhir kegiatan program atau proyek yang disebut dengan *post-test*.

Pengukuran (*measurement*) dapat didefinisikan sebagai angka-angka yang mengindikasikan perbedaan nilai diantara variabel. Dalam hal pengukuran terdapat empat skala yang dapat digunakan terutama dalam kegiatan penelitian. Empat jenis skala pengukuran yang dimaksud adalah nominal, ordinal, interval, dan rasio.

Skala nominal adalah angka-angka yang menunjukkan kategorisasi. Contoh data nominal dalam penelitian adalah jenis kelamin, ras, tipe sekolah, dan asal komunitas (kota, pinggiran, dll). Laki-laki dapat di beri kode nominal "1" dan Perempuan "2".

Data ordinal adalah bagian lain dari data kontinum. Selain memiliki atribut, juga memiliki peringkat atau urutan. Ia digunakan untuk mengurutkan objek dari yang paling rendah sampai yang paling tinggi, atau sebaliknya. Skala ordinal merupakan skala pengukuran yang sudah menyatakan peringkat antar tingkatan.

Data interval adalah data hasil pengukuran yang dapat diurutkan atas dasar kriteria tertentu serta menunjukkan semua sifat yang dimiliki oleh data ordinal. Kelebihan sifat data interval dibandingkan dengan data ordinal adalah memiliki sifat kesamaan jarak atau memiliki rentang yang sama antara data yang telah diurutkan.

Data rasio adalah data dengan skala rasio yang lebih menekankan pada serangkaian skala pengukuran variabel yang tidak hanya menghasilkan urutan variabel penelitian, tetapi juga membuat perbedaan antar variabel beserta informasi nilai *true zero* (nol sesungguhnya).

Latihan 7

1. Jelaskan apa yang dimaksud dengan tes, pengukuran, penilaian, dan evaluasi?
2. Sebutkan jenis-jenis penilaian!
3. Jelaskan langkah-langkah dalam penelitian evaluasi?

4. Sebutkan prinsip-prinsip evaluasi!
5. Jelaskan apa yang dimaksud dengan Validitas instrument?

Pembahasan Latihan Soal 7

1. Tes atau testing (test) merupakan prosedur sistematis yang direncanakan oleh evaluator guna membandingkan antar perilaku yang dievaluasi. Pengukuran (measurement) dapat didefinisikan sebagai angka-angka yang mengindikasikan perbedaan nilai diantara variabel. Penilaian (assessment) adalah proses mengumpulkan informasi melalui observasi yang sistematis untuk membuat keputusan mengenai individu. Evaluasi (evaluation) merujuk pada prosedur pengumpulan informasi dan menggunakannya untuk membuat keputusan.
2. Pusat kurikulum Badan Penelitian dan Pengembangan (Balitbang) Kementerian Pendidikan dan Kebudayaan, menggolongkan penilaian dalam berbagai jenis; Kuis, isian atau jawaban singkat, Pertanyaan lisan, Ulangan harian, Ulangan tengah semester, Tugas individu, Tugas kelompok, Ujian praktik, Laporan kerja praktik, dan Penilaian portofolio.
Sementara dilihat dari bentuknya Suharsimi Arikunto menyampaikan bentuk penilaian terdiri dari; Penilaian tes tertulis, Penilaian tes lisan, Penilaian unjuk kerja, Penilaian produk, Penilaian proyek, Penilaian portofolio, dan Penilaian diri.
3. Langkah-langkah dalam penelitian evaluasi adalah sebagai berikut;
 - a. Mengklasifikasi alasan penelitian,
 - b. Memilih model evaluasi,
 - c. Menentukan tujuan evaluasi,
 - d. Mengidentifikasi pemangku kepentingan yang hendak dilibatkan,
 - e. Mengidentifikasi subjek atau objek yang dievaluasi,
 - f. Mengembangkan desain penelitian evaluasi,
 - g. Pengambilan data di lapangan, dan
 - h. Melaporkan hasil evaluasi.
4. Prinsip-prinsip penting dalam evaluasi adalah sebagai berikut;
 - a. Jujur, hal ini terkait dengan para pihak yang terlibat dalam pengumpulan data, informasi dan keterangan-keterangan yang dibutuhkan agar dapat memberikan data yang sesuai fakta tanpa

- ada rekayasa. Begitu juga dengan evaluator yang memiliki tanggungjawab mengolah informasi harus secara objektif dalam proses pengolahan data yang dimaksud.
- b. Prinsip kedua adalah objektif. Bahwa semua pihak yang terlibat dalam evaluasi harus menyangkutkan penilaian atas data dan informasi sesuai apa adanya tanda ada intervensi atau dipengaruhi oleh factor-faktor lain diluar informasi dan kriteria yang ada.
 - c. Tanggungjawab, para pihak yang terlibat memberikan data dan informasi yang benar dan nyata serta bisa diberikan alasannya secara rasional.
 - d. Transparansi, hasil evaluasi dapat dikomunikasikan untuk memperoleh hasil yang lebih baik dan bisa dipertanggungjawabkan.
5. Alat ukur atau alat tes dikatakan valid jika instrument yang dikembangkan untuk melakukan tes dapat digunakan untuk mengukur apa yang seharusnya diukur. Sebagai contoh alat ukur berat badan tentu timbangan, untuk mengukur suhu tubuh thermometer, untuk mengukur kecepatan speedometer.

Test Formatif Pilihan Berganda

1. Bentuk tes yang termasuk kategori objective adalah....!
 - a. Pilihan berganda
 - b. Uraian
 - c. Isian singkat
 - d. Portofolio
2. Objek yang tidak termasuk sebagai masukan instrumental dalam kegiatan evaluasi adalah ...!
 - a. Materi kurikulum
 - b. Sarana Pendidikan
 - c. Pengelolaan
 - d. Lingkungan
3. Dalam kegiatan evaluasi, masukan mentah merupakan salah satu objek evaluasi, sebutkan yang termasuk masukan mentah?
 - a. Guru

- b. Siswa
 - c. Keluarga
 - d. Teman bermain
4. Suatu Tes disebut standar jika memiliki kriteria sebagai berikut!
 - a. Dapat terjadi hanya mencakup pengetahuan atau keterampilan yang sempit.
 - b. Didasarkan atas bahan dan tujuan khusus yang dirumuskan oleh guru untuk kelasnya sendiri.
 - c. Biasanya disusun sendiri oleh guru dengan sedikit atau tanpa bantuan orang lain/tenaga ahli.
 - d. Menggunakan butir-butir tes yang sudah diujicobakan, dianalisis dan direvisi sebelum menjadi sebuah tes.
 5. Sebuah tes dapat dikatakan baik sebagai alat pengukur, jika memenuhi syarat sebagai berikut, kecuali...!
 - a. Produktivitas
 - b. Praktikabilitas
 - c. Objektivitas
 - d. Reliabilitas

Kunci jawaban dan pedoman penskoran

1. A
2. D
3. B
4. D
5. A

Bobot

Latihan soal essay 75 % dengan bobot per soal 15.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{50}{75} \times 100 = \mathbf{66,67}$$

Tes Formatif Pilihan Berganda 25% dengan bobot per soal 5.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{20}{25} \times 100 = \mathbf{80}$$

$$66,67 \times 75\% = \mathbf{50}$$

$$80 \times 25\% = \mathbf{20}$$

$$50 + 20 = \mathbf{70}$$

$$\text{Nilai akhir} = \mathbf{70}$$

MODUL 8

PENGERTIAN KONSEP PENILAIAN DAN EVALUASI DALAM PEMBELAJARAN

Metode Pembelajaran	Estimasi waktu	Capaian Pembelajaran
-Presentasi virtual -Diskusi Virtual interaktif -Tanya jawab	120 menit X 1	- Mahasiswa memahami konsep penilaian dan evaluasi dalam pembelajaran

Materi 8. Pengertian konsep penilaian dan evaluasi dalam pembelajaran.

A. Penilaian

Penilaian menjadi bagian yang tidak terpisahkan dalam kegiatan evaluasi secara umum. Kegiatan penilaian melekat pada proses pembelajaran, karena penilaian merupakan salah satu alat ukur yang paling relevan (setidaknya hingga saat ini) untuk mengetahui perkembangan siswa dalam kegiatan pembelajaran. Penilaian tidak hanya dapat mengetahui prestasi siswa tetapi juga kelemahan siswa bahkan dapat menjadi indikator keberhasilan guru dalam mengajar.

Dilihat dari jenisnya pusat kurikulum Badan Penelitian dan Pengembangan (Balitbang) Kementerian Pendidikan dan Kebudayaan, penilaian terdiri atas;

1. Kuis, isian atau jawaban singkat,
2. Pertanyaan lisan,
3. Ulangan harian,
4. Ulangan tengah semester,
5. Tugas individu,
6. Tugas kelompok,
7. Ujian praktik,

8. Laporan kerja praktik, dan
9. Penilaian portofolio.

Sementara dilihat dari bentuknya Suharsimi Arikunto menyampaikan ada tujuh bentuk penilaian.

1. Penilaian tes tertulis,
2. Penilaian tes lisan,
3. Penilaian unjuk kerja,
4. Penilaian produk,
5. Penilaian proyek,
6. Penilaian portofolio, dan
7. Penilaian diri

Merujuk pada pernyataan Mertens *“performance assessment is a process for collecting information through systematic observation in order to make decisions about an individual. Performance assessment relies on the use of multiple types of assessments, not a single test or measurement device, and assessment occurs across time. The primary vehicle for assessment is the direct observation of performance in the form of behavior and products”*.⁶⁸ Penilaian kinerja adalah proses mengumpulkan informasi melalui observasi yang sistematis untuk membuat keputusan mengenai individu. Penilaian kinerja menyandarkan pada penggunaan berbagai tipe penilaian, bukan hanya satu bentuk tes atau satu alat pengukuran, dan penilaian dilakukan sepanjang waktu. Alat utama untuk melakukan penilaian kinerja adalah observasi langsung terhadap kinerja yang dapat dilihat melalui perilaku dan produk. Dapat dipahami bahwa penilaian kinerja ini bersifat komprehensif, karena melihat kinerja individu dengan menggunakan berbagai media atau alat untuk menentukan seberapa tinggi unjuk kerja seseorang yang dinilai.

Merten juga mengutip pernyataan Fuch & Deno, 1991, *“performance assessment or curriculum-based assessment is typically designed to collect information on the instructional needs and/or progress of students*

⁶⁸ Donna M. Mertens. *Research and Evaluation in Educational and Psychology: Integrating Diversity with Quantitative, Qualitative, and Mixed Methods*, (California: Sage Publications, Inc., 2005), p. 369.

through the continuous, direct observation of student performance”.⁶⁹ Penilaian kinerja atau penilaian berbasis kurikulum secara khas di desain untuk mengumpulkan informasi mengenai kebutuhan instruksional dan/atau perkembangan siswa secara berkelanjutan melalui observasi langsung terhadap kinerja siswa. Kata kuncinya adalah kebutuhan instruksional, bahwa dalam penilaian kinerja akan dilihat secara langsung hal-hal yang dibutuhkan untuk meningkatkan proses pembelajaran sehingga hasil dari pembelajaran dapat meningkat secara optimal sesuai yang telah ditetapkan sebelumnya. Karena didalam instruksional terdapat banyak hal yang saling bersinggungan berperan baik langsung maupun tidak langsung terhadap keberhasilan prose pembelajaran yang dilakukan guru. Dalam instruksional ada kurikulum yang menyangkut metode pembelajaran, buku yang digunakan, Teknik dan metode penilaian siswa, sarana, kompetensi guru, dan lain sebagainya.



Gambar 5. Ilustrasi Penilaian (Assessment)

Menurut McMilan penilaian adalah *“The shorter term, assess, is a synonyme for measure. In the context of classroom assessment, the term refers to entire process of measurement, evaluation, and finally, use of the information by teachers and students.”*⁷⁰ Istilah pendeknya adalah

⁶⁹ *Ibid.*, p. 369.

⁷⁰ James H. McMilan. *Educational research fundamentals for the consumer*, (United States: Pearson Education Inc., 2008), p. 132.

“menilai” merupakan sinonim dengan mengukur. Dalam konteks penilaian kelas, istilah tersebut merujuk pada keseluruhan proses pengukuran, evaluasi, dan akhirnya, menggunakan informasi yang terkumpul oleh guru dan siswa. *“Sometimes assessment means ‘evaluation,’ and sometimes it refers to the more specific process of diagnosing of individual difficulties, such as assessing for learning disabilities.”*⁷¹ Terkadang asesmen berarti evaluasi dan kadang-kadang merujuk pada proses khusus untuk mendiagnosis kesulitan individu, seperti menilai kesulitan belajar. Dalam permendikbud 2014, penilaian adalah proses mengumpulkan informasi/bukti melalui pengukuran, menafsirkan, mendeskripsikan, dan menginterpretasikan bukti-bukti hasil pengukuran. Sebagai contoh, dari 100 butir soal yang ada, Aji dapat menjawab dengan benar sebanyak 80 butir soal. Maka, dapat ditentukan bahwa Aji termasuk anak pandai. Dalam hal ini penilaian bersifat kualitatif. Penilaian bisa diartikan sebagai proses interpretasi data hasil pengukuran terhadap proses dan hasil pembelajaran yang awalnya berupa skor atau angka menjadi nilai berdasarkan prosedur tertentu.

B. Evaluasi Dalam Pembelajaran

*“Evaluation is a term that is sometime used interchangeably with measurement. Evaluation refers to procedures for collecting information and using the information to make decisions. Principles of measurement and specific instruments are part of evaluation.”*⁷² Evaluasi adalah sebuah istilah yang kadang-kadang digunakan secara bergantian dengan pengukuran. Evaluasi merujuk pada prosedur pengumpulan informasi dan menggunakannya untuk membuat keputusan. Prinsip pengukuran dan instrumennya merupakan bagian dari evaluasi. Dalam pandangan ini evaluasi disamakan dengan pengukuran dimana kedua-duanya merupakan proses mengumpulkan data dan informasi untuk kepentingan membuat keputusan terkait objek yang diukur.

⁷¹ *Ibid.* P. 132.

⁷² James H. McMilan. *Educational research fundamentals for the consumer*, (United States: Pearson Education Inc., 2008), p. 131.

Scriven menyebutkan “*What is evaluation? Synthesizing what the dictionaries and common usage tell us, it is the process of determining the merit, worth, or significance of things (near-synonyms are quality/value/importance).*”⁷³ Apa itu evaluasi? Sintesis apa yang kamus dan masyarakat sampaikan kepada kita, yaitu proses menentukan nilai, manfaat atau derajat signifikansi tentang sesuatu (mendekati sinonim mutu/nilai/hal yang penting). Kata kunci yang dapat diambil dari definisi ini adalah proses menentukan nilai, manfaat dan derajat signifikansi barang atau sesuatu. Tentu dibutuhkan langkah-langkah yang tepat dan sesuai tahapan dalam mendapatkan pengetahuan secara ilmiah. Hasil penelitian harus memiliki nilai tambah baik dari aspek keilmuan maupun dari sisi manfaat bagi objek penelitian dan juga subjek penelitian. Terutama bagi pengguna manfaat penelitian tersebut (para *stakeholder*).

Stufflebeam mendefinisikan evaluasi “*Evaluation is the process of delineating, obtaining, providing, and applying descriptive and judgmental information about the merit and worth of some object's goals, design, implementation, and outcomes to guide improvement decisions, provide accountability reports, inform institutionalization/ dissemination decisions, and improve understanding of the involved phenomena.*”⁷⁴ Evaluasi adalah proses mendeskripsikan, mendapatkan, menyediakan, dan menerapkan dan menyatakan pendapat tentang informasi mengenai nilai dan manfaat tujuan sesuatu, design, implementasi, dan hasil untuk membimbing kepada perbaikan keputusan, menyediakan laporan pertanggungjawaban, menyampaikan keputusan institusi, dan memperbaiki pemahaman fenomena yang meliputinya.

Siapa yang melakukan evaluasi adalah siapa saja yang memiliki kepentingan untuk memperoleh informasi terkait objek yang dimaksud. Evaluator dalam kegiatan evaluasi disebut sebagai subjek, sementara sasaran kegiatan evaluasi disebut sebagai objek evaluasi. Objek evaluasi bisa berupa apa saja yang dapat diamati atau hal-hal yang

⁷³ Michael Scriven, *International Handbook of Educational Evaluation*, T. Kellaghan, D.L. Stufflebeam (eds.), (Dordrecht: Kluwer Academic Publishers, 2003), P. 15.

⁷⁴ Daniel L. Stufflebeam. *International Handbook of Educational Evaluation*, T. Kellaghan & D.L. Stufflebeam (eds.), (Dordrecht: Kluwer Academic Publishers, 2003), P. 15.Ibid. p. 34.

menjadi perhatian untuk dievaluasi. Menurut Suharsimi Arikunto objek evaluasi adalah aspek-aspek yang berkaitan dengan transformasi diantaranya; (1) masukan mentah, (2) masukan instrumental, (3) masukan lingkungan, (4) proses transformasi, dan (5) keluaran, yaitu hasil dari transformasi.

Siswa yang dalam proses pembelajaran berstatus sebagai subjek, dalam evaluasi dia merupakan objek evaluasi, karena dicermati untuk diketahui kinerjanya ketika mengikuti pembelajaran.⁷⁵ Dalam pembahasan ini maka siswa sebagai masukan mentah yang perlu diolah oleh guru di dalam kelas maupun diluar kelas untuk diberikan pemahaman dan Pendidikan sehingga secara mental dan pengetahuan serta keterampilan menjadi berkembang. Kedua, masukan instrumental. Yang termasuk masukan instrumental adalah guru, materi kurikulum, sarana Pendidikan, dan pengelolaan, maka dalam penilaian perlu dicermati kinerja masing-masing unsur tersebut.⁷⁶

Ketiga, adalah masukan lingkungan. Diantaranya adalah lingkungan keluarga, meskipun secara lokasi lingkungan keluarga jaraknya jauh dengan sekolah tempat siswa belajar, tetapi lingkungan keluarga terutama dukungan orang tua dalam menyediakan sarana penunjang untuk belajar siswa selama dirumah sangat berperan dalam menentukan keberhasilan peserta didik di sekolah. Selain lingkungan keluarga ada lingkungan sekolah yang tentu saja berperan dominan terhadap kesuksesan belajar siswa. Serta yang terakhir adalah lingkungan bermain dan bergaul di masyarakat.

Proses transformasi merupakan serangkaian kegiatan *treatment* atau perlakuan yang dilakukan guru terhadap siswa dengan tujuan untuk meningkatkan pengetahuan, pemahaman dari tidak tahu menjadi tahu, dari tidak bisa menjadi bisa dan mengubah perilaku dari kurang baik menjadi baik begitu seterusnya. Pada tahap akhir adalah keluaran, hal ini berkaitan dengan hasil yang diperoleh atas proses yang telah dilakukan.

⁷⁵ Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, (Jakarta: PT Bumi Aksara, 2013), h. 31.

⁷⁶ Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, (Jakarta: PT Bumi Aksara, 2013), h. 10.

Menurut Sukardi beberpa prinsip-prinsip penting dalam evaluasi adalah sebagai berikut.

1. Jujur, hal ini terkait dengan para pihak yang terlibat dalam pengumpulan data, informasi dan keterangan-keterangan yang dibutuhkan agar dapat memberikan data yang sesuai fakta tanpa ada rekayasa. Begitu juga dengan evaluator yang memiliki tanggungjawab mengolah informasi harus secara objektif dalam proses pengolahan data yang dimaksud.
2. Prinsip kedua adalah objektif. Bahwa semua pihak yang terlibat dalam evaluasi harus menyandarkan penilaian atas data dan informasi sesuai apa adanya tanda ada intervensi atau dipengaruhi oleh factor-faktor lain diluar informasi dan krteria yang ada.
3. Tanggungjawab, para pihak yang terlibat memberikan data dan informasi yang benar dan nyata serta bisa diberikan alasannya secara rasional.
4. Transparansi, hasil evaluasi dapat dikomunikasikan untuk memperoleh hasil yang lebih baik dan bisa dipertanggunggugatkan.

Sukardi juga memberi penjelasan langkah-langkah yang ada dalam penelitian evaluasi sebagai berikut.

1. Mengklasifikasi alasan penelitian,
2. Memilih model evaluasi,
3. Menentukan tujuan evaluasi,
4. Mengidentifikasi pemangku kepentingan yang hendak dilibatkan,
5. Mengidentifikasi subjek atau objek yang dievaluasi,
6. Mengembangkan desain penelitian evaluasi,
7. Pengambilan data di lapangan, dan
8. Melaporkan hasil evaluasi.

Mertens (2005) memberikan penjelasan mengenai langkah-langkah dalam perencanaan evaluasi sebagai berikut.

Focusing the evaluation

- *Description of what is to be evaluated*
- *The purpose of the evaluation*
- *The stakeholders in the evaluation*
- *Constraints affecting the evaluation*
- *The evaluation questions*
- *Selection of an evaluation model*

Planning the evaluation

- *Data collection specification, analysis, interpretation, and use strategies*
- *Management of evaluation*
- *Meta-evaluation plans*

Implementing the evaluation

- *Completing the scope of work specified in the plan*

Pada tahap awal perlu dibahas mengenai apa yang akan dievaluasi. Hal ini penting supaya jelas dan tidak melebar kearah yang terlalu luas sehingga membingungkan dan tidak focus dengan objek evaluasi. Jika demikian, maka dapat dipastikan hasil evaluasi tidak akan akurat dan bias, pada akhirnya tidak bisa digunakan sebagai bahan pengambilan keputusan. Deskripsi objek evaluasi sesungguhnya ingin menjawab hal-hal yang menurut Mertens sebagai berikut:

1. *Is there a written description of what is to be evaluated?*
2. *What is the status of the evaluated? Relatively stable and mature? New? Developing? How long has the program been around?*
3. *In what context will (or does) the evaluand function?*
4. *Who is the evaluand designed to serve?*
5. *How does the evaluand work? Or how is it supposed to work?*
6. *What is it supposed to do?*
7. *What resources are being put into the evaluand (e.g., financial, time, staff, materials, etc.)?*
8. *What are the processes that make up the evaluand?*
9. *What outputs are expected? Or occur?*
10. *Why do you want to evaluate it?*

11. *Whos description of the evaluand is available to you at the start of the evaluation?*
12. *Whos description of the evaluand is needed to get a full understanding of the program to be evaluated?*

Pertanyaan-pertanyaan diatas ingin menjelaskan tentang apa yang akan dievaluasi, tujuan evaluasi, para stakeholder evaluasi yang akan dilaksanakan dan hal-hal yang kemungkinan akan menjadi penghambat dalam evaluasi. Berkaitan dengan pertanyaan evaluasi, merupakan pengembangan dari tujuan evaluasi yang menurut Mertens *“evaluation question can be derived from the statement of purpose for the evaluation, expanded on by holding brainstorming sessions with stakeholder groups, borrowed from previous evaluation studies, or generated by a theoretical framework that is relevant to the study”*.⁷⁷ Pertanyaan evaluasi dapat di rangkum dari pernyataan tujuan evaluasi, dikembangkan dengan brainstorming bersama dengan para *stakeholder*, mengambil dari penelitian terdahulu, atau dikembangkan berdasarkan kerangka teori yang relevan dengan penelitian dimaksud.

Berkenaan dengan pemilihan model evaluasi, peneliti perlu memilihnya karena akan menentukan arah evaluasi apakah sesuai dengan hasil yang diinginkan oleh stakeholder terkait. Maka menjadi penting bagi evaluator untuk memilih dan menetapkan model apa yang akan digunakan dalam evaluasi tersebut. Apakah menggunakan CIPP, model ketimpangan, model sistem evaluasi analisis dan lain sebagainya.

Pada tahap perencanaan evaluasi, evaluator akan melakukan proses pengumpulan data, analisis data, interpretasi data, dan menetapkan strateginya. Dalam pengumpulan data hal-hal yang perlu dilakukan adalah tentang bagaimana cara mengumpulkan data dimaksud, siapa yang akan melakukan, siapa atau apa yang menjadi sumber data/informasi, kapan waktunya dan dimana tempat pengumpulan data, dan bagaimana data-data dikumpulkan ke peneliti (jika

⁷⁷ Donna M. Mertens. *Research and Evaluation in Educational and Psychology: Integrating Diversity with Quantitative, Quaalitative, and Mixed Methods*, (California: Sage Publications, Inc., 2005), p. 69.

menggunakan angket misalnya, bagaimana hasil isian responden dikembalikan ke peneliti).

Analisis dan interpretasi data merupakan bagian penting yang akan dilakukan evaluator setelah data dan informasi terkumpul. Tahap ini membutuhkan Teknik dan juga waktu yang cukup untuk mengolah data sampai akhirnya dapat digunakan untuk mengambil keputusan terkait objek evaluasi.

Bagaimanapun bagusya perencanaan suatu evaluasi yang dilakukan, tetap saja memiliki potensi penyimpangan atau tidak sesuaiya antara hasil yang diharapkan dengan fakta yang terjadi. Maka, diperlukan disiapkan meta-evaluasi sebagai langkah antisipasi jika terjadi penyimpangan dalam proses evaluasi. Meta evaluasi sangat penting dilakukan oleh evaluator untuk menjaga semua rencana kegiatan dalam evaluasi berjalan pada jalur yang sudah direncanakan. Hasil evaluasi juga perlu dipantau agar sesuai dengan target yang sudah ditetapkan sebelumnya.

RANGKUMAN

Penilaian (*assessment*) adalah proses mengumpulkan informasi melalui observasi yang sistematis untuk membuat keputusan mengenai individu.

Evaluasi (*evaluation*) merujuk pada prosedur pengumpulan informasi dan menggunakannya untuk membuat keputusan. Prinsip-prinsip penting dalam evaluasi adalah jujur, objektif, tanggungjawab dan transparan.

Evaluator dalam kegiatan evaluasi disebut sebagai subjek, sementara sasaran kegiatan evaluasi disebut sebagai objek evaluasi. Objek evaluasi bisa berupa apa saja yang dapat diamati atau hal-hal yang menjadi perhatian untuk dievaluasi. Menurut Suharsimi Arikunto objek evaluasi adalah aspek-aspek yang berkaitan dengan transformasi diantaranya; (1) masukan mentah, (2) masukan instrumental, (3) masukan lingkungan, (4) proses transformasi, dan (5) keluaran, yaitu hasil dari transformasi.

Siswa yang dalam proses pembelajaran berstatus sebagai subjek, dalam evaluasi dia merupakan objek evaluasi, karena dicermati untuk diketahui kinerjanya ketika mengikuti pembelajaran.¹ Dalam pembahasan ini maka siswa sebagai masukan mentah yang perlu diolah oleh guru di dalam kelas maupun diluar kelas untuk diberikan pemahaman dan Pendidikan sehingga secara mental dan pengetahuan serta keterampilan menjadi berkembang.

Latihan 8

1. Mengapa proses pembelajaran perlu di evaluasi?
2. Sebutkan jenis-jenis penilaian!
3. Jelaskan langkah-langkah dalam penelitian evaluasi?
4. Sebutkan prinsip-prinsip evaluasi!

5. Jelaskan apa yang dimaksud dengan Validitas instrument?

Pembahasan Latihan Soal 8

1. Kegiatan penilaian melekat pada proses pembelajaran, karena penilaian merupakan salah satu alat ukur yang paling relevan (setidaknya hingga saat ini) untuk mengetahui perkembangan siswa dalam kegiatan pembelajaran. Penilaian tidak hanya dapat mengetahui prestasi siswa tetapi juga kelemahan siswa bahkan dapat menjadi indikator keberhasilan guru dalam mengajar.
2. Pusat kurikulum Badan Penelitian dan Pengembangan (Balitbang) Kementerian Pendidikan dan Kebudayaan, menggolongkan penilaian dalam berbagai jenis; Kuis, isian atau jawaban singkat, Pertanyaan lisan, Ulangan harian, Ulangan tengah semester, Tugas individu, Tugas kelompok, Ujian praktik, Laporan kerja praktik, dan Penilaian portofolio.

Sementara dilihat dari bentuknya Suharsimi Arikunto menyampaikan bentuk penilaian terdiri dari; Penilaian tes tertulis, Penilaian tes lisan, Penilaian unjuk kerja, Penilaian produk, Penilaian proyek, Penilaian portofolio, dan Penilaian diri.

3. Langkah-langkah dalam penelitian evaluasi adalah sebagai berikut;
 - a. Mengklasifikasi alasan penelitian,
 - b. Memilih model evaluasi,
 - c. Menentukan tujuan evaluasi,
 - d. Mengidentifikasi pemangku kepentingan yang hendak dilibatkan,
 - e. Mengidentifikasi subjek atau objek yang dievaluasi,
 - f. Mengembangkan desain penelitian evaluasi,
 - g. Pengambilan data di lapangan, dan
 - h. Melaporkan hasil evaluasi.
4. Prinsip-prinsip penting dalam evaluasi adalah sebagai berikut;
 - a. Jujur, hal ini terkait dengan para pihak yang terlibat dalam pengumpulan data, informasi dan keterangan-keterangan yang dibutuhkan agar dapat memberikan data yang sesuai fakta tanpa ada rekayasa. Begitu juga dengan evaluator yang memiliki tanggungjawab mengolah informasi harus secara objektif dalam proses pengolahan data yang dimaksud.

- b. Prinsip kedua adalah objektif. Bahwa semua pihak yang terlibat dalam evaluasi harus menyandarkan penilaian atas data dan informasi sesuai apa adanya tanpa ada intervensi atau dipengaruhi oleh factor-faktor lain diluar informasi dan kriteria yang ada.
 - c. Tanggungjawab, para pihak yang terlibat memberikan data dan informasi yang benar dan nyata serta bisa diberikan alasannya secara rasional.
 - d. Transparansi, hasil evaluasi dapat dikomunikasikan untuk memperoleh hasil yang lebih baik dan bisa dipertanggunggugatkan.
5. Alat ukur atau alat tes dikatakan valid jika instrument yang dikembangkan untuk melakukan tes dapat digunakan untuk mengukur apa yang seharusnya diukur. Sebagai contoh alat ukur berat badan tentu timbangan, untuk mengukur suhu tubuh thermometer, untuk mengukur kecepatan speedometer.

Test Formatif Pilihan Berganda

1. Bentuk penilaian berikut ini merupakan penilaian yang direkomendasikan oleh Suharsimi Arikunto, kecuali....
 - a. Penilaian melekat
 - b. Penilaian unjuk kerja
 - c. Penilaian produk
 - d. Penilaian diri
2. Objek yang tidak termasuk sebagai masukan instrumental dalam kegiatan evaluasi adalah ...!
 - a. Materi kurikulum
 - b. Sarana Pendidikan
 - c. Pengelolaan
 - d. Lingkungan
3. Dalam kegiatan evaluasi, masukan mentah merupakan salah satu objek evaluasi, sebutkan yang termasuk masukan mentah?
 - a. Guru
 - b. Siswa
 - c. Keluarga

- d. Teman bermain
4. Suatu Tes disebut standar jika memiliki kriteria sebagai berikut!
 - a. Dapat terjadi hanya mencakup pengetahuan atau keterampilan yang sempit.
 - b. Didasarkan atas bahan dan tujuan khusus yang dirumuskan oleh guru untuk kelasnya sendiri.
 - c. Biasanya disusun sendiri oleh guru dengan sedikit atau tanpa bantuan orang lain/tenaga ahli.
 - d. Menggunakan butir-butir tes yang sudah diujicobakan, dianalisis dan direvisi sebelum menjadi sebuah tes.
5. Sebuah tes dapat dikatakan baik sebagai alat pengukur, jika memenuhi syarat sebagai berikut, kecuali...!
 - a. Produktivitas
 - b. Praktikabilitas
 - c. Objektivitas
 - d. Reliabilitas

Kunci jawaban dan pedoman penskoran

1. A
2. D
3. B
4. D
5. A

Bobot

Latihan soal essay 75 % dengan bobot per soal 15.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{50}{75} \times 100 = \mathbf{66,67}$$

Tes Formatif Pilihan Berganda 25% dengan bobot per soal 5.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{20}{25} \times 100 = \mathbf{80}$$

$$66,67 \times 75\% = \mathbf{50}$$

$$80 \times 25\% = \mathbf{20}$$

$$50 + 20 = \mathbf{70}$$

$$\text{Nilai akhir} = \mathbf{70}$$

MODUL 9

MODEL-MODEL EVALUASI GOAL-BASED EVALUATION, GOAL FREE EVALUATION, FORMATIF & SUMATIF EVALUATION, DAN RESPONSIVE EVALUATION

Metode Pembelajaran	Estimasi waktu	Capaian Pembelajaran
-Presentasi virtual -Diskusi Virtual interaktif -Tanya jawab	120 Menit X 1	-Mahasiswa mampu memahami model-model evaluasi <i>Goal-Based Evaluation, Goal Free Evaluation, Formatif & Sumatif Evaluation, dan Responsive Evaluation.</i>

Materi 9. Model-model Evaluasi: *Goal-Based Evaluation, Goal Free Evaluation, Formatif & Sumatif Evaluation, dan Responsive Evaluation.*

A. Model-Model Evaluasi

Model evaluasi merupakan penjabaran teori evaluasi dalam praktik melaksanakan evaluasi. Model evaluasi membedakan antara evaluasi dengan penelitian murni dan penelitian terapan lainnya. Hanya evaluasi yang mempergunakan model evaluasi dalam melaksanakan penelitian.⁷⁸ Model menjadi ciri khas bagi evaluasi karena tidak ada didalam penelitian murni. Selain model evaluasi, istilah lain yang sama maknanya adalah pendekatan evaluasi. Madaus dan Kellaghan menyatakan,

Each evaluation approach has its particular strengths that can help illuminate different aspects of a program. Within the limitations of the budget, pick and choose features from various models that can provide the best evidence to answer questions about the project. For example, consider combining test data from a goals-oriented approach, resource allocation data from the

⁷⁸ Wirawan. Evaluasi: Teori, Model, Metodologi, Standar, Aplikasi dan Profesi, (Jakarta: PT Rajagrafindo Persada, 2016), h. 122.

decision-oriented approach, and observational and interview data from the naturalistic approach.”⁷⁹

Setiap pendekatan evaluasi memiliki kekuatan masing-masing yang dapat membantu menjelaskan berbagai aspek program yang berbeda. Dalam keterbatasan anggaran, ambil dan pilih ciri-ciri dari berbagai model yang bisa memberikan data terbaik untuk menjawab pertanyaan-pertanyaan tentang suatu proyek. Sebagai contoh, mempertimbangkan kombinasi data tes dari pendekatan berorientasi tujuan, data alokasi sumber daya dari pendekatan berbasis tujuan, dan observasi dan data interview dari pendekatan naturalistik.

The research worker, teacher, or school administrator confronted with the problem of appraisal is likely to think of one of these devices and fail to recognize that the choice of the particular method of evaluation should depend upon the effectiveness of that method for the particular problem under consideration. A clearer recognition of the common elements in all of these methods of evaluation is needed to provide a basis for choosing the techniques to use in a particular situation.⁸⁰

Peneliti, guru, atau tenaga administrasi sekolah menghadapi masalah penilaian yang mungkin saja salah satu dari alat berikut dan gagal untuk menetapkan pilihan terhadap metode evaluasi tertentu yang seharusnya didasarkan pada efektifitas metode tersebut untuk masalah tertentu yang akan dievaluasi. Penentuan elemen-elemen umum yang tegas dari metode evaluasi sangat diperlukan sebagai dasar pemilihan teknik yang akan digunakan dalam kondisi tertentu. Stufflebeam menekankan pentingnya menentukan metode yang akan digunakan dalam melakukan evaluasi karena dengan menetapkan metode yang jelas, berimbang pada pemilihan Teknik seperti apa yang akan digunakan sehingga tujuan evaluasi dapat dicapai sesuai target.

⁷⁹ George F. Madaus and Thomas Kellaghan. *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, Edited by Thomas Kellaghan, George F. Madaus, and Daniel L. Stufflebeam (New York: Kluwer Academic Publishers, 2002), P.25

⁸⁰ Educational Evaluation: Classic Works of Ralph W. Tyler Compiled and Edited by George F. Madaus and Daniel L. Stufflebeam, (Boston: Kluwer Academic Publishers, 1989) p.77.

B. Evaluasi Berbasis Tujuan (*Goal-Based Evaluation*)

Model evaluasi berbasis tujuan secara umum ingin mengukur apakah tujuan program atau proyek yang telah ditetapkan, dapat dicapai atau tidak. Model evaluasi ini memfokuskan pada mengumpulkan informasi yang bertujuan mengukur pencapaian tujuan kebijakan, program dan proyek untuk pertanggungjawaban dan pengambilan keputusan. Jika program tidak mempunyai tujuan, atau tidak mempunyai tujuan yang bernilai maka program tersebut merupakan program yang buruk.⁸¹ Suatu kebijakan, program, dan proyek harus memiliki tujuan yang jelas sehingga perkembangan implementasinya dapat diukur sudah sampai tingkat mana. Hal itu pula yang membedakan antara kebijakan, program, dan proyek legal yang memiliki standar dan dikembangkan secara resmi oleh Lembaga resmi dengan kegiatan informal yang sifatnya *coincidence* atau sekedar *trial and error*. Salah satu tokoh yang menjadi pelopor model evaluasi berbasis tujuan adalah Ralph W. Tyler.

	<p>BASIC PRINCIPLES of CURRICULUM and INSTRUCTION</p> <p>Ralph W. Tyler</p> <p>with a Foreword by Peter S. Hiebowitz</p>	<p>RALPH WINFRED TYLER</p> <p>Director Emeritus Center for Advanced Study in the Behavioral Sciences Born, Chicago, April 22, 1902 Education Doane College, A.B., 1921 University of Nebraska, A.M., 1923 University of Chicago, Ph.D., 1927</p>
<p>Tyler was born on April 22, 1902, in Chicago to a professional family. His father, William Augustus Tyler, had been raised in a farm, and had become a doctor. As the sixth of eight children, Tyler grew up in Nebraska where he recalled having to trap animals for food and wear donated clothing. He worked at various jobs while growing up, including his first job at age twelve in a creamery.</p> <p>His first teaching job was as a high school science teacher in Pierre, South Dakota. In 1923, Tyler wrote a science test for high school students which helped him "see the holes in testing only for memorization." He earned his master's degree from the University of Nebraska in 1923 and his Ph.D. from the University of Chicago in 1927.</p> <p>His graduate work at the University of Chicago connected him with notable educators Charles Judd and W. W. Charters, whose ideas influenced Tyler's later work in curriculum development and evaluation. Tyler's first appointment was at the University of North Carolina in 1927, where he worked with state teachers to improve curricula. Later in 1927, Tyler joined the faculty at Ohio State University, where he refined his innovative approach to testing while working with Charters, who was the director of the university's Bureau of Educational Research. Tyler helped Ohio State University faculty to improve their teaching and increase student retention.</p>		

Profil 3: Ralph Winfred Tyler. Sumber: https://en.wikipedia.org/wiki/Ralph_W._Tyler

⁸¹ Wirawan. Evaluasi: Teori, Model, Metodologi, Standar, Aplikasi dan Profesi, (Jakarta: PT Rajagrafindo Persada, 2016), h. 122.

Dalam implementasinya, evaluasi berbasis tujuan ini dimaksudkan untuk mengetahui apakah target program telah dicapai. Pihak-pihak yang memiliki kepentingan terkait evaluasi ini diantaranya adalah pengembang program, sponsor, manajer yang ingin tahu perkembangan yang telah dicapai tentang sasaran program yang telah ditetapkan sebelumnya.

Metode yang digunakan dalam evaluasi ini pada intinya adalah melibatkan tujuan operasional yang spesifik dan mengumpulkan dan menganalisa informasi terkait untuk mengetahui seberapa baik setiap tujuan yang telah ditetapkan dicapai. Tujuan umum dan penilaian kinerjanya juga dapat digunakan. Pengujian berbasis kriteria sangat relevan dengan pendekatan evaluasi model ini. Kriteria adalah standar yang menjadi ukuran ketercapaian implementasi program. Sudah sejauh mana program yang dilaksanakan mencapai tujuan yang ditetapkan sebelumnya, maka kriteria inilah yang menjadi alat ukur atau standarnya.

Madaus dan Kellaghan *“Goal-Achievement Model (of evaluation). The idea that merit of the program (or person) is to be equated with success in achieving a stated goal. Goal-Based Evaluation (GBE) is based and focused on knowledge of the goals and objectives of the program, person or product, (Scriven, 1980a, p. 59).⁸²* Model evaluasi berbasis tujuan. Idenya bahwa nilai suatu program atau seseorang adalah sama dengan keberhasilan dalam mencapai tujuan yang telah ditetapkan. Evaluasi berbasis tujuan didasarkan dan difokuskan pada pengetahuan tentang tujuan atau sasaran program, orang atau produk.

Stufflebeam menegaskan bahwa *“In this approach, some statement of objectives provides the advance organizer. The objectives may be mandated by the client, formulated by the evaluator, or specified by the service providers. The usual purpose of an objectives-based study is to determine whether the program’s objectives have been achieved. Typical audiences are program developers, sponsors, and managers who want*

⁸² George F. Madaus and Thomas Kellaghan. *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, Edited by Thomas Kellaghan, George F. Madaus, Daniel L. Stufflebeam (New York: Kluwer Academic Publishers, 2002), P.26.

to know the extent to which each stated objective was achieved."⁸³ Dalam pendekatan ini (*Goal-based evaluation*) pernyataan tujuan memberikan keluasan bagi organisator. Tujuan merupakan mandate dari klien, dirumuskan oleh evaluator, atau dispesifikasikan oleh penyedia jasa. Tujuan umum suatu penelitian berbasis tujuan adalah menentukan apakah tujuan program telah dicapai. Pihak-pihak yang berkepentingan dalam evaluasi adalah pengembang program, sponsor, dan manajer yang ingin tahu capaian perkembangan dari setiap pernyataan tujuan.

*The objectives-based approach is especially applicable in assessing tightly focused projects that have clear, supportable objectives. Even then, such studies can be strengthened by judging project objectives against the intended beneficiaries' assessed needs, searching for side effects, and studying the process as well as the outcomes.*⁸⁴ Pendekatan berbasis tujuan merupakan suatu hal khususnya dalam kegiatan menilai secara ketat focus pada proyek yang jelas, dan tujuannya dapat dilaksanakan. Evaluasi berbasis tujuan dapat diperperkuat dengan membandingkan tujuan proyek terhadap kebutuhan penilaian yang diinginkan.

Dikutip dari Wirawan Evaluasi: Teori, Model, Metodologi, Standar, Aplikasi dan Profesi (2016) halaman 125. Proses evaluasi berbasis tujuan dilaksanakan dengan tahapan sebagai berikut.

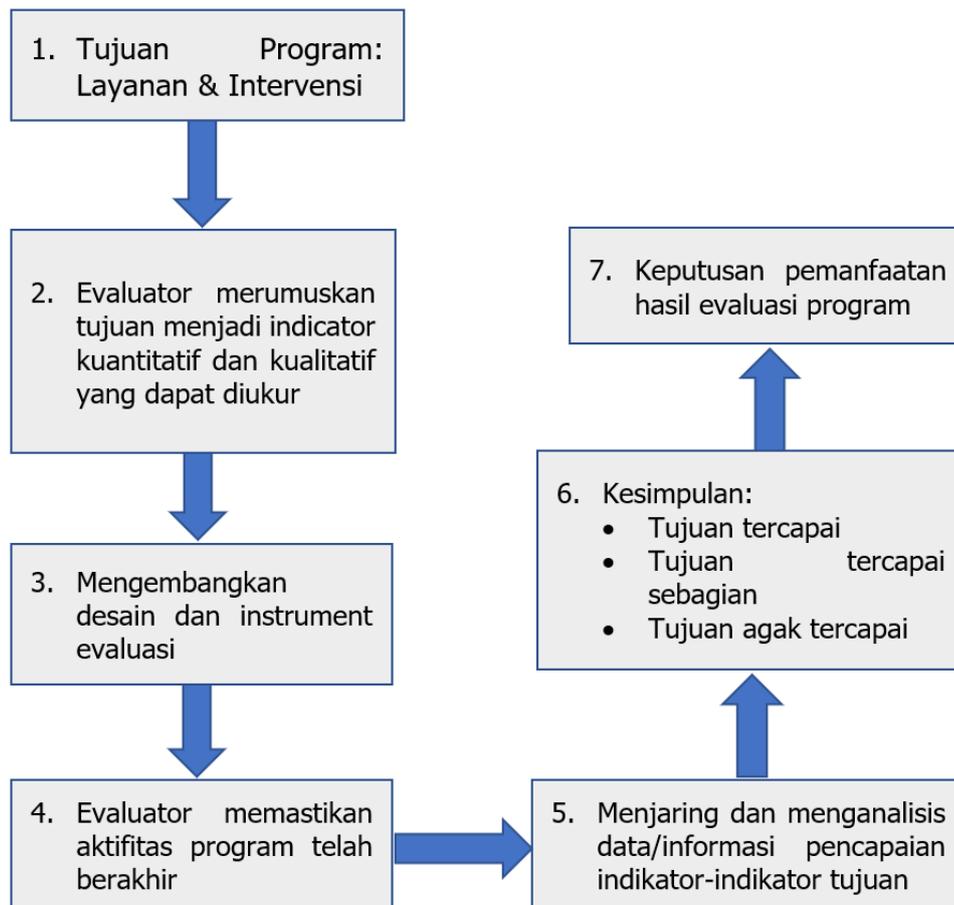
1. Mengidentifikasi tujuan, mengidentifikasi dan mendefinisikan tujuan intervensi, layanan dari program yang tercantum dalam rencana program. Tujuan program diuraikan dalam indikator-indikator kuantitas dan kualitas yang dapat diukur.
2. Merumuskan tujuan menjadi indikator-indikator. Tugas evaluator npada tahap ini adalah merumuskan tujuan program menjadi indikator-indikator kuantitatif dan nkualitatif yang dapat diukur.
3. Mengembangkan metode dan instrument untuk menjaring data. Evaluator menentukan metode yang akan digunakan dalam

⁸³ Daniel Stufflebeam. *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, Edited by Thomas Kellaghan, George F. Madaus, Daniel L. Stufflebeam (New York: Kluwer Academic Publishers, 2002), P.41.

⁸⁴ *Ibid.*, p. 41.

evaluasi apakah metode kuantitatif, kualitatif atau metode campuran. Dilanjutkan dengan mengembangkan instrument untuk menjaring data dan informasi.

4. Memastikan program telah berakhir dalam mencapai tujuan. Pada tahap ini, evaluator meastikan bahwa layanan, intervensi dari program telah dilaksanakan dan terdapat indicator pencapaian tujuan, pengaruh atau perubahan yang diharapkan.
5. Menjaring dan menganalisis data/informasi mengenai indicator-indikator program. Dengan menggunakan instrument evaluasi yang telah dikembangkan, evaluator menjaring data dan informasi yang hasilnya akan dianalisis. Proses analisis merujuk pada semua indicator program yang telah dirumuskan dalam indikator kuantitatif dan kualitatif seperti yang terdapat pada langkah kedua.
6. Kesimpulan. Berdasarkan hasil analisis data yang telah diolah, tahap kesimpulan mengukur hasil pencapaian program, atau pengaruh intervensi. Adakah perubahan yang diharapkan dari pelaksanaan program dan membandingkan dengan tujuan yang direncanakan dalam rencana program. Hasil analisis data juga digunakan untuk menentukan apakah terjadi ketimpangan.



Gambar 6: Proses Model Evaluasi Berbasis Tujuan

- Mengambil keputusan mengenai program. Dalam tahap ini akan ditentukan apakah program terlaksana sesuai rencana dan mendapat hasil yang diharapkan? Pertama, Jika iya, maka program bisa dilanjutkan atau dilaksanakan ditempat lain yang membutuhkan program yang sama. Kedua, Program bisa saja dihentikan seumpama tidak diperlukan lagi, karena program dimaksud telah mencapai tujuan sepenuhnya sehingga tidak ada lagi yang perlu dilakukan dengan program tersebut. Yang ketiga, jika program ternyata gagal tetapi masih dibutuhkan oleh masyarakat, maka program dapat dilanjutkan dengan perbaikan-perbaikan yang diperlukan setelah melalui proses analisis penyebab kegagalannya dan memodifikasinya.

Sebagai salah satu model evaluasi yang berkembang lebih awal dibanding model evaluasi lainnya, *goal-based evaluation model*

memiliki kelebihan dan kelemahan, Wirawan menyebutkan sebagai berikut.

Tabel 8: Kelebihan & Kelemahan Model *Goal-based Evaluation*

Keunggulan	Kelemahan
<p>a) <i>Demokratis</i>. Tujuan, layanan atau intervensi program merupakan hasil keputusan formal dari lembaga negara yang dipilih secara demokratis.</p> <p>b) <i>Imparsial</i>. Evaluasi merupakan bagian dari riset sosial yang bersifat imparsial, tidak memihak.</p> <p>c) Sederhana. Proses merancang dan melaksanakan mudah; biaya dan waktunya murah.</p>	<p>a) Tujuan tidak mudah dipahami. Penetapan tujuan seringkali tidak mudah dipahami, ambigu.</p> <p>b) Suatu tujuan berkaitan dengan ketidakpastian masa depan. Tujuan ditetapkan umumnya berhubungan dengan waktu sekarang. Namun, ketika program berjalan, terjadi perubahan situasi sehingga potensi terjadinya ketidaksesuaian antara tujuan dan kondisi yang dihadapi.</p> <p>c) Efek sampingan dari tujuan. Tidak sedikit munculnya efek samping dari sebuah tujuan yang ditetapkan.</p> <p>d) Adanya tujuan tersembunyi dari pengambil kebijakan. Seringkali tujuan suatu program hanyalah tujuan sementara dari tujuan yang hendak dicapai oleh pengambil kebijakan.</p>

C. **Goal Free Evaluation Model/Evaluasi Bebas Tujuan**

Model evaluasi bebas tujuan adalah model evaluasi dimana evaluator melakukan evaluasi tanpa mempunyai pengetahuan atau referensi dari *goal* dan *objective* serta pengaruh yang diharapkan oleh perancang program. Pengaruh program yang sesungguhnya mungkin berbeda atau lebih banyak atau lebih luas atau mungkin lebih sedikit dari tujuan yang dinyatakan dalam program. *Goal free evaluation model* berupaya mengukur keluaran dan pengaruh yang sesungguhnya tanpa dipengaruhi oleh tujuan dan pengaruh yang diharapkan dalam rencana program.⁸⁵

Scriven mendefinisikan evaluasi sebagai berikut “*What is evaluation? Synthesizing what the dictionaries and common usage tell us, it is the process of determining the merit, worth, or significance of things (near-synonyms are quality/value/importance).*”⁸⁶ Apakah Evaluasi itu? Menyintesis apa yang kamus dan masyarakat ceritakan kepada

⁸⁵ Wirawan. *Evaluasi: Teori, Model, Metodologi, Standar, Aplikasi dan Profesi*, (Jakarta: PT Rajagrafindo Persada, 2016), h. 127.

⁸⁶ Michael Scriven. *International Handbook of Educational Evaluation*, 15-30 T. Kellaghan & D.L. Stufflebeam (eds.) (Dordrecht: Kluwer Academic Publishers, 2003) p. 15.

kita, bahwa evaluasi adalah proses menentukan nilai, manfaat atau makna dari sesuatu (sinonimnya adalah kualitas/nilai/sesuatu yang penting). Ada tiga kata kunci dalam definisi ini, yaitu; nilai, manfaat dan makna penting sesuatu yang terkandung dalam proses evaluasi. Lebih lanjut Scriven mengatakan bahwa;

proposed goal-free evaluation as an approach in order to free the evaluator from the bias associated with evaluating only the objectives that program development personnel think are important. He reasoned that if a program has significant effects, they should be obvious to an observer who has not been informed about the intended effects. Goal-free evaluations are also an excellent means of identifying unintended side effects of a program, both positive and negative.⁸⁷

Evaluasi bebas tujuan menawarkan suatu pendekatan yang membebaskan evaluator dari bias terkait dengan kegiatan evaluasi yang hanya fokus pada pentingnya pemikiran pribadi dalam pengembangan program. Scriven beralasan, jika program memiliki pengaruh yang signifikan, maka akan menjadi jelas bagi peneliti yang sebelumnya tidak pernah diinformasikan mengenai pengaruh yang dimaksud/tujuan program tersebut. Evaluasi bebas tujuan juga merupakan alat yang sangat bagus untuk mengidentifikasi pengaruh sampingan dari program, baik pengaruh positif dan negatif. Pendapat Scriven ini sebagai salah satu pelopor evaluasi bebas tujuan melihat bahwa pengaruh program yang benar-benar nyata akan dapat diidentifikasi melalui proses evaluasi ini. Tujuan evaluasi yang ada dalam rencana program bukanlah efek sesungguhnya dari program yang dijalankan, meskipun tujuan yang dimaksud tetap harus ada. Tetapi dalam konteks evaluasi ini, evaluator harus dibebaskan dari pemikiran tentang tujuan yang telah direncanakan supaya kegiatan evaluasi dapat berjalan apa adanya. Hasilnya akan mencerminkan pengaruh yang sesungguhnya. Sekaligus akan ditemukan pengaruh sampingan yang tidak pernah ada dalam rencana.

⁸⁷ Donna M. Mertens and Amy T. Wilson. *Program Evaluation Theory and Practice: A Comprehensive Guide*, (New York: The Guilford Press A Division of Guilford Publications, Inc, 2019), P.137.

Scriven “has sharply criticized other views of evaluation and has used his critical analysis to extend his own position. He has charged that the Tylerian tradition (Scriven, 1991, 1993; Tyler, 1942), which sees evaluation as determining whether objectives have been achieved, is fundamentally flawed in that it is essentially value-free (meaning that the evaluator rather uncritically accepts the developer’s values as reflected in stated goals and, based on unvalidated goals, makes claims or conclusions about something’s quality or value that are not objectively defensible). He has argued that this approach is potentially invalid, because a developer’s goals may be immoral, unrealistic, unrepresentative of the assessed needs of consumers, mainly in the developer’s interest, or too narrow to encompass possibly crucial side effects.⁸⁸

Scriven mengkritik tajam pandangan lain tentang evaluasi dan menggunakan analisis kritisnya untuk menyampaikan posisinya. Scriven menyerang tradisi Tylerian yang melihat evaluasi sebagai proses menentukan apakah tujuan telah dicapai, adalah cacat karena sesungguhnya evaluasi harus bebas tujuan (yang berarti bahwa evaluator kurang kritis menerima nilai pengembang program seperti direfleksikan dalam pernyataan tujuan dan berdasarkan tujuan yang tidak divalidasi, mengklaim atau menyimpulkan tentang suatu kualitas atau nilai yang tidak bisa dipertahankan secara objektif). Scriven beralasan bahwa pendekatan seperti ini berpotensi invalid, karena tujuan pengembang program bisa jadi tidak sesuai, tidak realistis, tidak representative bagi kebutuhan konsumen atas penilaian, terutama maksud pengembang, atau terlalu sempit untuk mengcover pengaruh sampain yang mungkin terjadi.

Model evaluasi ini berusaha mengabaikan tujuan program yang telah ditetapkan sebelumnya, karena akan berpengaruh terhadap hasil evaluasi yang kurang menunjukkan pengaruh sesungguhnya dari program yang dijalankan. Evaluasi bebas tujuan ingin mengetahui *the real effect of the program* terhadap sasaran program.

⁸⁸ Daniel L. Stufflebeam and Chris L.S. Coryn. *Evaluation Theory, Models, and Applications*, (San Fransisco: Jossey-Bass A Wiley Brand), P.344.

According to this approach the evaluator purposely remains ignorant of a program's printed goals and searches for all effects of a program irrespective of its developer's objectives. There are no side effects to examine, since data about all effects, whatever the intent of the program, are equally admissible. If a program is doing what it was supposed to do, then the evaluation should confirm this, but the evaluator will also be more likely to uncover unanticipated effects that the goal-based evaluators would miss because of their preoccupation with stated goals.⁸⁹

Menurut pendekatan ini evaluator secara sengaja mengabaikan tujuan program yang telah ditetapkan dan mencari seluruh pengaruh program terlepas dari tujuan yang dikembangkan penyusun program. Tidak ada pengaruh sampingan yang dapat untuk diuji, karena data mengenai semua pengaruh, apakah tujuan program, sesuai dengan yang diinginkan. Jika program melakukan apa yang seharusnya dilakukan, maka evaluasi seharusnya dapat mengonfirmasi ini, tetapi evaluator juga sangat mungkin bisa lebih membuka pengaruh-pengaruh yang tidak direncanakan, yang evaluator berbasis tujuan bisa saja terlewatkan karena terlalu asyik dengan tujuan yang telah ditetapkan.

Scriven adalah salah satu tokoh pengembang evaluasi bebas tujuan sebagai upaya untuk mengekspresikan bahwa evaluasi bebas tujuan memfasilitasi evaluator menemukan pengaruh yang sesungguhnya dari suatu program. Bahwa setiap program bisa jadi telah memiliki tujuan, tetapi tujuan program yang dimaksud belum tentu mencerminkan pengaruh program yang murni muncul dari implementasi program tersebut. Sebagai contoh, jika salah satu tujuan dibukanya sekolah terbuka dan universitas terbuka adalah pemerataan Pendidikan bagi seluruh masyarakat, maka dalam evaluasi bebas tujuan seorang evaluator tidak perlu memedulikan tujuan tersebut. Evaluator hanya perlu fokus pada pengaruh yang muncul dengan adanya sekolah terbuka dan universitas terbuka terhadap masyarakat. Sekalipun hasil

⁸⁹ Daniel L. Stufflebeam Anthony J. Shinkfield. *Systematic Evaluation A Self-Instructional Guide to Theory and Practice*, (Boston: Kluwer-Nijhoff Publishing a member of the Kluwer Academic Publishers Group,1988), P. 317.

evaluasinya menunjukkan adanya indikator tentang pemerataan Pendidikan bagi masyarakat, tetapi pengaruh tersebut mungkin hanya salah satu saja. Pengaruh lain yang muncul bisa jadi tidak pernah dideskripsikan dalam rencana pembukaan sekolah dan universitas terbuka.

Mengutip penjelasan Wirawan bahwa suatu program dapat memiliki tiga jenis pengaruh yaitu;

1. Pengaruh sampingan yang negatif. Merupakan pengaruh sampingan yang tidak dikehendaki program. Kebijakan sekolah gratis sesungguhnya dimaksudkan untuk memangkas kesenjangan antara yang kaya dan yang kurang mampu dalam mendapatkan Pendidikan. Ternyata, efek samping negative yang muncul adalah sebagian siswa dan orang tua menjadi kurang gregat dan termotivasi untuk berprestasi bahkan cenderung malas belajar karena tidak menanggung biaya Pendidikan. Sehingga ketika tidak berprestasi dan bahkan tidak naik ke jenjang/level yang lebih tinggi mereka tidak terlalu peduli.
2. Pengaruh positif yang ditetapkan oleh tujuan program. Suatu program dilaksanakan tentu memiliki tujuan. Tujuan program yang diharapkan oleh pembuat kebijakan dapat dipastikan sesuatu yang positif. Seperti contoh di atas tentang program sekolah gratis memiliki tujuan yang sangat mulia agar akses Pendidikan dapat dirasakan oleh seluruh lapisan masyarakat, baik yang kurang mampu dan yang mampu mendapatkan pelayanan Pendidikan sama. Bahwa ada pengaruh sampingan yang negative merupakan hal yang tidak diinginkan dalam perencanaan program.
3. Pengaruh positif sesuai dengan tujuan program. Adalah pengaruh positif yang diharapkan oleh perancang program. Misalnya, kandidat kepala daerah yang akan maju dalam pilkada menyampaikan program sekolah gratis bagi seluruh masyarakat tanpa memandang latar belakang dan status ekonomi untuk pemerataan Pendidikan terutama bagi golongan masyarakat yang

kurang mampu. Sesungguhnya para politisi ini memiliki tujuan tersembunyi agar dianggap berpihak kepada rakyat dan dipilih menjadi kepala daerah di wilayah pemilihannya.

Menurut Scriven yang dikutip oleh Stufflebeam dan Coryn “*Advantages of goal-free evaluation, according to Scriven (1973, 1974), are that it is less intrusive than goals-based evaluation; more adaptable to midstream goal shifts; better at finding side effects; less prone to social, perceptual, and cognitive biases; more professionally challenging; and more equitable in considering a wide range of values.*”⁹⁰ Keuntungan dari evaluasi bebas tujuan adalah bahwa gangguan lebih sedikit dibandingkan evaluasi berbasis tujuan: lebih dapat diadaptasi dalam perubahan tujuan arus utama; lebih baik dalam menemukan pengaruh sampingan; cenderung lebih mudah bagi social, pemahaman, bias kognitif; secara profesional lebih menantang, dan lebih tepat sebagai pertimbangan nilai yang lebih luas. Gangguan lebih sedikit dibanding evaluasi berbasis tujuan artinya bahwa evaluator tidak dibebani oleh tujuan program yang telah ditentukan. Evaluator hanya focus pada identifikasi pengaruh apa saja yang muncul dari implementasi program. Evaluator juga menjadi lebih fleksibel menyesuaikan ketika terjadi perubahan tujuan utama dari program.

*Goal-free evaluation is an innovative approach that is helpful in implementing the consumer-oriented approach to evaluation. In our evaluation practice, we have found that goal-free evaluation provides important supplementary information, expands the sources of evaluative information, is especially good in turning up unexpected findings, is a relatively low-cost procedure, and is welcomed and appreciated by clients.*⁹¹ Evaluasi bebas tujuan adalah pendekatan inovatif yang sangat berguna dalam implementasi evaluasi dengan pendekatan yang berorientasi pada pelanggan. Dalam praktik evaluasi yang dilakukan, kami menemukan bahwa evaluasi bebas tujuan menyediakan informasi tambahan yang sangat penting, mengembangkan sumber informasi

⁹⁰ Daniel L. Stufflebeam and Chris L.S. Coryn. *Evaluation theory, models, and applications*, (San Fransisco: Jossey-Bass A Wiley Brand), P. 348.

⁹¹ Daniel L. Stufflebeam and Chris L.S. Coryn. *Evaluation theory, models, and applications*, (San Fransisco: Jossey-Bass A Wiley Brand), P. 348

evaluasi yang lebih luas, terutama sangat cocok untuk menemukan hal-hal yang tidak pernah diduga sebelumnya, secara prosedur biaya lebih efisien, dan terbuka serta apresiasi dari klien. Penjelasan di atas membuka wawasan kita tentang evaluasi bebas tujuan yang lebih fleksibel menerima perubahan yang terjadi dalam proses evaluasi. Memungkinkan evaluator menemukan informasi tambahan yang tidak bisa didapatkan jika menggunakan model lain. Evaluasi bebas tujuan juga terbuka dengan informasi dan masukan-masukan tambahan yang muncul belakangan saat kegiatan evaluasi berlangsung.

Menurut Brandon W. Youker, Allyssa Ingraham, dan Nicholas Bayer (2014), terdapat empat langkah untuk menemukan *goal* dan pengaruh program yang sesungguhnya dalam *Goal Free Evaluation Model*.

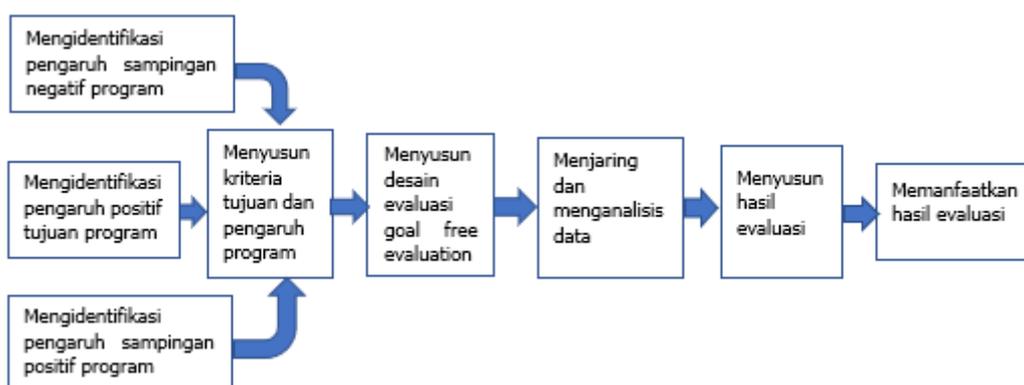
1. Meneliti dan mengidentifikasi pengaruh program yang relevan tanpa merujuk goal dan objektif yang ada di rencana program.
2. Mengidentifikasi apa yang terjadi tanpa merujuk pada tujuan dan sasaran program.
3. Menentukan pengaruh apa yang muncul dan secara logis disebabkan oleh program dan intervensi program.
4. Menentukan derajat pengaruh positif, negatif atau netral dari program.

	<p>Scriven's Background</p> <p>Scriven is a philosopher of science and an expert in critical thinking. He was born in England and raised in Australia. He earned his bachelor's degree in mathematics and his master's degree in applied mathematics and symbolic logic at the University of Melbourne. Subsequently, he completed his PhD in philosophy of science at Oxford University. He has served at ten universities in Australia, New Zealand, and the United States. From 2004 to 2007 he was professor of philosophy and associate director of the Evaluation Center at Western Michigan University as well as director of the IDPE program. Currently he holds a professorial appointment at Claremont Graduate University.</p>
---	---

Profil 4: MICHAEL SCRIVEN

Scriven berargumen apakah evaluasi program harus dipandu dengan tujuan-tujuan, beliau meyakini, bahwa evaluator harus mencari seluruh hasil atau efek dari implementasi program, menilai kebutuhan pelanggan, dan menggunakan serangkaian penilaian untuk sampai pada kesimpulan tentang nilai dan manfaat suatu program. Ketergantungan pada tujuan yang telah ditetapkan sebelumnya berpotensi memunculkan bias terhadap hasil yang diperoleh dalam evaluasi.

Seperti halnya model evaluasi yang lain, evaluasi bebas tujuan juga dapat dibuatkan skema pelaksanaannya sebagai berikut.



Gambar 7: Pelaksanaan Model Evaluasi Bebas Tujuan

Salah satu kesulitan dalam menggunakan goal-free evaluation adalah bagaimana caranya menentukan makna nilai terhadap temuan evaluasi. Jika hasilnya teridentifikasi tidak mencapai sesuatu yang diharapkan, maka bagaimana seseorang akan dapat memilah mana hal-hal yang diinginkan dan mana hal atau efek yang tidak diinginkan? Jawabannya adalah harus membandingkan antara hasil yang dicapai dengan kebutuhan pelanggan. Tetapi jika diperlukan, perbedaan antara sesuatu yang real dan sesuatu yang ideal dan apakah sesuatu yang ideal itu merupakan tujuan, bukankah asesmen kebutuhan berdasarkan target, dan untuk itu apakah evaluasi bebas tujuan juga bukan berbasis tujuan? Jawabannya adalah bukan. Alasannya adalah sebagai bentuk kritik terhadap *Goal-based Evaluation* (GBE) bahwa pengembang evaluasi berbasis tujuan sama sekali tidak konsisten dengan beberapa hal yang bersifat ideal seperti sesuatu yang ada dalam demokrasi. Dalam hal tertentu, ia (pengembang GBE) menjaga konsep

klasik tentang perlunya perbedaan antara sesuatu yang real dan sesuatu yang ideal adalah salah, karena sesuatu yang ideal sering tidak realistis. Karena kebutuhan konsumen merupakan konsep mendasar dalam pendekatan evaluasi ini, ia (pengembang GBE) dan pengikutnya secara berkelanjutan mengembangkan dan meneliti konsep ini.

D. Evaluasi Formatif dan Sumatif

Evaluasi formatif dan sumatif merupakan salah satu model evaluasi yang dikembangkan oleh Scriven selain model bebas tujuan. Pada awalnya Scriven menggunakan istilah *outcome evaluation of an intermediate stage in development of the teaching instrument* pada tahun 1967. Evaluasi formatif adalah evaluasi yang dilaksanakan pada saat program sedang berjalan sebagai upaya untuk mengukur perkembangan program. Evaluasi sumatif dilakukan saat program sudah selesai. Dalam konteks sekolah evaluasi sumatif dilakukan di akhir semester setelah seluruh program pembelajaran dan proses pembelajaran dilakukan.

Tujuan evaluasi formatif tersebut adalah mengetahui seberapa jauh program yang dirancang dapat berlangsung, sekaligus mengidentifikasi hambatan. Perbaikan dapat segera dilakukan ketika hambatan dan kendala yang ada teridentifikasi. Evaluasi sumatif bertujuan untuk mengukur ketercapaian program. Fungsi evaluasi sumatif dalam evaluasi program pembelajaran adalah sebagai sarana untuk mengetahui posisi atau kedudukan individu di dalam kelompoknya.⁹²

Stufflebeam & Shlnkfeld "*Formative evaluation is an integral part of the development process. It provides continual feedback to assist in planning and then producing some object. In curriculum development it addresses questions about content validity, vocabulary level, usability, appropriateness of media, durability of materials, efficiency, staffing, and other matters. In general, formative evaluation is done to help staff to improve whatever they are operating or developing.*"⁹³ Evaluasi formatif

⁹² Suharsimi Arikunto dan Cepi Safruddin Abdul Jabar. *Evaluasi Program Pendidikan: Pedoman Teoritis, Praktis Bagi Mahasiswa dan Praktisi Pendidikan*, (Jakarta: Bumi Aksara, 2009), h. 42.

⁹³ Daniel L. Stufflebeam and Anthony J. Shlnkfeld. *Systematic Evaluation A Self-Instructional Guide to Theory and Practice*, (Norwel: Kluwer-Nijhoff Publishing, 1988), P. 315.

adalah bagian integral proses pengembangan. Evaluasi ini memberikan umpan balik secara berkelanjutan untuk membantu dalam perencanaan dan memperbaiki objek. Dalam pengembangan kurikulum, evaluasi formatif membicarakan tentang validitas isi, level perbendaharaan kata, kegunaan, kesesuaian media, daya tahan material, efisiensi, staffing, dan masalah-masalah lain. Secara umum, evaluasi formatif dilakukan untuk membantu staf memperbaiki apakah mereka menjalankan atau mengembangkan tugasnya sesuai target.

Madaus & Kellaghan menyebutkan “Formative evaluation is conducted during the development or improvement of a program or product (or person, etc.). It is an evaluation which is conducted for the in-house staff of the program and normally remains in-house; but it may be done by an internal or an external evaluator or (preferably) a combination. The distinction between formative and summative has been well summed up in a sentence of Bob Stake: “When the cook tastes the soup, that’s formative; when the guests taste the soup, that’s summative.”⁹⁴

Evaluasi formatif dilakukan selama program atau produk (atau orang, dll) dalam pengembangan atau perbaikan. Merupakan evaluasi yang dilakukan bagi staf program dan normalnya di dalam lingkup organisasi; tetapi bisa juga dilakukan oleh internal maupun eksternal evaluator atau kombinasi keduanya.

Stufflebeam *“If the information is used to guide development, the evaluation is formative. If it is used to sum up the value of something, the evaluation is summative. In these respects, the same data may be used for either formative or summative evaluation. Evaluation in its formative application is an integral part of the development process. It provides continual feedback to assist in planning, developing, and delivering a program or service.”*⁹⁵ Jika informasi yang digunakan untuk membantu pengembangan, evaluasi itu formatif. Jika informasinya digunakan

⁹⁴ George F. Madaus and Thomas Kellaghan. *Evaluation Models Viewpoints on Educational and Human Services Evaluation*, Edited by Thomas Kellaghan, George F. Madaus, Daniel L. Stufflebeam, (New York: Kluwer Academic Publishers, 2002), P.28

⁹⁵ Daniel L. Stufflebeam and Chris L.S. Coryn. *Evaluation theory, models, and applications*, (San Fransisco: Jossey-Bass A Wiley Brand), Pp. 345-346.

untuk menyimpulkan nilai sesuatu, evaluasi tersebut adalah sumatif. Dalam hal ini, data yang sama bisa digunakan untuk melakukan evaluasi formatif dan sumatif. Evaluasi dalam penerapan formatif adalah bagian integral dari proses pengembangan. Yang menyediakan umpan balik berkelanjutan untuk membantu perencanaan, pengembangan, dan menyampaikan program atau pelayanan.

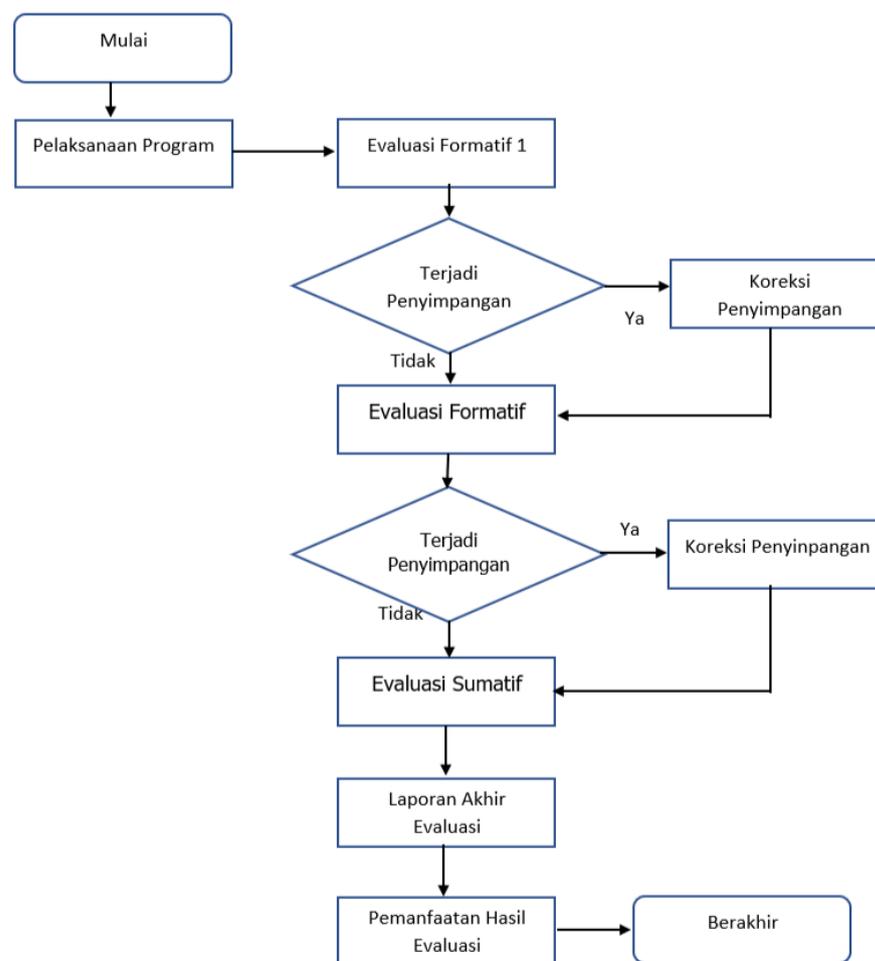
Wirawan dalam bukunya *Evaluasi: Teori, Model, Metodologi, Standar, Aplikasi dan Profesi* menyebutkan tujuan dilakukannya evaluasi formatif.

1. Untuk mengukur hasil pelaksanaan program secara periodik.
2. Untuk mengukur apakah klien/pertisipan bergerak kearah tujuan yang direncanakan.
3. Untuk mengukur apakah sumber-sumber telah dipergunakan sesuai dengan rencana.
4. Untuk menentukan koreksi apa yang harus dilakukan jika terjadi penyimpangan. Penyimpangan bisa dalam bentuk positif dan negative. Penyimpangan positif adalah apabila target terlampaui, artinya program dilaksanakan lebih cepat dari yang direncanakan. Ini merupakan penyimpangan yang diharapkan. Penyimpangan negatif adalah penyimpangan yang tidak diharapkan dan harus dihindari. Jika terjadi penyimpangan, sumber harus dilakukan penghematan pada kegiatan termin berikutnya. Dasampping itu harus dilakukan audit kegiatan dan penggunaan anggaran.
5. Memberikan balikan. Evaluasi formatif memberikan balikan secara terus menerus untuk memperbaiki perencanaan, standar prosedur operasi, penggunaan sumber-sumber dan perkembangan pelaksanaan program.

Selanjutnya Stuffkebeam dan Shinkfield mengatakan bahwa,

In the summative role, evaluation "may serve to enable administrators to decide whether the entire finished curriculum, refined by use of the evaluation process in its first (formative) role, represents a sufficiently significant advance on the available

alternatives to justify the expense of adoption by a school system." Summative evaluation usually should be performed by an external evaluator in order to enhance objectivity, and the findings should be made public. This type of evaluation searches for all effects of the object and examines them against the assessed needs of the relevant consumers.⁹⁶



Gambar 8: Diagram Aliran Proses Evaluasi Formatif dan Sumatif
Peran evaluasi sumatif bisa memberi pelayanan untuk memungkinkan administrator menentukan apakah keseluruhan kurikulum telah selesai, diperbaiki melalui penerapan proses evaluasi dalam kegiatan evaluasi awal (formatif), mewakili perkembangan yang cukup signifikan sebagai alternatif untuk menentukan biaya adopsi oleh system sekolah.

⁹⁶ Daniel L. Stufflebeam and Anthony J. Shlnkfield. *Systematic Evaluation A Self-Instructional Guide to Theory and Practice*, (Norwel: Kluwer-Nijhoff Publishing, 1988), P. 315-316.

Evaluasi sumatif biasanya dilakukan oleh evaluator eksternal guna meningkatkan objektivitas, dan hasil temuannya dapat dipublikasi. Evaluasi bentuk ini ingin mencari semua pengaruh dari program dan mengujinya dengan menilai kebutuhan pelanggan yang relevan. Evaluasi sumatif dilaksanakan pada akhir program. Evaluasi ini mengukur kinerja akhir objek evaluasi. Menurut Steven et al. dikutip Wirawan tujuan evaluasi sumatif adalah sebagai berikut;

1. Menentukan kesuksesan keseluruhan program;
2. Menentukan apakah tujuan umum dan tujuan khusus telah dicapai;
3. Menentukan apakah klien mendapat keuntungan dari program;
4. Menentukan komponen mana paling efektif dan komponen mana yang kurang efektif;
5. Menentukan apakah ada keluaran yang tidak diantisipasi;
6. Menentukan cost benefit program;
7. Mengomunikasikan temuan evaluasi kepada para pemangku kepentingan program.
8. Mengambil keputusan apakah program harus dihentikan, dikembangkan, dihentikan dan dilaksanakan ditempat lain.

Evaluasi sumatif berupaya mengukur indikator-indikator berikut:

1. Hasil dan pengaruh layanan atau intervensi program;
2. Mengukur persepsi klien mengenai layanan dan intervensi program;
3. Menentukan *cost effectiveness*, *cost efficiency* dan *cost benefit*.

E. Evaluasi Responsif

Salah satu pelopor evaluasi responsive adalah Robert Stake (1975), yang pada mulanya menamai *Countenance of Educational Evaluation*. Pada awalnya Stake berpikir tentang bagaimana mengevaluasi program yang merupakan pengembangan ilmu pengetahuan sosial empiris dan psikometrik, dimana objektivitas dan personalisasinya lebih dihargai. Dalam upayanya melakukan evaluasi perubahan kurikulum, Stake menemukan bahwa tidak satupun desain maupun tes yang digunakan untuk mengumpulkan data dapat memberi jawaban yang memadai

atas pertanyaan-pertanyaan penting. Sehingga beliau merespon untuk mengembangkan model evaluasi responsif. *'Responsive evaluation was my response to "preordinate evaluation", prior selection and final measurement of a few outcome criteria. Over the years I came to be comfortable with the idea that disciplining impressions and personal experience led to better understanding of merit and worth than using needs to identify improvement with strict controls on bias.'*⁹⁷ Evaluasi responsif adalah respon saya terhadap "evaluasi preordinat", seleksi awal dan pengukuran akhir terhadap beberapa kriteria hasil yang diinginkan.

Daniel Stufflebeam dan Anthony J. Shinfield (1985) menamai evaluasi responsif: *Client-centered evaluation* atau evaluasi yang berpusat pada klien.

Responsive evaluation is an approach, a predisposition, to the evaluation of educational and other programs. Compared to most other approaches it draws attention to program activity, to program uniqueness, and to the social plurality of its people. This same predisposition toward merit and shortcoming can be built into or can be recognized in other approaches, such as a stakeholder evaluation or connoisseurship evaluation. A responsive evaluation is a search and documentation of program quality. The essential feature of the approach is a responsiveness to key issues or problems, especially those recognized by people at the site. It is not particularly responsive to program theory or stated goals but more to stakeholder concerns.⁹⁸

⁹⁷ Robert Stake. *International Handbook of Educational Evaluation Part One: Perspectives*, Editors: Thomas Kellaghan & Daniel L. Stufflebeam, (Dordrecht: Kluwer Academic Publishers, 2003), P.65.

⁹⁸ Robert Stake, *International Handbook of Educational Evaluation*, edited by T. Kellaghan, D.L. Stufflebeam (eds.) (Dordrecht: Kluwer Academic Publishers, 2003), p. 63.

Robert Stake was born in December 18, 1927, Nebraska, America. He is Director of the Center for Instructional Research and Curriculum Evaluation at the University of Illinois at Urbana-Champaign. He is one of several educational researchers who created theory and practice for educational program evaluation in the 1960s. His responsive evaluation approach emphasizes the study of classroom experience, personal interaction, and institutional processes and contexts, often in the form of case studies. Among the evaluative studies he has directed are studies in science and arts education; model programs; and conventional teaching, including higher education, special education and, with Bernadine Evans Stake, gender equity. He is a recipient of the Special Career Award in Qualitative Inquiry from the International Congress for Qualitative Inquiry, the Lazarsfeld Award from the American Evaluation Association, and the Presidential Citation from the American Educational Research Association, and holds honorary doctorates from the University of Uppsala, Sweden, and the University of Valladolid, Spain.



FORMAL EDUCATION

1948 B.A. Honours School of Mathematics, University of Melbourne, Australia

1950 M.A. Combined Honours School of Mathematics and Philosophy, University of Melbourne. Thesis title: "The Scope and Limitations of Formal Logic"

1956 D.Phil., School of Literae Humaniores, Oxford University. Supervisor, Gilbert

Ryle; examiners, H. H. Price and R. B. Braithwaite. Thesis title: "Explanations" (a study of the logic of explanations in the sciences and the humanities)

2013 D. Ed. (Honoris causa) University of Melbourne

Profile 5: Robert Stake

Evaluasi responsif adalah suatu pendekatan, kecenderungan, untuk evaluasi Pendidikan dan program lainnya. Dibandingkan dengan pendekatan lain, evaluasi responsif memberi perhatian pada aktifitas program, keunikan program, dan keragaman sosial masyarakatnya. Kecenderungan ini mirip dengan nilai dan kelemahannya dapat diperbaiki didalam atau dapat dikenali di pendekatan evaluasi lain, seperti evaluasi pemangku kepentingan atau evaluasi konosersip. Evaluasi responsive adalah proses pencarian dan pendokumentasian kualitas program. Esensi pendekatan evaluasi responsif adalah respon terhadap isu-isu inti atau masalah, terutama isu atau masalah yang menjadi perhatian masyarakat di tempat kegiatan evaluasi. Bukan responsive terhadap teori program atau tujuan yang telah ditentukan tetapi lebih kepada hal yang menjadi perhatian pemangku kepentingan. Kata kuncinya adalah bahwa evaluasi ini fokus pada aktifitas program apakah memiliki pengaruh terhadap sasaran program. Perhatian pada keunikan program dan keragaman sosial masyarakatnya berkenaan dengan nilai yang ada didalam program, seberapa besar pengaruhnya

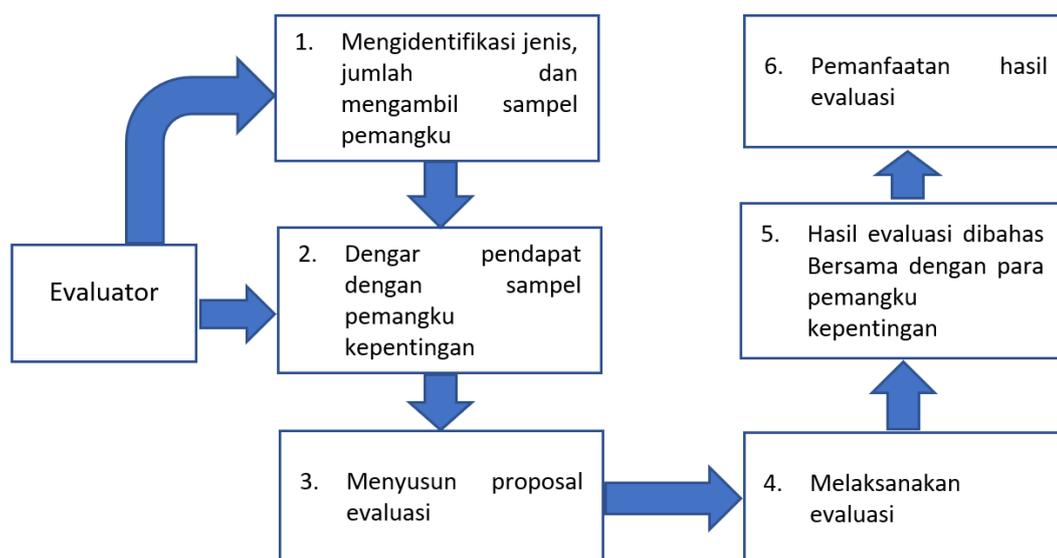
atau efek yang ditimbulkan terhadap sasaran program baik efek positif maupun negatif.

Mertens dan Wilson mengutip pernyataan Stake sebagai pengembang evaluasi responsif; According to Stake (1991), a responsive evaluation includes an agreement to begin the study with observations, the time and place for which are negotiated with stakeholders. The evaluator will use the data from the observations to prepare brief reports that include narrative data, product displays (e.g., student work), and graphs. Stakeholders are then asked to indicate what aspects of the preliminary report are of value to them and what diversity of opinions exists among them in this respect. Stakeholders are further asked to react to the accuracy of the reports, the importance of the various findings, and their relevance. All of this can be done fairly informally; the evaluator does keep a written record of what is presented and of the stakeholders' reactions.⁹⁹

Menurut Stake, evaluasi responsif meliputi kesepakatan untuk memulai penelitian dengan observasi, waktu dan tempat dapat dinegosiasikan dengan para pemangku kepentingan. Evaluator akan menggunakan data dari hasil observasi untuk mempersiapkan laporan singkat yang meliputi data naratif, menampilkan produk (karya siswa), dan grafik. Kemudian, Para pemangku kepentingan diminta untuk menunjukkan aspek apa saja dalam laporan awal yang bernilai bagi mereka dan apa perbedaan pendapat yang ada diantara mereka. Lebih lanjut, para pemangku kepentingan diminta untuk memberi reaksi terhadap akurasi laporan, pentingnya berbagai penemuan, dan relevansinya. Semuanya dapat dilakukan dengan adil secara informal; evaluator menyimpan rekaman tulisan mengenai hal yang muncul dan reaksi para pemangku kepentingan.

Wirawan menjabarkan proses pelaksanaan Model Evaluasi Responsif sebagai berikut.

⁹⁹ Donna M. Mertens and Amy T. Wilson. *Program Evaluation Theory and Practice: A Comprehensive Guide*, (New York: The Guilford Press A Division of Guilford Publications, Inc, 2019), P.145.



Gambar 9: Proses Model Evaluasi Responsif

1. Evaluator mengidentifikasi jenis dan jumlah setiap pemangku kepentingan (responden). Jika jenisnya terlalu banyak, maka harus dirangking berdasarkan pentingnya setiap pemangku kepentingan bagi program. Evaluasi memiliki keterbatasan sumber dan waktu pelaksanaan.
2. Melakukan dengar pendapat dengan pemangku kepentingan. Evaluator dapat mengunjungi sampel pemangku kepentingan secara langsung dan berbincang-bincang dengan mereka. Dengar pendapat merupakan bagian dari penelitian pendahuluan.
3. Menyusun proposal. Proposal evaluasi disusun dengan memerhatikan pendapat para pemangku kepentingan.
4. Melaksanakan evaluasi. Kegiatan evaluasi dilakukan dengan menjalin komunikasi dengan para pemangku kepentingan dan pimpina serta staf program.
5. Membahas hasil evaluasi dengan para pemangku kepentingan. Proses ini untuk mendapatkan masukan, kritik, dan saran sebanyak mungkin dari para pemangku kepentingan termasuk pimpinan dan staf program.

6. Pemanfaatan hasil evaluasi. Evaluator merekomendasikan hasil evaluasi untuk diterima dan digunakan oleh para pemangku kepentingan.¹⁰⁰

RANGKUMAN

Evaluasi berbasis tujuan ingin mengukur apakah tujuan program atau proyek yang telah ditetapkan, dapat dicapai atau tidak. Model evaluasi ini memfokuskan pada mengumpulkan informasi yang bertujuan mengukur pencapaian tujuan kebijakan, program dan proyek untuk pertanggungjawaban dan pengambilan keputusan. Jika program tidak mempunyai tujuan, atau tidak mempunyai tujuan yang bernilai maka program tersebut merupakan program yang buruk.

Evaluasi bebas tujuan adalah model evaluasi dimana evaluator melakukan evaluasi tanpa mempunyai pengetahuan atau referensi dari *goal* dan *objective* serta pengaruh yang diharapkan oleh perancang program. Pengaruh program yang sesungguhnya mungkin berbeda atau lebih banyak atau lebih luas atau mungkin lebih sedikit dari tujuan yang dinyatakan dalam program. *Goal free evaluation model* berupaya mengukur keluaran dan pengaruh yang sesungguhnya tanpa dipengaruhi oleh tujuan dan pengaruh yang diharapkan dalam rencana program.

Evaluasi formatif adalah evaluasi yang dilaksanakan pada saat program sedang berjalan sebagai upaya untuk mengukur perkembangan program. Tujuan evaluasi formatif tersebut adalah mengetahui seberapa jauh program yang dirancang dapat berlangsung, sekaligus mengidentifikasi hambatan. Perbaikan dapat segera dilakukan ketika hambatan dan kendala yang ada teridentifikasi.

Evaluasi sumatif dilakukan saat program sudah selesai. Dalam konteks sekolah evaluasi sumatif dilakukan di akhir semester setelah seluruh program pembelajaran dan proses pembelajaran dilakukan. Evaluasi sumatif bertujuan untuk mengukur ketercapaian program. Fungsi evaluasi sumatif dalam evaluasi program pembelajaran adalah sebagai sarana untuk mengetahui posisi atau kedudukan individu di dalam kelompoknya.

Evaluasi responsive adalah proses pencarian dan pendokumentasian kualitas program. Esensi pendekatan evaluasi responsif adalah respon terhadap isu-isu inti atau masalah, terutama isu atau masalah yang menjadi perhatian masyarakat di tempat kegiatan evaluasi. Bukan responsive terhadap teori program atau tujuan yang telah ditentukan tetapi lebih kepada hal yang menjadi perhatian pemangku kepentingan. Perhatian pada keunikan program dan keragaman sosial masyarakatnya berkenaan dengan nilai yang ada didalam program, seberapa besar pengaruhnya terhadap sasaran program baik efek positif maupun negatif.

¹⁰⁰ Wirawan. Evaluasi, teori, Model, Metodologi, Aplikasi dan Profesi, (Jakarta: Raja Grafindo Persada, 2016), hh. 135-136.

Latihan 9

1. *Goal-Based Evaluation* adalah...!
2. *Free Goal Evaluation* adalah...!
3. *Responsive Evaluation* adalah...!
4. *Sumatif Evaluation* adalah...!
5. *Formatif Evaluation* adalah...!

Pembahasan Latihan Soal 9

1. Model evaluasi berbasis tujuan secara umum ingin mengukur apakah tujuan program atau proyek yang telah ditetapkan, dapat dicapai atau tidak. Model evaluasi ini memfokuskan pada mengumpulkan informasi yang bertujuan mengukur pencapaian tujuan kebijakan, program dan proyek untuk pertanggungjawaban dan pengambilan keputusan. Jika program tidak mempunyai tujuan, atau tidak mempunyai tujuan yang bernilai maka program tersebut merupakan program yang buruk
2. Evaluasi bebas tujuan adalah model evaluasi dimana evaluator melakukan evaluasi tanpa mempunyai pengetahuan atau referensi dari *goal* dan *objective* serta pengaruh yang diharapkan oleh perancang program. *Goal free evaluation model* berupaya mengukur keluaran dan pengaruh yang sesungguhnya tanpa dipengaruhi oleh tujuan dan pengaruh yang diharapkan dalam rencana program.
3. Evaluasi responsive adalah proses pencarian dan pendokumentasian kualitas program. Esensi pendekatan evaluasi responsif adalah respon terhadap isu-isu inti atau masalah, terutama isu atau masalah yang menjadi perhatian masyarakat di tempat kegiatan evaluasi. Bukan responsive terhadap teori program atau tujuan yang telah ditentukan tetapi lebih kepada hal yang menjadi perhatian pemangku kepentingan.
4. Evaluasi formatif adalah evaluasi yang dilaksanakan pada saat program sedang berjalan sebagai upaya untuk mengukur perkembangan program. Tujuan evaluasi formatif tersebut adalah mengetahui seberapa jauh program yang dirancang dapat berlangsung, sekaligus mengidentifikasi hambatan. Perbaikan dapat segera dilakukan ketika hambatan dan kendala yang ada teridentifikasi.
5. Evaluasi sumatif dilakukan saat program sudah selesai. Dalam konteks sekolah evaluasi sumatif dilakukan di akhir semester setelah seluruh

program pembelajaran dan proses pembelajaran dilakukan. Evaluasi sumatif bertujuan untuk mengukur ketercapaian program. Fungsi evaluasi sumatif dalam evaluasi program pembelajaran adalah sebagai sarana untuk mengetahui posisi atau kedudukan individu di dalam kelompoknya

Tes Formatif Pilihan Berganda

1. Nama lain evaluasi responsive adalah...!
 - a. *Countenance of Educational Evaluation*
 - b. *Coincidence Evaluation*
 - c. *Post Academic Evaluation*
 - d. *Formative Academic Evaluation*
2. Evaluasi formatif adalah evaluasi yang dilaksanakan pada saat...!
 - a. Program belum berjalan sebagai upaya persiapan agar berhasil sesuai harapan.
 - b. Program direncanakan dan pada saat program dimulai untuk mengukur kesiapan program.
 - c. Program sedang berjalan sebagai upaya untuk mengukur perkembangan program.
 - d. Program sudah selesai dilaksanakan dan pertanggungjawaban selesai di susun.
3. Evaluasi sumatif adalah evaluasi yang dilakukan saat program....!
 - a. Sudah selesai.
 - b. Sedang berjalan.
 - c. Belum berjalan.
 - d. Masih direncanakan.
4. Proses evaluasi berbasis tujuan dilaksanakan dengan tahapan sebagai berikut!
 - a. Mengidentifikasi tujuan, Merumuskan tujuan menjadi indikator, Mengembangkan metode dan instrument, Memastikan program telah mencapai tujuan, Menjaring dan menganalisis data program, Kesimpulan.
 - b. Meneliti dan mengidentifikasi pengaruh program yang relevan tanpa merujuk goal dan objektif yang ada di rencana program.

- c. Menentukan pengaruh apa yang muncul dan secara logis disebabkan oleh program dan intervensi program.
 - d. Mengidentifikasi apa yang terjadi tanpa merujuk pada tujuan dan sasaran program.
5. Kegiatan evaluasi yang dilakukan oleh evaluator tanpa mempunyai pengetahuan atau referensi dari tujuan serta pengaruh yang diharapkan oleh perancang program disebut...!
- a. *Goal-Based evaluation*
 - b. *Sumatif Evaluation*
 - c. *Responsive Evaluation*
 - d. *Goal Free Evaluation*

Kunci Jawaban dan Pedoman Penskoran

1. A
2. C
3. A
4. A
5. D

Bobot

Latihan soal essay 75 % dengan bobot per soal 15.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{50}{75} \times 100 = \mathbf{66,67}$$

Tes Formatif Pilihan Berganda 25% dengan bobot per soal 5.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{20}{25} \times 100 = \mathbf{80}$$

$$66,67 \times 75\% = \mathbf{50}$$

$$80 \times 25\% = \mathbf{20}$$

$$50 + 20 = \mathbf{70}$$

$$\text{Nilai akhir} = \mathbf{70}$$

MODUL 10

MODEL-MODEL EVALUASI CIPP, EVALUASI ADVERSARY, EVALUASI MODEL KETIMPANGAN ATAU DISCREPANCY EVALUATION MODEL, EVALUASI SISTEM ANALISIS

Metode Pembelajaran	Estimasi waktu	Capaian Pembelajaran
-Presentasi virtual -Diskusi Virtual interaktif -Tanya jawab	120 Menit X 1	-Mahasiswa mampu memahami model-model evaluasi CIPP, Evaluasi <i>adversary</i> , Evaluasi model ketimpangan atau <i>Discrepancy evaluation Model</i> , evaluasi system analisis.

Materi 10. Model-model Evaluasi CIPP, Evaluasi *adversary*, Evaluasi model ketimpangan atau *Discrepancy evaluation Model*, evaluasi system analisis.

A. Evaluasi CIPP

Model evaluasi CIPP dianggap kerangka yang paling komprehensif dibanding dengan model lain. Salah satu alasannya adalah didalam CIPP memuat aspek *Context, Input, Proses, dan Product* sebagai objek evaluasi. Model CIPP dikembangkan oleh Daniel L. Stufflebeam pada tahun 1960an, dengan tujuan untuk membantu memperbaiki dan pertanggungjawaban proyek-proyek sekolah negeri yang dibiayai oleh pemerintah Federal Amerika Serikat, terutama sebagai panduan untuk memperbaiki pengajaran dan pembelajaran disekolah. Semakin lama model CIPP semakin berkembang dan digunakan di seluruh Amerika Serikat dan beberapa negara lain dan diterapkan di bidang-bidang lain. Model CIPP berangkat dari fisosofi *learning by doing*- yakni suatu usaha berkelanjutan untuk mengidentifikasi dan memperbaiki kesalahan-kesalahan yang terjadi dalam kegiatan evaluasi, untuk menginventarisir dan menguji prosedur-prosedur baru yang dibutuhkan, dan untuk memelihara dan menggabungkan praktik-praktik yang efektif.

Stufflebeam mendefinisikan evaluasi sebagai *“Our operational definition of evaluation states that evaluation is the systematic process of delineating, obtaining, reporting, and applying descriptive and judgmental information about some object’s merit, worth, probity, feasibility, safety, significance, and/or equity.”*¹⁰¹ Definisi operasional yang dikemukakan oleh Stufflebeam tentang evaluasi adalah proses yang sistematis dalam menggambarkan, mengumpulkan, melaporkan, dan mendeskripsikan dan proses menentukan informasi mengenai nilai, manfaat, kejujuran, kemungkinan untuk dilakukan, keamanan, signifikansi, dan kewajaran/keadilan.



Daniel L. Stufflebeam

Stufflebeam, Daniel LeRoy was born on September 19, 1936 in Waverly, Iowa, United States. Son of LeRoy and Melva Stufflebeam.

Education

Bachelor, State University Iowa, 1958. Master of Science, Purdue University, 1962. Doctor of Philosophy, Purdue University, 1964.

Daniel L. Stufflebeam, PhD, was Distinguished University Professor Emeritus at Western Michigan University (WMU) until his death in 2017. He founded and directed The Evaluation Center at The Ohio State University and moved it to WMU, where he directed it until 2002. He also designed the WMU Interdisciplinary PhD Program in Evaluation. The developer of the *CIPP Evaluation Model*, Dr. Stufflebeam led development of the General Educational Development Tests (GED) and other standardized tests; founded the Joint Committee on Standards for Educational Evaluation and was principal author of the Committee’s original standards; and designed and directed the national Center for Research on Educational Accountability and Teacher Evaluation. He led development of many evaluation systems and lectured, advised, and conducted evaluations in more than 20 countries. The author of 21 books and about 100 journal articles and book chapters (appearing in eight languages), Dr. Stufflebeam received the Paul F. Lazarsfeld Evaluation Theory Award from the American Evaluation Association, among numerous other honors.

Profil 6: Daniel L. Stufflebeam

Dua kata kunci merit dan worth diperkuat oleh definisi evaluasi yang disampaikan Scriven (1967b) *“defined evaluation as a method of determining the merit or worth of an evaluand (the project, program, or other entity being evaluated), arguing that the placement of value on*

¹⁰¹ Daniel L. Stufflebeam and Chris L. S. Coryn. *Evaluation Theory, Models, and Applications*, (San Fransisco: Jossey-Bass A Wiley Brand), p.14.

*something should be appraised based on these two dimensions.”*¹⁰² Evaluasi sebagai metode untuk menentukan nilai atau manfaat evaluand (proyek, program, atau hal lain yang dievaluasi) alasannya bahwa penempatan nilai sesuatu itu seharusnya di ukur berdasarkan pada dua dimensi ini (*merit & worth*).

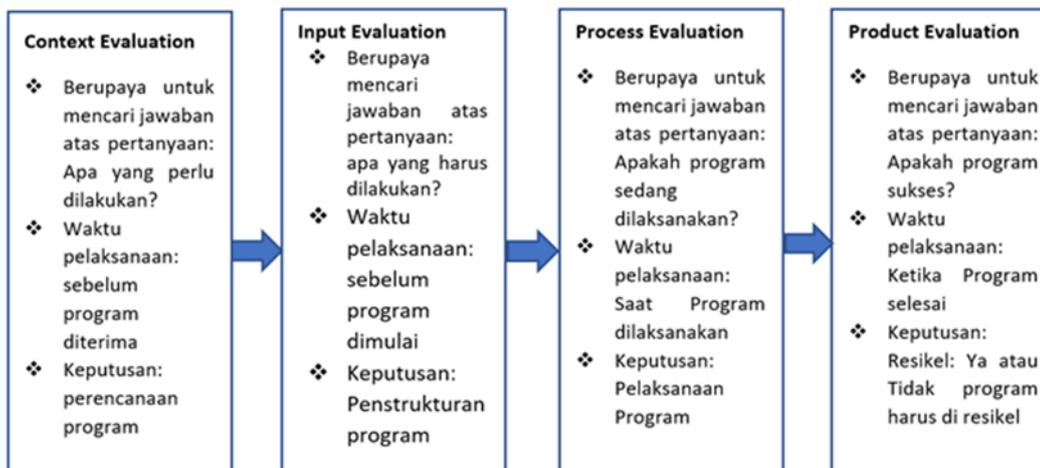
Merit adalah kualitas sesuatu yang bersifat absolut maupun yang relative, baik secara instrinsik atau yang berkaitan dengan kriteria tertentu. Sebagai contoh, pelayanan pendukung bagi warga penyandang disabilitas yang memenuhi kebutuhan pribadinya merupakan karakteristik intrinsik yang memiliki nilai dalam berbagai kontek. **Worth** adalah hasil evaluasi dan merujuk pada nilai objek evaluasi dalam kontek tertentu. Worth dan merit tidak saling ketergantungan satu sama lain. Objek evaluasi (misalnya layanan tambahan) bisa saja memiliki nilai intrinsic tetapi bisa saja hanya memiliki manfaat sedikit (disekolah misalnya tidak ada siswa penyandang disabilitas, tetapi ada fasilitas penunjang untuk siswa berkebutuhan khusus).

Model CIPP memiliki empat jenis evaluasi yaitu: Evaluasi kontek (*context*), Evaluasi Masukan (*input*), Evaluasi Proses (*process*), dan Evaluasi Produk (*product*). Daniel Stufflebeam, “*In context evaluations, evaluators assess needs, problems, assets, and opportunities, plus relevant contextual conditions and dynamics. Decision makers use context evaluations to define goals and set priorities and to make sure program goals are targeted to address significant, assessed needs and problems.*”¹⁰³ Dalam evaluasi kontek, evaluator menilai kebutuhan, masalah, asset, dan kemungkinan-kemungkinan, ditambah kondisi kontekstual yang relevan dan dinamis. Pengambil keputusan menggunakan evaluasi kontek untuk mendefinisikan tujuan dan prioritas-prioritas dan untuk memastikan tujuan program tercapai secara signifikan, penilaian kebutuhan dan masalah. Mengutip dari Wirawan dalam buku *Evaluasi; Teori, Model, Metodologi, Standar,*

¹⁰² Donna M. Mertens and Amy T. Wilson. *Program Evaluation Theory and Practice: A Comprehensive Guide*, (New York: The Guilford Press A Division of Guilford Publications, Inc., 2019), P. 6.

¹⁰³ Daniel L. Stufflebeam and Chris L. S. Coryn. *Evaluation Theory, Models, and Applications*, (San Fransisco: Jossey-Bass A Wiley Brand), P. 312.

Aplikasi dan Profesi (2016), pendekatan model CIPP dapat di ilustrasikan sebagai berikut:



Gambar 10: Model Evaluasi CIPP

Stufflebeam dan Coryn menyampaikan bahwa tujuan utama evaluasi kontek adalah:

- *Set boundaries around and describe the setting for the intended program or other improvement effort.* Menetapkan Batasan-batasan yang ada disekelilingnya dan mendeskripsikan posisi program atau untuk upaya perbaikan yang lain.
- *Identify intended beneficiaries and assess their needs.* Mengidentifikasi penerima program dan menilai kebutuhannya.
- *Identify problems or barriers to meeting the assessed needs.* Mengidentifikasi masalah-masalah atau hambatan untuk memenuhi kebutuhan yang dinilai.
- *Identify relevant, accessible assets and funding opportunities that could be used to address the targeted needs.* Mengidentifikasi hal-hal yang relevan, asset yang dapat diakses dan peluang pendanaan yang dapat digunakan untuk memenuhi kebutuhan yang ditargetkan.
- *Provide a basis for setting improvement-oriented goals.* Menyediakan dasar untuk menetapkan tujuan berorientasi perbaikan.

- *Assess the clarity and appropriateness of improvement-oriented goals.* Menilai kejelasan dan kesesuaian tujuan yang berorientasi perbaikan.
- *Provide a basis for judging outcomes of the subject program.*¹⁰⁴
Menyediakan landasan untuk menilai hasil pelaksana program

*In input evaluations, evaluators assist with program planning by identifying and assessing alternative approaches and subsequently assessing procedural plans, staffing provisions, and budgets for their feasibility and potential cost-effectiveness in regard to meeting targeted needs and achieving goals.*¹⁰⁵ Dalam evaluasi *input*, evaluator membantu perencanaan program dengan mengidentifikasi dan menilai pendekatan alternative dan kemudian menilai prosedur perencanaan, menyusun syarat kepegawaian, dan anggaran untuk hal-hal yang mungkin terjadi dan potensi efektifitas biaya dalam rangka memenuhi target yang dibutuhkan dan mencapai tujuan.

Evaluator melakukan evaluasi input dalam beberapa tahap, yang terjadi secara tidak berurutan. Evaluator bisa saja memulai kegiatan dengan mereview pernyataan tentang hal-hal yang akan dilakukan dalam upaya memenuhi kebutuhan dan tujuan. Proses yang dimaksud termasuk sejumlah komponen-komponen sebagai berikut:

- *Reviewing relevant literature.* Mereview literatur yang relevan.
- *Visiting exemplary programs.* Mengunjungi program percontohan.
- *Consulting experts and government representatives.* Berkonsultasi dengan para ahli dan perwakilan pemerintah.
- *Querying pertinent information services (especially those on the World Wide Web).* Layanan informasi yang berhubungan dengan pertanyaan (khususnya mengenai jaringan dunia maya).
- *Reviewing a pertinent article in Consumer Reports or a similar publication that critically reviews available products and services.* Mereview artike-artikel yang berhubungan dengan laporan

¹⁰⁴ Daniel L. Stufflebeam and Chris L. S. Coryn. *Evaluation Theory, Models, and Applications*, (San Fransisco: Jossey-Bass A Wiley Brand), P. 320.

¹⁰⁵ *Ibid.*, P. 312.

pelanggan atau sejenis publikasi yang secara kritis mereview produk dan layanan yang ada.

- *Inviting proposals from involved staff.*¹⁰⁶ Mengundang proposal dari para staff yang terlibat.

Evaluator perlu mengorganisasikan informasi dalam ruang perencanaan yang khusus, melibatkan kelompok studi khusus untuk menginvestigasi informasi atau melakukan seminar perencanaan khusus untuk menganalisis materi. Evaluator menggunakan informasi untuk menilai apakah terdapat solusi strategis yang bisa diterima. Evaluator akan menghitung pendekatan yang menjanjikan/relevan dengan kriteria seperti berikut ini:

- *Responsiveness to assessed needs of targeted beneficiaries.* Responsive untuk menilai kebutuhan yang ditargetkan.
- *Responsiveness to targeted problems in the organization.* Responsive terhadap persoalan-persoalan organisasi.
- *Use of special funding programs or other relevant opportunities.* Menggunakan pendanaan program khusus atau peluang lain yang relevan.
- *Potential effectiveness.* Potensi efektifitas.
- *Cost.* Biaya.
- *Political viability.* Kelangsungan politis.
- *Administrative feasibility.* Mengerjakan administratif.
- *Potential for important impacts outside the local area.* Potensi pengaruh dari luar yang cukup kuat bagi area didalam.

Setelah evaluasi *input* diselesaikan, maka tahap selanjutnya dalam pendekatan CIPP ini yang akan dilakukan oleh evaluator adalah mengevaluasi proses. Dalam evaluasi proses, Stufflebean dan Coryn mengatakan bahwa *“In process evaluations, evaluators monitor, document, assess, and report on the implementation of program plans. Such evaluators provide feedback throughout a program’s implementation and later report on the extent to which the program was carried out as intended and required. Program staff use periodic process*

¹⁰⁶ Daniel L. Stufflebeam and Chris L. S. Coryn. *Evaluation Theory, Models, and Applications*, (San Fransisco: Jossey-Bass A Wiley Brand), P. 324.

*evaluation reports to take stock of their progress, identify implementation issues, and adjust their plans and performance to ensure program quality and on-time delivery of services.*¹⁰⁷

Dalam evaluasi proses, evaluator memonitor, mendokumentasikan, menilai, dan melaporkan berdasarkan implementasi rencana program. Evaluator memberikan umpan balik menyeluruh terkait implementasi program dan laporan berikutnya berdasarkan perkembangan program yang dilaksanakan sesuai yang dimaksud dan dibutuhkan. Staff program menggunakan laporan evaluasi proses secara periodic untuk menampung perkembangannya, mengidentifikasi isu implementasi, dan menyelaraskan rencana mereka dan kinerjanya guna memastikan kualitas program dan layanan pengiriman yang tepat waktu.

Selanjutnya adalah evaluasi produk yang di arahkan untuk mencari jawaban atas pertanyaan: apakah programnya berhasil? Stufflebeam dan Coryn menyatakan bahwa;

In product evaluations, evaluators identify and assess costs and outcomes—intended and unintended, short term and long term. They provide feedback during a program’s implementation on the extent to which program goals are being addressed and achieved. At the program’s end, product evaluation helps identify and assess the program’s full range of accomplishments. Program staff use interim product evaluation feedback to maintain focus on achieving important outcomes and to identify and address deficiencies in the program’s progress toward achieving important outcomes.¹⁰⁸

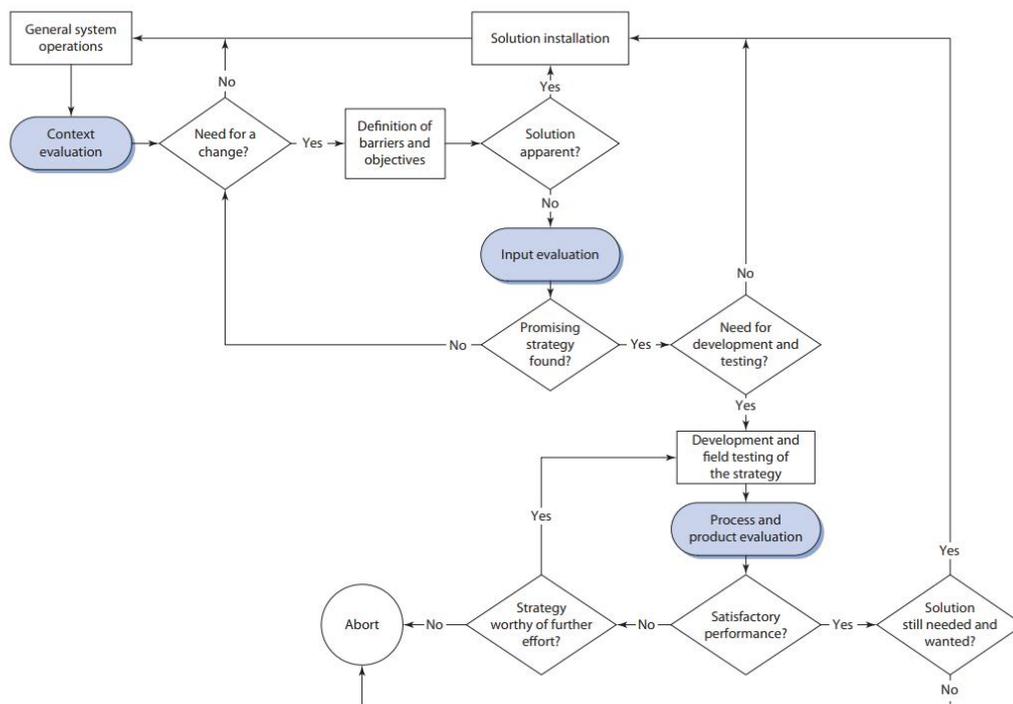
Pada tahap evaluasi produk, evaluator mengidentifikasi dan menilai biaya dan hasilnya-baik yang sesuai harapan maupun yang tidak sesuai harapan, dalam jangka pendek dan jangka Panjang. Evaluator memberikan balikan selama implementasi program berdasarkan perkembangan tujuan program yang dicapai. Pada akhir program,

¹⁰⁷ Daniel L. Stufflebeam and Chris L. S. Coryn. *Evaluation Theory, Models, and Applications*, (San Fransisco: Jossey-Bass A Wiley Brand), P. 312.

¹⁰⁸ Daniel L. Stufflebeam and Chris L. S. Coryn. *Evaluation Theory, Models, and Applications*, (San Fransisco: Jossey-Bass A Wiley Brand), P. 313.

evaluasi produk membantu mengidentifikasi dan menilai tingkat pencapaian program secara utuh. Staff program menggunakan balikan evaluasi produk sementara untuk menjaga focus pada pencapaian hasil yang signifikan dan untuk mengidentifikasi serta mengetahui kekurangan-kekurangan perkembangan program untuk mencapai hasil yang diharapkan.

Menurut Stufflebeam Model Evaluasi CIPP bersifat linear, yang berarti bahwa evaluasi *Input* harus didahului dengan evaluasi *Context*, Evaluasi *Process* didahului oleh Evaluasi *Input*, dan Evaluasi *Product* dilakukan setelah Evaluasi *Process*.



Gambar 11: Flowchart of a CIPP Evaluation in Fostering and Assessing System Improvement

Stufflebeam mengatakan bahwa Model CIPP merupakan evaluasi dengan pendekatan sistem sosial. Sistem sosial adalah serangkaian aktivitas yang saling terkait yang berfungsi secara Bersama untuk mencapai misi dan tujuan yang telah ditentukan dalam konteks tertentu. Dalam perspektif ini, evaluasi model CIPP cocok untuk mengembangkan dan membantu pencapaian tujuan dan perbaikan yang berkelanjutan.

B. Evaluasi Adversary

Evaluasi *adversary* merupakan model evaluasi yang menyerupai proses pengadilan. Dalam konteks hukum, cara kerja *adversary* dikembangkan sebagai proses mendengarkan untuk menentukan nilai suatu kasus dimana pihak oposan ikut terlibat. Dalam evaluasi pendidikan, model *adversary* dikembangkan pada awal 1970an. Menurut Thomas Owens salah satu pengembang evaluasi *adversary*, pendekatan *adversary* memiliki metodologi yang dibutuhkan untuk menyediakan informasi bagi pengambil keputusan untuk merencanakan proyek dan untuk menetapkan tujuan mereka dan pengaruh yang tidak diinginkan yang telah diimplementasikan.

Tujuan utama model evaluasi *Adversary* adalah untuk mengurangi potensi bias dengan membentuk 2 evaluator yang berbeda. Kedua evaluator yang dimaksud adalah evaluator pro dan evaluator kontra. Kedua evaluator sepakat mengenai isu yang akan diselesaikan dan menyiapkan pangkalan data umum mengenai isu tersebut. Kemudian melakukan pengumpulan data khusus sesuai dengan tugas masing-masing. Selanjutnya, kedua evaluator mendiskusikan data umum dan data khusus, terutama hasil wawancara untuk mendukung argumentasi masing-masing.¹⁰⁹

Stufflebeam & Shinkfield mengutip Owens mengenai karakteristik evaluasi *adversary* yang teridentifikasi sesuai/cocok untuk pengambilan keputusan dalam Pendidikan. Karakteristik yang dimaksud adalah sebagai berikut.

1. Aturan-aturan yang dibuat untuk menangani cara kerja *adversary* harus benar-benar fleksibel.
2. Aturan yang rumit mengenai bukti diganti dengan evaluasi bebas bukti yang semata-mata hanya berdasarkan pada apakah buktinya dianggap relevan oleh petugas pemeriksaan.

¹⁰⁹ Wirawan. *Evaluasi: Teori, Model, Metodologi, Standar, Aplikasi dan Profesi*, (Jakarta: Rajagrafindo Persada, 2016), h. 149.

3. Kedua belah pihak diwajibkan sebelum ujicoba untuk menyampaikan kepada petugas pemeriksaan mengenai fakta-fakta yang relevan, alat bukti, dan nama saksi.
4. Salinan pembiayaan dilengkapi ke petugas pemeriksa dan terdakwa sebelum diuji, dan tergugat memiliki opsi menerima biaya-biaya tertentu yang terdahulu dan memberi tantangan kepada yang lain.
5. Para saksi diberikan izin untuk mengujinya secara bebas dan memeriksa secara silang.
6. Para ahli dipanggil untuk mengujinya, bahkan sebelum ujicoba.
7. Konferensi sebelum ujicoba yang dilakukan oleh petugas pemeriksaan dengan kedua belah pihak cenderung untuk membuat ujicoba kurang perjuangan dalam ketajaman berpikir dan lebih kepada pencarian untuk fakta-fakta yang relevan.
8. Lebih jauh bahwa kedua belah pihak terlibat, kelompok menarik yang lainnya bisa saja diizinkan untuk berpartisipasi.

Owens dalam bukunya Stufflebeam & Shinkfield, *Systematic Evaluation A Self-Instructional Guide to Theory and Practice* (p.269-270) menyarankan tujuh cara dalam menggunakan evaluasi adversary untuk tujuan evaluasi Pendidikan. Secara garis besar penulis mencoba menerjemahkan secara bebas gambaran ketujuh saran yang disampaikan Owens.

Exploring the values of a new or existing curriculum. Evaluasi perbandingan sering gagal menentukan dan menginvestigasi teori, rasional, dan asumsi yang ada dibalik kurikulum. Debat terbuka mengenai relevansi kurikulum tertentu bagi anak-anak sekarang dan kebutuhan dimasa depan, kebutuhan social, atau yang berkaitan dengan integritas mungkin menjadi langkah yang sangat berguna sebelum menentukan apakah bijak atau tidak untuk mengembangkan atau menerapkan kurikulum tersebut.

Estimating the congruence between an innovation and the existing system. Inovasi dapat diperkenalkan tanpa referensi kepada warga sekolah secara lebih luas. Pada pemeriksaan pihak kontra, bukti-bukti mungkin dibuat seperti pengaruh yang mengganggu inovasi terhadap

system secara keseluruhan, dengan demikian dapat menyingkap potensi penghalang bagi keberhasilan program. Selanjutnya, keputusan dapat dibuat untuk mencegah atau mengatasi masalah yang ada.

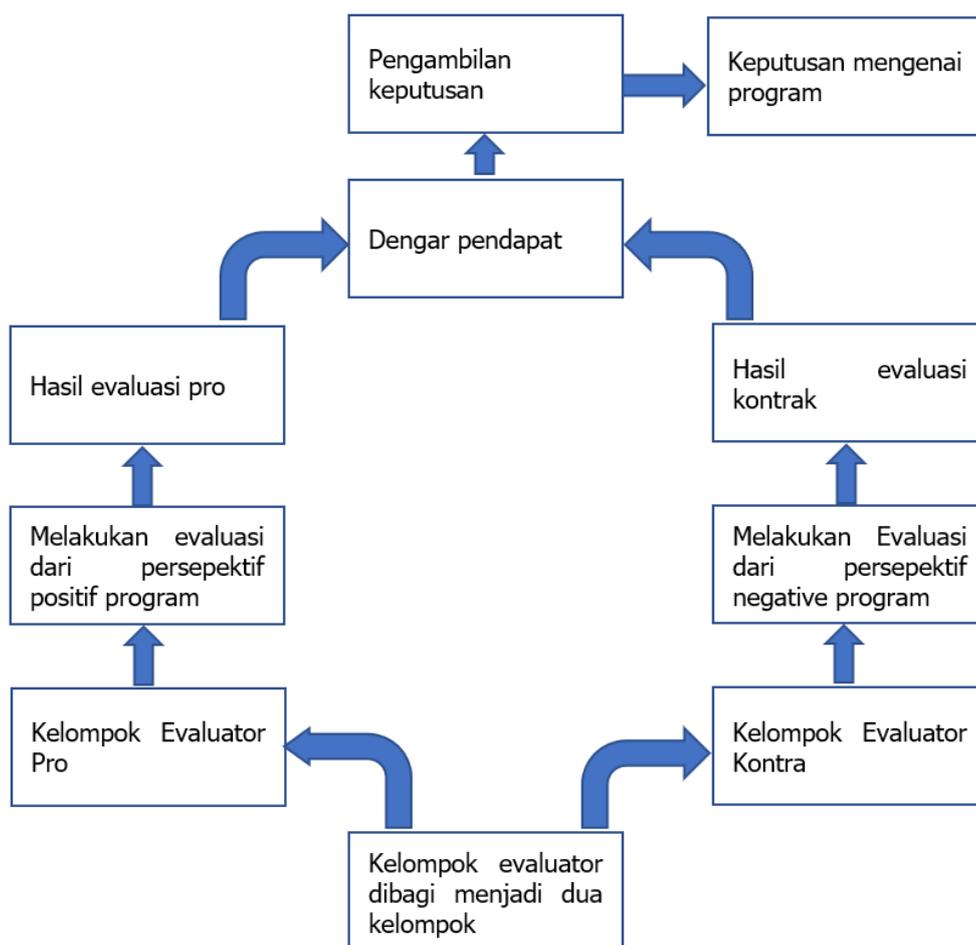
Revealing the different interpretations of the same data by various representatives. Owens mengemukakan bahwa jurinya bisa terdiri dari perwakilan siswa, guru, staf administrasi, orang tua, dan komunitas sekolah. Dengan berbagai latar belakang juri, dapat menginterpretasikan data yang sama secara berbeda; perbedaan ini, dikomunikasikan kepada pengambil keputusan, dapat memperbaiki komunikasi mereka dengan kelompoknya. Pengambilan keputusan secara Bersama-sama juga memungkinkan untuk dilakukan.

Informing teachers, supervisors, and administrators. Sebagai observer atau peserta, pendidik akan mendapatkan keuntungan pengetahuan baru tentang program yang menjadi objek pemeriksaan.

Resolving disputes about performance contracts. Perdebatan mengenai keberhasilan penyelesaian kontrak kinerja tetap akan muncul. Pemeriksaan adversary akan membantu menyelesaikan perselisihan berdasarkan testimoni ahli yang berasal dari berbagai sumber.

Arriving at the decision to be implemented. Pengambil keputusan bidang Pendidikan bisa menjadi staf pemeriksaan, dimana kasus putusannya adalah keputusan. Atau ia menginginkan untuk melibatkan anggota juri, seperti guru dan perwakilan siswa dalam pengambilan keputusan. Bisa juga pengambil keputusan mempertimbangkan putusan juri sebagai salah satu bagian informasi yang akan digunakan bersama dengan yang lainnya, seperti alasan-alasan pokok dalam pemeriksaan dan data yang terkumpul melalui pengujian.

Owens menganggap bahwa cara kerja evaluasi adversary menjadi sangat relevan terhadap implementasi evaluasi yang dikembangkan oleh Daniel Stufflebeam, Robert Stake ("countenance" model, 1967), dan Malcolm Provus. Secara umum model evaluasi Adversari dilakukan melalui aktifitas sebagai berikut.



Gambar 12: Proses Model Evaluasi Adversari

1. *Membentuk dua atau lebih Tim Evaluator yang independent.* Pada tahap ini evaluator dibagi menjadi dua kelompok yaitu kelompok yang pro dan kelompok kontra. Jika diperlukan dapat dibentuk kelompok netral. Masing-masing kelompok menyusun proposal evaluasi berdasarkan perspektif kelompok masing-masing. Kelompok yang pro menyusun rencana evaluasi dengan asumsi bahwa program berhasil mencapai tujuan dan mampu mengubah keadaan social. Sedangkan kelompok kontra menyusun rencana evaluasi dengan asumsi bahwa program gagal mencapai tujuan dan tidak mampu melakukan perubahan sosial.
2. *Melakukan Evaluasi.* Kelompok pro melakukan evaluasi dari persepektif positif atau keberhasilan program mencapai tujuannya dan memiliki dampak positif terhadap sasaran program atau masyarakat. Data-data yang dikumpulkan merupakan data

dan informasi yang mendukung argumentasi keberhasilan program. Disisi lain, kelompok kontra mengevaluasi program dari sudut pandang negative yang menunjukkan kegagalan program dalam mengubah masyarakat atau kontribusi program yang dijalankan minimal atau bahkan tidak berdampak.

3. *Merumuskan hasil evaluasi.* Kedua kelompok merumuskan hasil hasil evaluasinya dari perspektif masing-masing. Meskipun rumusan hasil evaluasi dibuat berdasarkan sudut pandang yang berbeda, tetapi kedua kelompok melakukannya dengan penuh kehati-hatian dan teliti berdasarkan data dan informasi yang diperoleh dilapangan. Dengan keshahihan sumber data dan informasi, maka rumusan hasil evaluasi dapat dipertanggungjawabkan. Meskipun satu objek evaluasi dengan dua versi rumusan hasil yang berbeda.
4. *Dengar pendapat.* Kegiatan dengar pendapat dilakukan oleh kedua kelompok pro dan kontra Bersama-sama dengan pengambil keputusan. Dalam kegiatan ini disampaikan argumen-argumen dari masing-masing kelompok dengan bukti dan data yang mendukung masing-masing kesimpulan yang dibuat. Data dan informasi yang disampaikan setiap kelompok dimaksudkan untuk memengaruhi pengambil keputusan.
5. *Keputusan mengenai program.* Dengan memperhatikan nilai manfaat, mudharat, keunggulan dan kelemahan, serta potensi dari masing-masing masukan yang disampaikan oleh kelompok pro dan kelompok kontra, pengambil keputusan harus menentukan sikap. Pengambil keputusan dengan segala pertimbangan berdasarkan data, informasi, dan bukti-bukti yang ada, akan memutuskan mengenai program.¹¹⁰

Memerhatikan alur evaluasi adversary yang ada, menunjukkan adanya potensi pembiayaan yang cukup mahal karena harus ada dua kelompok evaluator yang pro dan kontra, dan jika ingin hasil yang lebih baik maka dapat ditambah satu kelompok lagi yaitu kelompok netral.

¹¹⁰ Wirawan. *Evaluasi: teori, Model, Metodologi, Standar, Aplikasi dan Profesi*, (Jakarta: Rajagrafindo Persada, 2016), hh. 149-150.

Jika menggunakan pendekatan ini, para evaluator dari masing-masing kelompok juga diharapkan memiliki kompetensi dan kapasitas yang setara sehingga tidak ada dominasi dari salah satu kelompok. Potensi ini bisa berdampak kurang baik pada hasil evaluasi yang akan memengaruhi keputusan yang akan diambil. Meskipun disisi lain, pendekatan adversary memberikan peluang bagi pengambil keputusan karena adanya alternative-alternatif yang menjadi pilihan dalam membuat keputusan mengenai program.

C. Evaluasi Ketimpangan

Evaluasi model ketimpangan atau *Discrepancy evaluation Model* (DEM) dikembangkan oleh Malcolm M. Provus pada tahun 1966, dengan dua tujuan; memberi informasi terhadap penilaian program dan untuk pengembangan program. Sejak saat itu, DEM terus direvisi secara periodik sesuai dengan perkembangan dan pengalaman para praktisinya. Model yang dirancang khusus untuk mengevaluasi program Pendidikan khususnya di *Pittsburgh Public School*. Ketika kita akan mengevaluasi benda jadi misalnya (bisa orang, program, atau kendaraan) maka harus membandingkannya dengan standar. Dengan standar kita mempunyai alat ukur, deskripsi, atau representasi kualitas atau karakteristik yang seharusnya dimiliki objek evaluasi. Deskripsi tentang bagaimana sesuatu itu disebut standar. Pada saat kita memiliki gambaran yang jelas tentang bagaimana objek itu seharusnya, kita bisa melanjutkan untuk menemukan apakah objek tersebut benar-benar sudah sesuai dengan yang seharusnya. Saat kita terlibat bdalam kegiatan untuk menemukan karateristik aslinya dari suatu objek yang dievaluasi, kita berbicara mengenai ukuran kinerja. Dengan demikian, evaluasi ketimpangan adalah persoalan membandingkan standar dengan kinerja. Evaluasi model ketimpangan adalah upaya membandingkan mengenai “apa” (kinerja) yang sesungguhnya dengan apa yang seharusnya atau “standar”. Jika ditemukan perbedaan antara standar dengan kinerja yang ada, perbedaan ini disebut sebagai ketimpangan.

Steinmetz mengatakan “*There is another term involved in the comparison between S and P. We say that the comparison yields Discrepancy (D) information, and thus we can speak of evaluation as being a matter of making judgments about the worth or adequacy of an object based upon D information between S and P.*”¹¹¹ Terdapat istilah lain dalam membandingkan antara standar (S) dan *performance* (P). Bahwa perbandingan akan menghasilkan informasi *discrepancy* (D), dengan demikian kita dapat mengatakan bahwa evaluasi adalah persoalan membuat keputusan tentang nilai atau kesesuaian objek evaluasi berdasarkan informasi antara standar dan kinerja (*performance*).

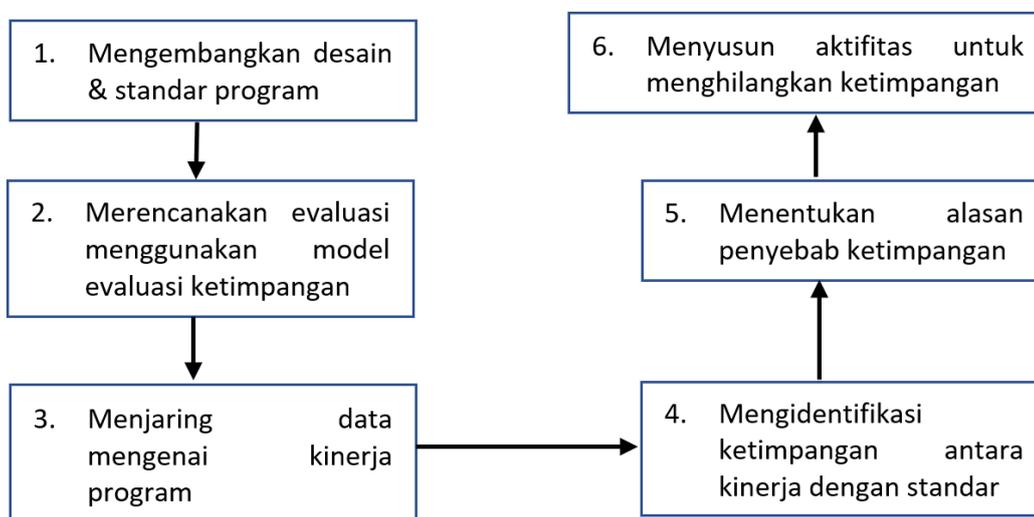
Mengutip pernyataan Wirawan bahwa Provus mendefinisikan evaluasi sebagai proses kesesuaian program terhadap standar-standar program kemudian menentukan apakah suatu ketimpangan (*discrepancy*) terjadi antara sejumlah aspek dari program dengan standar-standar yang mengatur aspek-aspek program tersebut, dan memakai informasi mengenai ketimpangan tersebut untuk mendefinisikan kelemahan-kelemahan program.¹¹² Model evaluasi ketimpangan memerlukan enam langkah untuk melaksanakannya;

1. Mengembangkan suatu desain dan standar-standar yang menspesifikasi karakteristik implementasi ideal dari objek evaluasi: kebijakan, program, atau proyek.
2. Merencanakan evaluasi menggunakan model evaluasi diskrepansi. Menentukan informasi yang diperlukan untuk membandingkan implementasi yang sesungguhnya dengan standar yang mendefinisikan kinerja objek evaluasi.
3. Menjaring kinerja objek evaluasi yang meliputi pelaksanaan program, hasil-hasil kuantitatif dan kualitatif.
4. Mengidentifikasi ketimpangan-ketimpangan antara standar dengan pelaksanaan dan hasil pelaksanaan objek evaluasi yang sesungguhnya dan menentukan rasio ketimpangan.

¹¹¹ Andrés Steinmetz. *Evaluation Models Viewpoints on Educational and Human Services Evaluation*, edited by Daniel L. Stufflebeam, George F. Madaus, dan Thomas Kellaghan, (New York: Kluwer Academic Publishers, 2002), P.128.

¹¹² Wirawan. *Evaluasi: Teori, Model, Metodologi, Standar, Aplikasi dan Profesi*, (Jakarta: Rajagrafindo Persada, 2016), h. 151.

5. Menentukan penyebab ketimpangan dengan membandingkan antara standar kinerja dengan kinerja objek evaluasi.
6. Menghilangkan ketimpangan dengan membuat perubahan-perubahan terhadap implementasi objek evaluasi.¹¹³



Gambar 13: Proses Model Evaluasi Ketimpangan

Model evaluasi ketimpangan mempunyai keunggulan dan kelemahan seperti berikut.

Keunggulan:

1. Model evaluasi ketimpangan sederhana dan mudah dilaksanakan.
2. Model ketimpangan memperkenalkan konsep manajemen informasi, proses ruangan kelas dan pentingnya membangun pangkalan data yang menghubungkan karakteristik siswa dan kinerja siswa.
3. Model ini mengembangkan hubungan yang terus-menerus antara staf evaluator permanen dan perencana dan pengembangan program.
4. Model ini menggunakan evaluasi formatif untuk merevisi dan mengoreksi program untuk mengarahkan kembali pada awal pengembangan dan instalasi program.

¹¹³ Wirawan. *Evaluasi: Teori, Model, Metodologi, Standar, Aplikasi dan Profesi*, (Jakarta: Rajagrafindo Persada, 2016), hh. 151-152.

Kelemahan Model ketimpangan adalah;

1. Untuk melaksanakan model ini memerlukan waktu yang Panjang karena pertanyaan yang harus dijawab secara berurutan untuk sampai kepada keputusan. Salah satu tujuan dari pertanyaan adalah mengikutsertakan staf program dalam proses evaluasi.
2. Taksonomi yang diidentifikasi sebagai suatu standar untuk definisi program merupakan daftar kategori. Standar sesungguhnya, seperti jumlah dan waktu pelatihan para guru harus dikembangkan di tempat lain.¹¹⁴

D. Evaluasi Sistem Analisis

Model evaluasi system analisis atau disebut juga dengan *Management Evaluation Model* merupakan evaluasi yang berbasis pada teori system. Pada awalnya teori system diungkapkan oleh Karl Luwig von Bertalanffy (1951). Sistem adalah satu kesatuan yang utuh. Diantara bagian-bagian yang ada dalam sistem memiliki keterkaitan satu sama lain dan saling ketergantungan. Setiap system diikat oleh tempat dan waktu, dipengaruhi oleh lingkungannya, dimaknai oleh struktur dan tujuannya, serta diwujudkan sesuai fungsinya masing-masing. Tujuan teori system adalah memperagakan system yang dinamis, hambatan-hambatan, kondisi, dan untuk menguraikan prinsip-prinsip seperti tujuan, ukuran, metode, dan alat yang dapat diterapkan pada sistem lain dilevel yang sama. Teori umum system adalah tentang mengembangkan konsep dan prinsip-prinsip secara luas yang dapat digunakan sebagai penyeimbang bagi prinsip-prinsip khusus dan konsep pengetahuan tertentu.

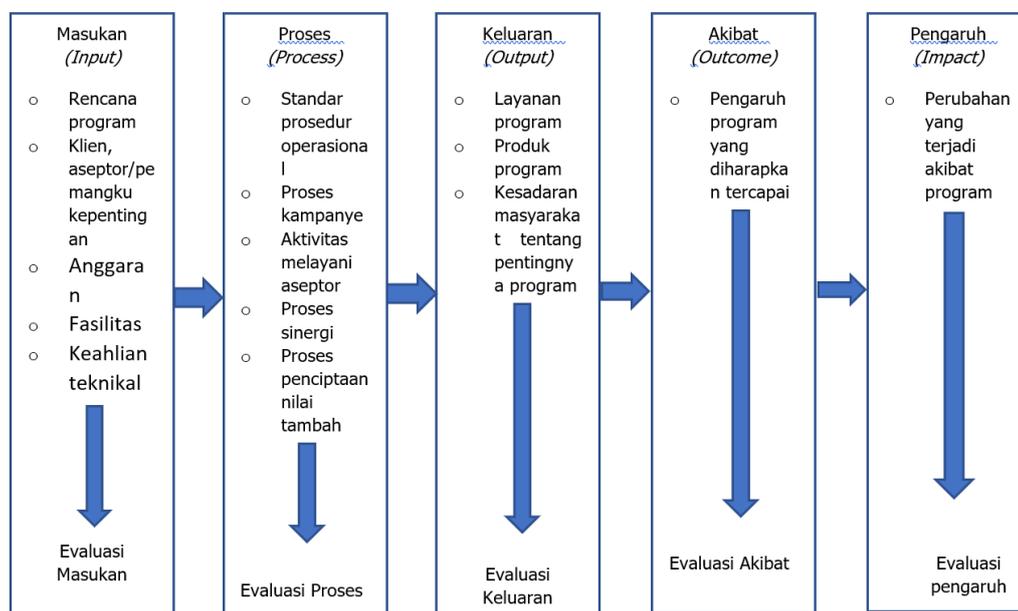
Menurut Wirawan suatu sistem program terdiri dari sejumlah unit, dimensi atau subsistem program yang masing-masing mempunyai fungsi tertentu. Setiap unit diikat dengan ikatan sinergi yang mengikat setiap unit untuk menjadi kesatuan program. Fungsi dari ikatan sinergi adalah:

¹¹⁴ Wirawan. *Evaluasi: Teori, Model, Metodologi, Standar, Aplikasi dan Profesi*, (Jakarta: Rajagrafindo Persada, 2016), hh. 153.

1. Mengikat semua unit menjadi satu kesatuan sistem agar aktivitas-aktivitasnya dapat bergerak secara Bersama-sama ke arah tujuan program.
2. Menciptakan sinergi positif, yaitu produksi semua unit program yang bekerja dalam system program lebih besar daripada produksi masing-masing unit dijumlahkan-jika tidak bekerja dalam kesatuan system program. Misalnya, jika unit 1, unit 2, unit 3, dan unit 4 bekerja sendiri-sendiri produktifitas masing-masing 25, 40, 15, dan 50 = 130. Sinergi dari keempat unit bisa lebih besar dari 130 bahkan bisa mencapai 150.¹¹⁵

Dalam model evaluasi system analisis terdapat empat jenis evaluasi yaitu: evaluasi masukan (*input evaluation*), evaluasi proses (*process evaluation*), evaluasi keluaran (*output evaluation*), evaluasi akibat (*outcome evaluation*), dan evaluasi pengaruh (*impact evaluation*).

Secara umum gambaran model evaluasi sistem analisis adalah sebagai berikut.



Gambar 14: Model Evaluasi Sistem Analisis

¹¹⁵ Wirawan. Evaluasi: Teori, Model, Metodologi, Standar, Aplikasi dan Profesi, (Jakarta: Rajagrafindo Persada, 2016), h. 155.

Menurut Wirawan penjelasan dari setiap tahapan proses evaluasi system analisis sebagai berikut:

1. *Evaluasi masukan.* Masukan pada umumnya berasal dari lingkungan eksternal. Tujuan evaluasi masukan adalah untuk menjangkau, menganalisis dan menilai kecukupan kuantitas dan kualitas masukan yang diperlukan untuk merencanakan dan melaksanakan program.
2. *Evaluasi proses.* Evaluasi proses memfokuskan pada pelaksanaan program dan sering menyediakan informasi mengenai kemungkinan program diperbaiki. Evaluasi ini merupakan evaluasi formatif yang berupaya mencari jawaban atas pertanyaan:
 - Apakah standar prosedur operasi perlu diubah?
 - Apakah proses pelaksanaan program mencapai tujuan?
 - Apakah layanan program telah mencapai target dan berlangsung sesuai yang direncanakan?
 - Apakah semua faktor masukan dan proses berhasil bersinergi dan menghasilkan nilai tambah yang diharapkan?Evaluasi proses merupakan katalis untuk pembelajaran dan pertumbuhan yang berkelanjutan.
3. *Evaluasi keluaran.* Evaluasi keluaran mengukur dan menilai keluaran dari pada program yaitu produk yang dihasilkan program. Keluaran berupaya menjawab pertanyaan:
 - Berapa banyak dan berapa baik produk dari program?
 - Berapa banyak dan berapa lama orang yang mendapatkan layanan program?
 - Berapa lama (jam) klien mendapatkan layanan program?
4. *Evaluasi akibat.* Evaluasi akibat mengukur apakah klien yang mendapat layanan program berubah. Evaluasi akibat berupaya menjawab pertanyaan sebagai berikut.
 - Apakah aktifitas program telah mengubah para klien seperti yang diharapkan?

- Apakah aktifitas program mempunyai pengaruh sampingan yang tidak diperhitungkan sebelumnya?
 - Siapa dan berapa banyak dari klien yang merespon positif dan negative terhadap aktifitas program?
5. Evaluasi pengaruh. Evaluasi pengaruh menilai perubahan yang terjadi terhadap klien atau para pemangku kepentingan sebagai akibat dari intervensi yang dilakukan program. Evaluasi ini mengukur pengaruh program sebagai hasil program dalam jangka Panjang.

Sekilas model evaluasi sistem analisis mirip dengan model evaluasi CIPP. Diantara yang membedakan adalah dalam model CIPP evaluasi akibat dan pengaruh menjadi satu dalam evaluasi produk, sementara dalam evaluasi system analisis masing-masing evaluasi akibat dan pengaruh dilakukan secara terpisah. Dalam evaluasi system analisis setiap jenis evaluasi dapat dilakukan secara parsial. Dalam CIPP keempat jenis evaluasi (*context, Input, Process, Product*) menjadi satu rangkaian kegiatan yang harus dilakukan secara linear.

RANGKUMAN

Model evaluasi CIPP dianggap kerangka yang paling komprehensif dibanding dengan model lain. Salah satu alasannya adalah didalam CIPP memuat aspek *Context, Input, Proses, dan Product* sebagai objek evaluasi. Model CIPP berangkat dari fisosofi *learning by doing*- yakni suatu usaha berkelanjutan untuk mengidentifikasi dan memperbaiki kesalahan-kesalahan yang terjadi dalam kegiatan evaluasi, untuk menginventarisir dan menguji prosedur-prosedur baru yang dibutuhkan, dan untuk memelihara dan menggabungkan praktik-praktik yang efektif.

Evaluasi *adversary* merupakan model evaluasi yang menyerupai proses pengadilan. Dalam kontek hukum, cara kerja *adversary* dikembangkan sebagai proses mendengarkan untuk menentukan nilai suatu kasus dimana pihak oposan ikut terlibat. Tujuan utama model evaluasi Adversary adalah untuk mengurangi potensi bias dengan membentuk 2 evaluator yang berbeda. Kedua evaluator yang dimaksud adalah evaluator pro dan evaluator kontra. Kedua evaluator sepakat mengenai isu yang akan diselesaikan dan menyiapkan pangkalan data umum mengenai isu tersebut. Kemudian melakukan pengumpulan data khusus sesuai dengan tugas masing-masing. Selanjutnya, kedua evaluator mendiskusikan data umum dan data khusus, terutama hasil wawancara untuk mendukung argumentasi masing-masing.

Evaluasi ketimpangan adalah persoalan membandingkan standar dengan kinerja. Evaluasi model ketimpangan adalah upaya membandingkan mengenai “apa” (kinerja) yang sesungguhnya dengan apa yang seharusnya atau “standar”. Jika ditemukan perbedaan antara standar dengan kinerja yang ada, perbedaan ini disebut sebagai ketimpangan.

Model evaluasi system analisis atau disebut juga dengan *Management Evaluation Model* merupakan evaluasi yang berbasis pada teori sistem. Pada awalnya teori system diungkapkan oleh Karl Luwig von Bertalanffy (1951). Sistem adalah satu kesatuan yang utuh. Diantara bagian-bagian yang ada dalam sistem memiliki keterkaitan satu sama lain dan saling ketergantungan. Setiap system diikat oleh tempat dan waktu, dipengaruhi oleh lingkungannya, dimaknai oleh struktur dan tujuannya, serta diwujudkan sesuai fungsinya masing-masing. Tujuan teori system adalah memperagakan system yang dinamis, hambatan-hambatan, kondisi, dan untuk menguraikan prinsip-prinsip seperti tujuan, ukuran, metode, dan alat yang dapat diterapkan pada sistem lain dilevel yang sama.

Latihan 10

1. Jelaskan Salah satu tujuan utama evaluasi adversary?
2. Evaluasi ketimpangan adalah...!
3. Jelaskan hal-hal yang perlu dilakukan oleh evaluator saat melaksanakan evaluasi proses dalam model CIPP!
4. Jelaskan keunggulan-keunggulan model evaluasi ketimpangan!
5. Evaluasi pengaruh dalam model evaluasi sistem analisis berusaha melakukan....!

Pembahasan Latihan 10

1. Tujuan utama model evaluasi Adversary adalah untuk mengurangi potensi bias dengan membentuk 2 evaluator yang berbeda. Kedua evaluator yang dimaksud adalah evaluator pro dan evaluator kontra. Kedua evaluator sepakat mengenai isu yang akan diselesaikan dan menyiapkan pangkalan data umum mengenai isu tersebut. Kemudian melakukan pengumpulan data khusus sesuai dengan tugas masing-masing. Selanjutnya, kedua evaluator mendiskusikan data umum dan data khusus, terutama hasil wawancara untuk mendukung argumentasi masing-masing.
2. Evaluasi model ketimpangan adalah upaya membandingkan mengenai “apa” (kinerja) yang sesungguhnya dengan apa yang seharusnya atau “standar”. Jika ditemukan perbedaan antara standar dengan kinerja yang ada, perbedaan ini disebut sebagai ketimpangan.
3. Evaluator memonitor, mendokumentasikan, menilai, dan melaporkan berdasarkan implementasi rencana program. Evaluator memberikan umpan balik menyeluruh terkait implementasi program dan laporan berikutnya berdasarkan perkembangan program yang dilaksanakan sesuai yang dimaksud dan dibutuhkan. Staff program menggunakan laporan evaluasi proses secara periodic untuk menampung perkembangannya, mengidentifikasi isu implementasi, dan menyelaraskan rencana mereka dan kinerjanya guna memastikan kualitas program dan layanan pengiriman yang tepat waktu.
4. Keunggulan evaluasi model ketimpangan;
 - a. Model evaluasi ketimpangan sederhana dan mudah dilaksanakan.

- b. Model ketimpangan memperkenalkan konsep manajemen informasi, proses ruangan kelas dan pentingnya membangun pangkalan data yang menghubungkan karakteristik siswa dan kinerja siswa.
 - c. Model ini mengembangkan hubungan yang terus-menerus antara staf evaluator permanen dan perencana dan pengembangan program.
 - d. Model ini menggunakan evaluasi formatif untuk merevisi dan mengoreksi program untuk mengarahkan kembali pada awal pengembangan dan instalasi program.
5. Evaluasi pengaruh menilai perubahan yang terjadi terhadap klien atau para pemangku kepentingan sebagai akibat dari intervensi yang dilakukan program. Evaluasi ini mengukur pengaruh program sebagai hasil program dalam jangka Panjang.

Tes Formatif Pilihan Berganda

1. Model Evaluasi yang menyerupai proses pengadilan adalah...!
 - a. Evaluasi sistem analisis
 - b. Management evaluation model
 - c. Evaluasi Adversary
 - d. Evaluasi ketimpangan
2. Model evaluasi yang terdiri dari evaluasi masukan (*input evaluation*), evaluasi proses (*process evaluation*), evaluasi keluaran (*output evaluation*), evaluasi akibat (*outcome evaluation*), dan evaluasi pengaruh (*impact evaluation*) merupakan model evaluasi...!
 - a. CIPP
 - b. Evaluasi Sistem Analisis
 - c. Evaluasi Model Ketimpangan
 - d. Evaluasi Adversary
3. Berikut ini adalah tujuan utama evaluasi kontek dalam CIPP, kecuali...!
 - a. Menetapkan Batasan-batasan yang ada disekelilingnya dan mendeskripsikan posisi program atau untuk upaya perbaikan yang lain.
 - b. Mengidentifikasi masalah-masalah atau hambatan untuk memenuhi kebutuhan yang dinilai.

- c. Menggunakan pendanaan program khusus atau peluang lain yang relevan.
- d. Mengidentifikasi penerima program dan menilai kebutuhannya.
4. Model evaluasi yang mengharuskan membentuk dua atau lebih Tim Evaluator yang independent merupakan ciri-ciri model evaluasi...!
 - a. Discrepancy Evaluation Model
 - b. Management Evaluation Model
 - c. Impact Evaluation
 - d. Adversary Evaluation Model
5. Tokoh utama pengembang evaluasi model CIPP adalah...
 - a. Daniel L. Stuffelbeam
 - b. Michael Scriven
 - c. Karl Luwig von Bertalanffy
 - d. Malcolm M. Provus

Kunci Jawaban Tes Formatif dan Pedoman Penskoran

1. C
2. B
3. C
4. D
5. A

Bobot

Latihan soal essay 75 % dengan bobot per soal 15.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{50}{75} \times 100 = \mathbf{66,67}$$

Tes Formatif Pilihan Berganda 25% dengan bobot per soal 5.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{20}{25} \times 100 = \mathbf{80}$$

$$66,67 \times 75\% = \mathbf{50}$$

$$80 \times 25\% = \mathbf{20}$$

$$50 + 20 = \mathbf{70}$$

$$\text{Nilai akhir} = \mathbf{70}$$

MODUL 11

MODEL EVALUASI BENCHMARKING, EVALUASI KOTAK HITAM (BLACK BOX EVALUATION MODEL), EVALUASI KONOSERSIP DAN KRITIKISME, UTILIZATION-FOCUSED EVALUATION

Metode Pembelajaran	Estimasi waktu	Capaian Pembelajaran
-Presentasi virtual -Diskusi Virtual interaktif -Tanya jawab	120 Menit X 1	Mahasiswa mampu memahami model-model evaluasi Benchmarking, Evaluasi Kotak Hitam (<i>Black Box Evaluation Model</i>), Evaluasi Konosersip dan Kritikisme, <i>Utilization-focused evaluation</i> .

Materi 11. Model-model Evaluasi, Evaluasi Benchmarking, Evaluasi Kotak Hitam (*Black Box Evaluation Model*), Evaluasi Konosersip dan Kritikisme, *Utilization-focused evaluation*.

A. Evaluasi Benchmarking

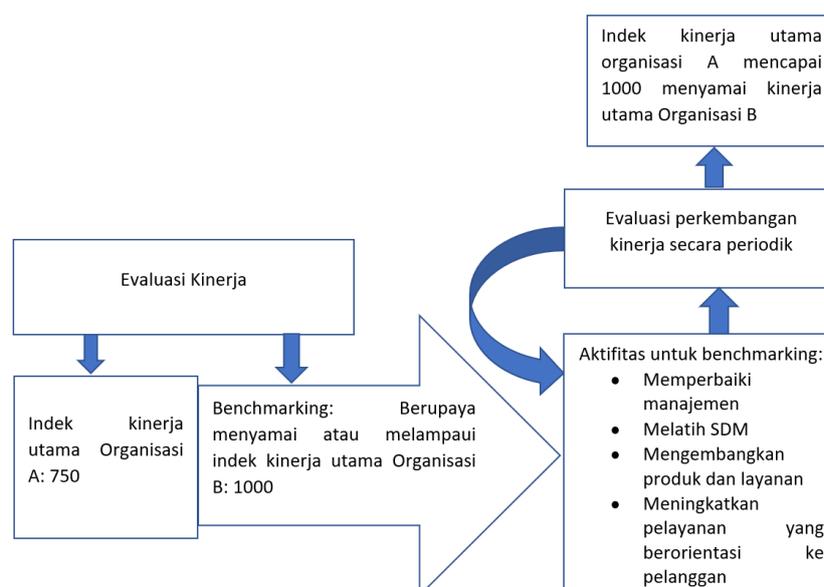
Kata *Benchmark* dilihat dari kamus bahasa Inggris memiliki makna “*something that serves as a standard by which others may be measured or judged*” sesuatu yang menjadi standar dimana pihak lain dapat mengukurnya atau menilainya. Berangkat dari makna benchmark kita diajak untuk memahami bahwa evaluasi benchmarking adalah proses evaluasi yang dilakukan dengan cara membandingkan suatu kinerja atau organisasi dengan kinerja atau organisasi lain yang dianggap memiliki kualitas lebih baik. Dalam manajemen strategik organisasi, organisasi yang kinerjanya terbaik dianggap sebagai “*best practice benchmarking*” dan dipergunakan sebagai proses untuk mengevaluasi berbagai aspek organisasi dalam hubungannya dengan praktik terbaik. Organisasi yang melakukan benchmarking membandingkan kinerja dirinya dengan organisasi yang kinerjanya terbaik dan berupaya menyamainya atau

melebihinya melalui perencanaan, pelaksanaan aktifitas (proyek), dan mengevaluasi hasilnya.¹¹⁶

Setiap organisasi menggunakan berbagai cara yang berbeda-beda dalam melakukan benchmarking. Prosedur yang dilakukan tergantung pada objek dan kondisi organisasi yang melakukan benchmarking. Secara umum langkah-langkah benchmarking yang identik dilakukan adalah;

1. Mengidentifikasi objek benchmarking. Organisasi mengevaluasi objek benchmarking yang tergantung pada lima jenis benchmarking (*proces benchmarking, financial benchmarking, product benchmarking, functional benchmarking, dan performance benchmarking*).
2. Mengidentifikasi organisasi lain yang sama. Mengidentifikasi organisasi lain yang bisnisnya, proses produksinya, produknya dan sebagainya yang sama dengan organisasi yang akan di benchmarking.

Ilustrasi tahapan proses benchmarking dapat dilihat pada bagan berikut.



Gambar 15: Contoh Proses Benchmarking

¹¹⁶ Wirawan. Evaluasi: Teori, Model, Metodologi, Standar, Aplikasi dan Profesi, (Jakarta: PT Rajagrafindo Persada, 2016), h. 156.

3. Mengidentifikasi lain yang menjadi pemimpin. Setelah langkah kedua dilewati dan mendapatkan 50 organisasi yang memiliki ciri dan karakter yang sama, kemudian diidentifikasi lagi organisasi mana yang menjadi pemimpin. Artinya yang terbaik dari 50 organisasi yang telah diidentifikasi sama. Setelah diperoleh satu organisasi terbaik, maka itu yang dijadikan rujukan untuk melakukan benchmarking
4. Mengunjungi organisasi rujukan. Pada tahap ini tim benchmarking mengunjungi organisasi rujukan untuk membicarakan maksud dan tujuan serta meminta izin melakukan benchmarking di organisasi tersebut.
5. Mensurvei praktik terbaik dari organisasi rujukan. Tim benchmarking melakukan survei, observasi, dan wawancara dengan pimpinan organisasi rujukan benchmarking. Hasilnya adalah indikator-indikator objek benchmarking terbaik yang akan menjadi contoh perubahan.
6. Menyusun rencana benchmarking. Tahap ini disusun rencana-rencana yang akan dilakukan untuk menyamai kinerja dan fungsi organisasi rujukan.
7. Melaksanakan dan mengevaluasi pelaksanaan rencana. Mengeksekusi rencana benchmarking dan mengevaluasi secara formatif dan sumatif hasil benchmarking.¹¹⁷

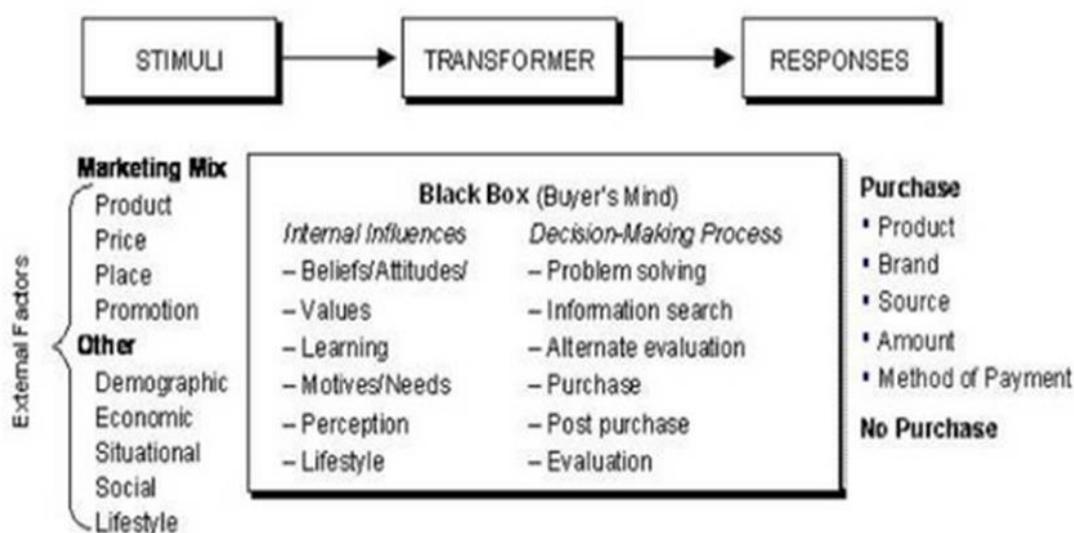
B. Evaluasi Kotak Hitam (*Black Box Evaluation Model*)

Kotak hitam adalah sesuatu yang tidak bisa dibuka sehingga tidak bisa dilihat apa isi didalamnya. Kemungkinannya hanya ditebak isinya berdasarkan masukan dan keluaran. Karena hanya bisa diidentifikasi dari masukan dan keluaran, maka apa yang sesungguhnya terjadi didalam prosesnya tidak diketahui.

Dalam konteks perilaku pelanggan, *black box* adalah pikiran manusia. Stimulus sebagai masukan (*Input*) sementara respon pelanggan adalah

¹¹⁷ Wirawan. Evaluasi: Teori, Model, Metodologi, Standar, Aplikasi dan Profesi, (Jakarta: PT Rajagrafindo Persada, 2016), h. 160-161.

keluaran (*Output*). Organisasi atau Lembaga dapat menduga atau memrediksi bagaimana otak manusia bekerja melalui respon tertentu yang diberikan. Dalam hal ini suatu institusi dapat memberikan stimulus yang sesuai kepada pelanggan dan berharap mendapatkan respon yang diharapkan dari pikiran manusia. Jika respon yang diinginkan ternyata tidak dicapai, dapat dianggap bahwa stimulusnya tidak tepat. Model kotak hitam (*Black Box*) menunjukkan interaksi stimulus, karakteristik konsumen, proses pengambilan keputusan dan respon pelanggan.



Gambar 16: Black Box Model

Model kotak hitam menganggap respon pembeli sebagai respon yang muncul secara sadar, proses keputusan yang rasional, dengan asumsi bahwa pembeli telah mengetahui masalah yang ada. Pembeli dalam konteks *black box* terdiri dari dua komponen – karakteristik pembeli dan proses keputusan pembeli.

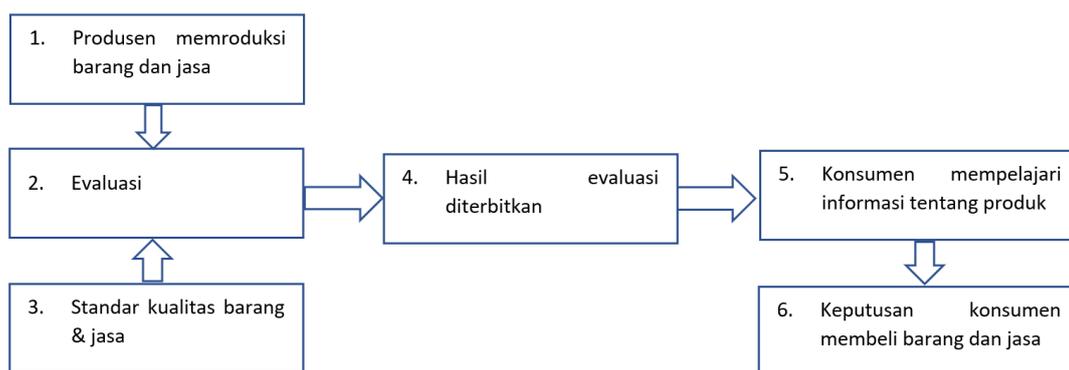
Karakteristik pembeli bisa dalam bentuk personal, sosial, budaya dan psikologi. Semua ini ada di dalam diri pembeli dan bersifat internal dan penjual tidak bisa berharap terlalu banyak dapat mengubahnya. Proses pengambilan keputusan yang dilakukan pembeli secara umum terdiri dari:

- *Pengenalan masalah:* Pelanggan/calon pembeli mencoba mengidentifikasi kebutuhannya akan sebuah produk.

- *Pencarian informasi*: ketika dibutuhkan yang dimaksud sudah ditemukan, pelanggan akan mencari informasi lebih jauh mengenai berbagai produk yang dapat memuaskan kebutuhannya. Misalnya, jika pelanggan sudah mengidentifikasi bahwa kebutuhannya adalah membeli laptop, maka akan mencari informasi lebih jauh mengenai berbagai jenis laptop yang sesuai dengan keinginannya.
- *Mengevaluasi pilihan*: Dari berbagai jenis Laptop yang ada di pasaran, kemudian dibandingkan berdasarkan parameter tertentu yang sesuai dengan keinginan pengguna. Modelnya, fitur-fiturnya, desainnya dll.
- *Keputusan membeli*: Berdasarkan parameter yang ada, calon pembeli akan menggolongkan produk mana yang paling sesuai dengan kebutuhan, kemudian memutuskan membeli salah satu yang paling cocok dengan kriteria yang diinginkan.
- *Perilaku pasca pembelian*: Pengguna bisa jadi ingin merevisi produk yang telah dibeli atau ingin menyocokkan kriteria pilihannya berdasarkan kinerja produk.

Di negara-negara maju sejumlah pabrik memproduksi piring, mangkok, sendok, dan sebagainya yang bahannya atau hiasannya terbuat dari emas 12 karat. Produk tersebut dievaluasi untuk mengetahui apakah bahannya benar dari emas 12 karat atau tidak dan daya tahannya seperti apa. Hasil evaluasi kemudian diterbitkan dan disebarakan kepada konsumen. Dengan demikian, konsumen mempunyai informasi yang cukup mengenai produk tersebut yang dapat digunakan untuk mengambil keputusan membeli atau tidak membeli produk tersebut.¹¹⁸ Proses pelaksanaan *Black Box Evaluation Model* dapat dilihat pada gambar berikut.

¹¹⁸ Wirawan. *Evaluasi: Teori, Model, Metodologi, Standar, Aplikasi dan Profesi*, (Jakarta: PT Rajagrafindo Persada, 2016), h. 162.



Gambar 17: Proses Black Box Evaluation Model

Black box evaluation sangat menolong konsumen dalam membeli produk (barang atau jasa). Konsumen sudah mempunyai informasi untuk mengambil keputusan tentang barang atau jasa yang akan dibelinya. Disisi lain *Black Box Evaluation* juga memiliki kelemahan, yaitu memerlukan banyak ahli mengenai berbagai produk sehingga berimplikasi pada kebutuhan akan laboratorium yang banyak pula. Selain itu, membutuhkan biaya yang cukup besar untuk melakukan evaluasi terhadap barang dan jasa yang jenisnya sangat banyak.

C. Evaluasi Konosersip dan Kritikisme

Evaluasi Konosersip diperkenalkan oleh Eisner tahun 1979an sebagai bentuk aplikasi dari metodologi gaya filosofis fenomenologi. Eisner mendeskripsikan peran evaluator sebagai kritik Pendidikan yang ditugaskan untuk menulis dengan gaya yang memungkinkan pembaca terlibat dalam peristiwa yang terdapat didalam kehidupan ruang kelas mengenai kritik mana yang perlu disuarakan. Dengan keterlibatan memungkinkan para pembaca mengetahui secara emosional tentang aspek kehidupan ruang kelas. Eisner menggambarkan tugas evaluator seolah-olah membawa sudut pandang baru yang bisa melihat hal-hal yang selama ini diabaikan atau terlupakan oleh seseorang yang menggunakannya didalam kelas atau dalam kontek tertentu dimana evaluasi dilakukan.

Menurut Eisner, konosersip dan kritikisme Pendidikan adalah dua proses evaluasi yang berakar dari seni. *“The aim of connoisseurship is to engender an awareness of the qualities that constitute some process*

or object and to grasp their significance. Connoisseurship pertains to matters of awareness and therefore to the creation of consciousness. Consciousness of the qualities that constitute novels, symphonies, visual works of art, and dance performances is not an automatic consequence of maturation."¹¹⁹



Elliot Wayne Eisner (March 10, 1933 – January 10, 2014) was a professor of Art and Education at the Stanford Graduate School of Education, and was one of the United States' leading academic minds. He was active in several fields including arts education, curriculum reform, qualitative research, and was the recipient of a University of Louisville Grawemeyer Award in 2005 for his work in education ^[3] as well as the Brock International Prize in 2004.

Eisner was born in Chicago, Illinois on March 10, 1933, to a family of Russian Jewish immigrants. His father, Louis Eisner (originally Leibl Iznuk), was born in the shtetl of Pavoloch in the Russian Empire (now Ukraine), and immigrated to America in 1909. He was an Oxen harness maker and a leatherworker, as well as a member of the International Fur & Leather Workers Union. His union experiences, and later work at Chicago's Platt Luggage Factory, instilled in him Socialist leanings. Louis Eisner's union activity provided him an opportunity to meet Eugene Debs at a Socialist convention for his campaign for the Election of 1920. His mother, Eva Perzov (originally Chava Perzovsky), was a stenographer from the town of Chechersk (in present-day Belarus).

Elliot Eisner received his M.A (1958) and Ph.D. (1962) in education from the University of Chicago, where he studied with Joseph Schwab, Bruno Bettelheim, Benjamin Bloom and Phillip Jackson. He was appointed Associate Professor of Education and Art at Stanford University in 1965.

Profil 7: Elliot Wayne Eisner

Tujuan konosersip adalah untuk mendorong kesadaran mutu yang terdapat beberapa proses atau objek dan untuk memahami signifikansinya. Konosersip berkaitan dengan masalah kesadaran dan penciptakan keasadaran. Kesadaran akan mutu yang terdapat dalam novel, Simponi, karya seni visual, dan tari pertunjukan bukan konsekuensi otomatis atas kematangan.

Dalam model konosersip, konoser yang berfungsi sebagai evaluator, menjadi katalisator untuk evaluasi dan instrument utama dan mengarahkan pengumpulan, menganalisis dan menginterpretasikan data. Menurut Wirawan untuk menjadi konoser, orang perlu memenuhi kriteria antara lain;

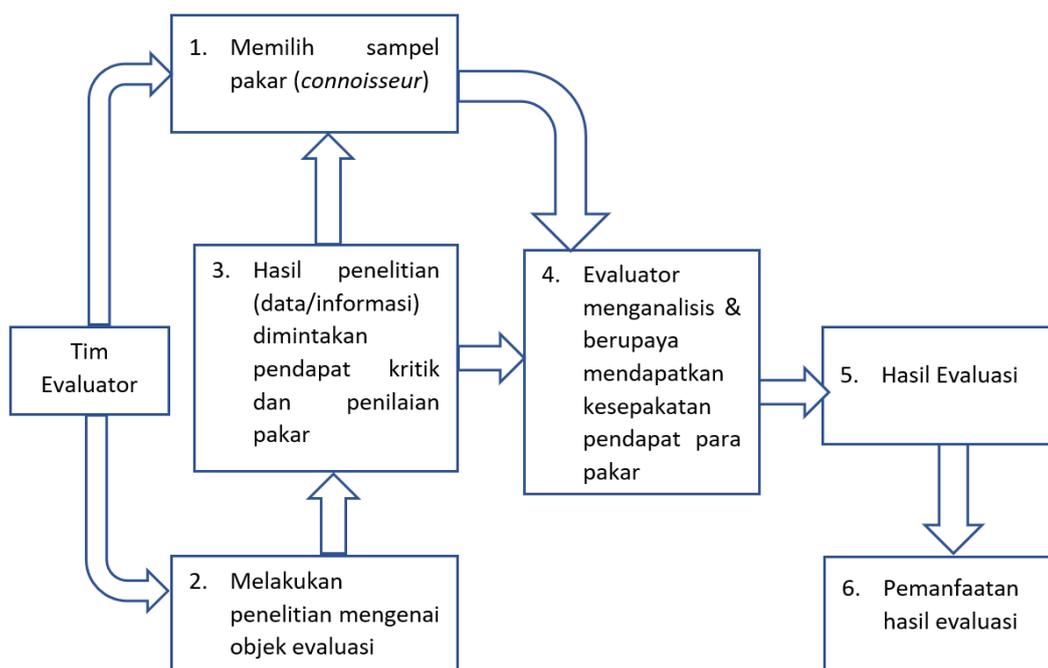
1. *Pengetahuan.* Seorang konoser memerlukan pengetahuan dan keterampilan mengenai benda atau keadaan yang dinilai.

¹¹⁹ Elliot Eisner. *International Handbook of Educational Evaluation Part One: Perspectives*, Editors: Thomas Kellaghan, Daniel L. Stufflebeam, Lori A. Wingate, (DORDRECHT: Kluwer Academic Publishers, 2003), P.153.

2. *Pengalaman.* Seorang konoser memerlukan pengalaman Panjang mengenai bidang kerjanya. Ia bukan saja menguasai teori tetapi juga seorang praktisi yang telah berpengalaman dalam bidang kerjanya.
3. *Pengakuan.* Seorang konoser memerlukan pengakuan dari masyarakat atau asosiasi konoser yang umumnya dalam bentuk sertifikat.¹²⁰

Proses evaluasi konosersip dapat digambarkan sebagai berikut.

1. *Tim evaluator memilih sampel pakar.* Evaluator perlu memilih jenis pakar yang dibutuhkan sesuai jenis kepakarannya dan yang paling berpengalaman diantara yang ada.
2. *Tim evaluator melakukan penelitian.* Tim evaluator menjangring dan mentabulasi dan menganalisis data mengenai kenijakan, program atau proyek.



Gambar 18: Proses Model Evaluasi Konosersip

3. *Hasil penelitian diserahkan kepada para pakar.* Hasil yang diserahkan ke para pakar dimintai pendapat dan penilaian.

¹²⁰ Wirawan. Evaluasi: Teori, Model, Metodologi, Standar, Aplikasi dan Profesi, (Jakarta: PT Rajagrafindo Persada, 2016), h. 164.

4. *Tim evaluator berupaya mendapatkan kesepakatan para pakar.* Sering para pakar mempunyai pendapat dan penilaian yang berbeda mengenai data dan hasil informasi yang diserahkan kepadanya.
5. *Hasil evaluasi.* Dalam keadaan para pakar tidak mencapai kesepakatan akhir pendapat dan penilaian mengenai hasil evaluasi, Tim evaluator menyusun hasil evaluasi berdasarkan perbedaan pendapat tersebut.¹²¹

Evaluasi konosersip memiliki keunggulan jika dibandingkan dengan model evaluasi yang lain, diantaranya adalah:

1. Evaluasi dilaksanakan dengan proses yang lebih saintifik dibanding model evaluasi lain.
2. Desain dan metodenya saintifik.
3. Hasil evaluasi dimintakan pendapat saintifik kepada para pakar.

Selain keunggulan, model evaluasi konosersip juga memiliki kelemahan;

1. Memerlukan waktu dan biaya lebih banyak dibanding model evaluasi lain.
2. Evaluator berfungsi sebagai penjaring data dan menganalisis data bukan penentu akhir dari hasil evaluasi. Tetapi kelemahan ini dapat diatasi jika melibatkan para pakar dibidangnya masing-masing sebagai anggota tim evaluasi.

D. Evaluasi Terfokus Utilisasi

MICHAEL QUINN PATTON mengatakan bahwa *Utilization-focused evaluation* dimulai dengan premis bahwa evaluasi harus ditentukan oleh utilitas dan kegunaannya yang sesungguhnya. Evaluator harus memfasilitasi proses evaluasi dan mendesain evaluasi dengan penuh kehati-hatian mengenai bagaimana segala sesuatunya akan dilakukan, mulai dari awal hingga akhir karena semuanya itu akan berpengaruh dalam penggunaan. Dan hal ini sesuai dengan standar evaluasi yang di

¹²¹ Wirawan. *Evaluasi: Teori, Model, Metodologi, Standar, Aplikasi dan Profesi*, (Jakarta: PT Rajagrafindo Persada, 2016), h. 164-66

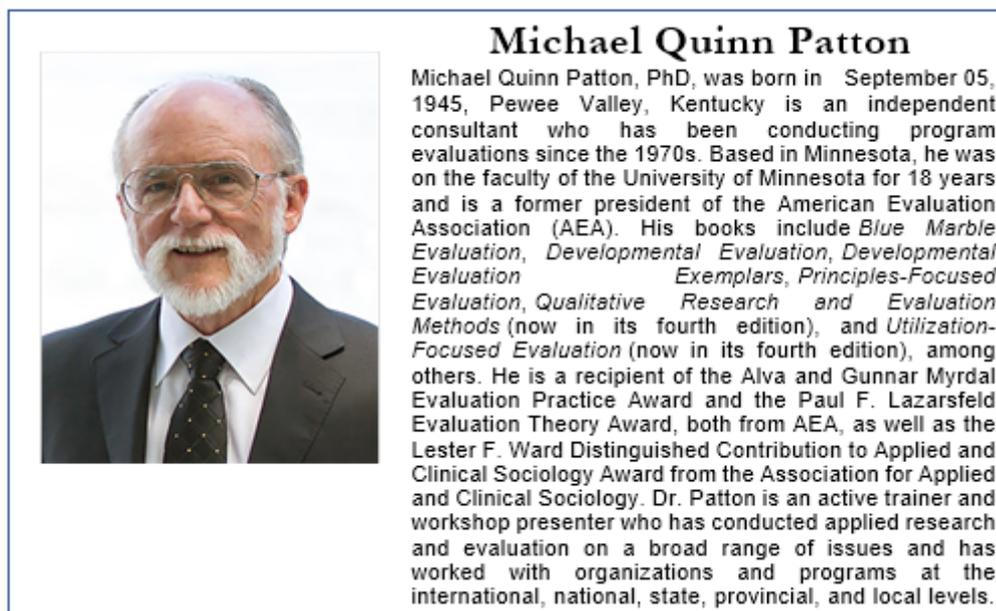
keluarkan oleh komite standar evaluasi yang diadopsi oleh asosiasi evaluasi Amerika bahwa evaluasi harus di tentukan oleh *utility, feasibility, propriety, dan accuracy*.

*Utilization-focused evaluation is concerned with how real people in the real world apply evaluation findings and experience the evaluation process. Therefore, the focus in utilization-focused evaluation is on intended use by intended users.*¹²² Evaluasi terfokus Utilisasasi berkaitan dengan bagaimana masyarakat dalam dunia nyata menerapkan penemuan evaluasi dan mengalami proses evaluasi. Oleh karena itu, focus dalam *utilization-fodused evaluation* berdasarkan tujuan penggunaan oleh penggunanya. Sejatinya tidak ada evaluasi yang bebas nilai, karenanya, *utilization-focused evaluation* ingin menjawab pertanyaan-pertanyaan tentang nilai siapa yang akan menguasai evaluasi melalui kegiatan identifikasi yang jelas, tujuan pokok pengguna yang memiliki tanggungjawab untuk menerapkan temuan evaluasi dan mengimplementasikan rekomendasi. Pada intinya, evaluasi terfokus utilisasi dilandasi pemahaman bahwa manfaat evaluasi lebih penting dari sekedar harapan atau sesuatu yang masih asumsi. Penggunaannya harus direncanakan dan difasilitasi.

Menurut MICHAEL QUINN PATTON (2003), evaluasi terfokus utilisasi dapat dilakukan dengan beberapa langkah. Berikut langkah-langkah evaluasi terfokus utilisasi menurut Patton yang penulis terjemahkan secara bebas.

Pertama, Calon pengguna evaluasi harus diidentifikasi. Para pengguna evaluasi ini dikumpulkan Bersama-sama atau diorganisir dalam beberapa cara, jika memungkinkan (misalnya, gugus tugas evaluasi dari stakeholder utama), untuk bekerja sama dengan evaluator dan berbagi dalam membuat keputusan mengenai evaluasi.

¹²² Michael Quinn Patton, *International Handbook of Educational Evaluation Part One: Perspectives*, Editors: Thomas Kellaghan, Daniel L. Stufflebeam, Lori A. Wingate, (DORDRECHT: Kluwer Academic Publishers, 2003), P.223



Profil 8: Michael Quinn Patton

Kedua, evaluator dan calon pengguna berkomitmen terhadap penggunaan evaluasi dan menentukan focus evaluasi, sebagai contoh, formatif, sumatif, atau sentral pengetahuan. Pertanyaan-pertanyaan evaluasi yang prioritas akan menjadi pertimbangan yang cukup penting berkaitan dengan pencapaian sasaran, implementasi program, dan/atau teori program tindakan. Pilihan evaluasi sangat luas, sangat banyak tipe evaluasi yang perlu didiskusikan. Evaluator bekerja Bersama dengan calon pengguna untuk menentukan prioritas penggunaan dengan memperhatikan pertimbangan politik dan etika. Dengan gaya yang interaktif dan responsif, evaluator membantu calon pengguna menjawab pertanyaan-pertanyaan: penggunaan yang diharapkan sudah ditetapkan, apakah evaluasinya bermanfaat? Seluas apa dan dengan cara apa calon pengguna berkomitmen untuk menggunakannya?

Ketiga, seluruh tahapan proses melibatkan metode, pengukuran dan keputusan desain. Calon pengguna utama dilibatkan dalam memutuskan metode yang akan digunakan agar supaya mereka memahami sepenuhnya kekuatan dan kelemahan temuan evaluasi yang akan mereka gunakan. Berbagai pilihan yang bisa menjadi pertimbangan: data kualitatif dan kuantitatif; naturalistic, eksperimen, dan desain kuasi-eksperimen; penuh harapan dan pendekatan

probalistik sampling; menekankan pada generalisasi yang lebih besar dan lebih sedikit; dan alternative-alternatif berkaitan dengan potensi ancaman *validitas, reliabilitas, dan utilitas*.

Keempat, ketika datanya telah dikumpulkan dan diatur untuk dianalisis, maka langkah keempat ini dapat dilakukan. Calon pengguna secara aktif dan terlibat secara langsung dalam menerjemahkan temuan evaluasi, mengambil keputusan berdasarkan data, dan melaksanakan rekomendasi. Selanjutnya, strategi-strategi yang spesifik dapat diformalisasikan mengingat temuan actual dan evaluator dapat memfasilitasi kelanjutan penggunaan aktualnya.

Akhirnya, keputusan tentang diseminasi laporan evaluasi dapat disebarluaskan apakah komitmen awal telah dibuat sebelumnya dalam perencanaan untuk tujuan penggunaan. Penguatan ini berbeda antara tujuan penggunaan oleh calon pengguna (*planned utilization*) dengan diseminasi yang bersifat umum sebagai pertanggungjawaban terhadap masyarakat luas (dimana harapan dan sesuatu yang tidak diinginkan dapat terjadi).

Wirawan menggambarkan pemanfaatan hasil evaluasi menggunakan *Utilization-Focused Evaluation* sebagai berikut.

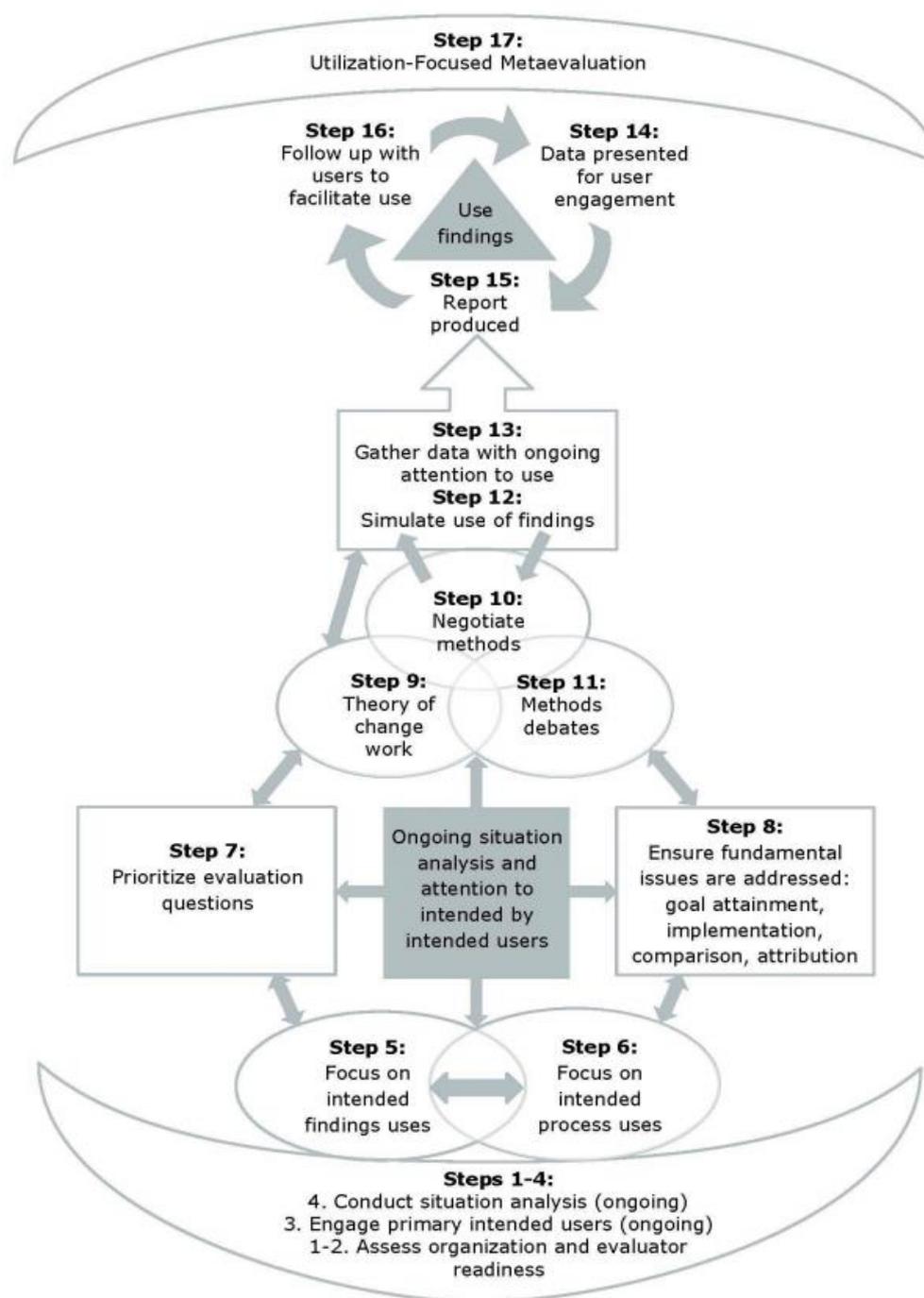
Tabel 9: Utilisasi Hasil Evaluasi Oleh Para Pemangku Kepentingan

Temuan Evaluasi	Membantu	Membuat Keputusan Mengenai
Data/informasi	Penyandang dana	Melanjutkan, Memperluas, Mempromosikan Program
	Pimpinan & Staf Program	Perbaiki, Mempercepat, Layanan-layanan baru
	Masyarakat	Dukungan, Menemukan Sumber-sumber, Kemitraan
	Para Administrator	Penstafan, Pengembangan Program di kemudian hari, Koordinasi, Kemitraan

Michael Quinn Patton (2013) mengembangkan *checklist* yang dapat digunakan oleh evaluator dalam melakukan evaluasi dengan *model Utilization-Focused Evaluation*. Terdapat tujuh belas langkah yang penting untuk di ikuti secara berurutan dan linear. Tidak dianjurkan untuk melompat ke langkah berikutnya sebelum melakukan tahapan sebelumnya. Setiap langkah harus dilakukan secara berurutan mulai dari langkah pertama, kedua, dan seterusnya. Berikut adalah tujuh

belas langkah dalam *Utilization-Focused Evaluation* yang dikembangkan Patton.

- Step 1 Assess and build program and organizational readiness for utilization-focused evaluation.
- Step 2 Assess and enhance evaluator readiness and competence to undertake a utilizationfocused evaluation.
- Step 3 Identify, organize, and engage primary intended users.
- Step 4 Conduct situation analysis with primary intended users
- Step 5 Identify primary intended uses by establishing the evaluation's priority purposes.



Gambar 19: Langkah-langkah Evaluasi Terfokus Utilisasi (*utilization-focused evaluation*)

Step 6 Consider and build in process uses if appropriate.

Step 7 Focus priority evaluation questions.

- Step 8 Check that fundamental areas for evaluation inquiry are being adequately addressed.
- Step 9 Determine what intervention model or theory of change is being evaluated.
- Step 10 Negotiate appropriate methods to generate credible findings and support intended use by intended users.
- Step 11 Make sure intended users understand potential controversies about methods and their implications.
- Step 12 Simulate use of findings.
- Step 13 Gather data with ongoing attention to use.
- Step 14 Organize and present the data for use by primary intended users.
- Step 15 Prepare an evaluation report to facilitate use and disseminate significant findings to expand influence.
- Step 16 Follow up with primary intended users to facilitate and enhance use.
- Step 17 Meta-evaluation of use: Be accountable, learn, and improve.¹²³

¹²³ Michael Quinn Patton, *Utilization-Focused Evaluation (U-FE) Checklist*, Evaluation Checklists Project www.wmich.edu/evalctr/checklists. Kamis, 4/3/2021. Pukul. 4:54 PM

RANGKUMAN

Berangkat dari makna benchmark kita diajak untuk memahami bahwa evaluasi benchmarking adalah proses evaluasi yang dilakukan dengan cara membandingkan suatu kinerja atau organisasi dengan kinerja atau organisasi lain yang dianggap memiliki kualitas lebih baik. Dalam manajemen stratejik organisasi, organisasi yang kinerjanya terbaik dianggap sebagai “*best practice benchmarking*” dan dipergunakan sebagai proses untuk mengevaluasi berbagai aspek organisasi dalam hubungannya dengan praktik terbaik.

Dalam konteks perilaku pelanggan, *black box* adalah pikiran manusia. Stimulus sebagai masukan (*Input*) sementara respon pelanggan adalah keluaran (*Output*). Organisasi atau Lembaga dapat menduga atau memrediksi bagaimana otak manusia bekerja melalui respon tertentu yang diberikan. Dalam hal ini suatu institusi dapat memberikan stimulus yang sesuai kepada pelanggan dan berharap mendapatkan respon yang diharapkan dari pikiran manusia. Jika respon yang diinginkan ternyata tidak dicapai, dapat dianggap bahwa stimulusnya tidak tepat. Model kotak hitam (*Black Box*) menunjukkan interaksi stimulus, karakteristik konsumen, proses pengambilan keputusan dan respon pelanggan.

Evaluasi Konosersip diperkenalkan oleh Eisner tahun 1979an sebagai bentuk aplikasi dari metodologi gaya filosofis fenomenologi. Peran evaluator sebagai kritik Pendidikan yang ditugaskan untuk menulis dengan gaya yang memungkinkan pembaca terlibat dalam peristiwa yang terdapat didalam kehidupan ruang kelas mengenai kritik mana yang perlu disuarakan. Dengan keterlibatan memungkinkan para pembaca mengetahui secara emosional tentang aspek kehidupan ruang kelas. Tugas evaluator seolah-olah membawa sudut pandang baru yang bisa melihat hal-hal yang selama ini diabaikan atau terlupakan oleh seseorang yang menggunakannya didalam kelas atau dalam konteks tertentu dimana evaluasi dilakukan.

Dalam model konosersip, konoser yang berfungsi sebagai evaluator, menjadi katalisator untuk evaluasi dan instrument utama dan mengarahkan pengumpulan, menganalisis dan menginterpretasikan data.

Utilization-focused evaluation dimulai dengan premis bahwa evaluasi harus di tentukan oleh utilitas dan kegunaannya yang sesungguhnya. Evaluator harus memfasilitasi proses evaluasi dan mendesain evaluasi dengan penuh kehati-hatian mengenai bagaimana segala sesuatunya akan dilakukan, mulai dari awal hingga akhir karena semuanya itu akan berpengaruh dalam penggunaan.

Latihan 11

1. Konoser adalah...
2. Evaluasi benchmarking adalah....
3. Sebutkan kelemahan-kelemahan evaluasi konosersip.
4. Jelaskan argumentasi yang berkaitan dengan penamaan *Black Box Evaluation*.
5. Sebutkan kriteria untuk menjadi evaluator konosersip.

Pembahasan Soal Latihan 11

1. Seorang evaluator yang menjadi katalisator untuk evaluasi dan sebagai instrument utama dan mengarahkan pengumpulan, menganalisis dan menginterpretasikan data.
2. Evaluasi benchmarking adalah proses evaluasi yang dilakukan dengan cara membandingkan suatu kinerja atau organisasi dengan kinerja atau organisasi lain yang dianggap memiliki kualitas lebih baik.
3. Model evaluasi konosersip memiliki kelemahan;
 - a. Memerlukan waktu dan biaya lebih banyak disbanding model evaluasi lain.
 - b. Evaluator berfungsi sebagai penjaring data dan menganalisis data bukan penentu akhir dari hasil evaluasi. Tetapi kelemahan ini dapat diatasi jika melibatkan para pakar dibidangnya masing-masing sebagai anggota tim evaluasi.
4. *Black box* adalah pikiran manusia. Stimulus sebagai masukan (*Input*) sementara respon pelanggan adalah keluaran (*Output*). Organisasi atau Lembaga dapat menduga atau memprediksi bagaimana otak manusia bekerja melalui respon tertentu yang diberikan. Dalam hal ini suatu institusi dapat memberikan stimulus yang sesuai kepada pelanggan dan berharap mendapatkan respon yang diharapkan dari pikiran manusia. Jika respon yang diinginkan ternyata tidak dicapai, dapat dianggap bahwa stimulusnya tidak tepat. Model kotak hitam (*Black Box*) menunjukkan interaksi stimulus, karakteristik konsumen, proses pengambilan keputusan dan respon pelanggan.
5. Kriteria untuk menjadi konoser antara lain;
 - a. Pengetahuan. Seorang konoser memerlukan pengetahuan dan keterampilan mengenai benda atau keadaan yang dinilai.

- b. Pengalaman. Seorang konoser memerlukan pengalaman Panjang mengenai bidang kerjanya. Ia bukan saja menguasai teori tetapi juga seorang prktisi yang telah berpengalaman dalam bidang kerjanya.
- c. Pengakuan. Seorang konoser memerlukan pengakuan dari masyarakat atau asosiasi konoser yang umumnya dalam bentuk sertifikat

Tes Formatif Pilihan Berganda

1. *Utilization-focused evaluation* dimulai dengan premis bahwa evaluasi harus di tentukan oleh utilitas dan kegunaannya yang sesungguhnya. Pernyataan tersebut memiliki makna bahwa...
 - a. Proses evaluasi harus mempertimbangkan tujuan pelaksanaan program/kegiatan menurut pembuat kebijakan.
 - b. Proses evaluasi dilakukan berdasarkan pesanan klien untuk memenuhi syarat kinerja.
 - c. Proses evaluasi dilakukan jika program tidak berjalan sesuai harapan.
 - d. Proses evaluasi berpegang pada keinginan evaluator.
2. Pernyataan berikuit merupakan keunggulan yang dimiliki Evaluasi konosersip, kecuali...!
 - a. Desain dan metodenya saintifik.
 - b. Hasil evaluasi dimintakan pendapat saintifik kepada para pakar.
 - c. Memerlukan waktu dan biaya lebih banyak dibanding model evaluasi lain.
 - d. Evaluasi dilaksanakan dengan proses yang lebih saintifik dibanding model evaluasi lain.
3. Model evaluasi yang berusaha memahami karakteristik calon pelanggan merujuk pada model evaluasi....!
 - a. Konosersip
 - b. *Utilization-focused evaluation*
 - c. *Benchmarking evaluation*
 - d. *Black box evaluation*

4. Proses evaluasi yang dilakukan dengan mengunjungi dan membandingkan kinerja organisasi sendiri dengan organisasi lain yang berkinerja terbaik adalah....!
 - a. Evaluasi Konosersip dan Kritikisme
 - b. *Utilization-focused evaluation*
 - c. *Benchmarking evaluation*
 - d. *Black box evaluation*
5. Salah satu tokoh pengembang evaluasi konosersip adalah...
 - a. Michael Quinn Patton
 - b. Elliot Wayne Eisner
 - c. Malcolm M. Provus
 - d. Michael Scriven

Kunci Jawaban dan Pedoman Penskoran

1. A
2. C
3. D
4. C
5. B

Bobot

Latihan soal essay 75 % dengan bobot per soal 15.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{50}{75} \times 100 = 66,67$$

Tes Formatif Pilihan Berganda 25% dengan bobot per soal 5.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{20}{25} \times 100 = 80$$

$$66,67 \times 75\% = \mathbf{50}$$

$$80 \times 25\% = \mathbf{20}$$

$$50 + 20 = \mathbf{70}$$

$$\text{Nilai akhir} = \mathbf{70}$$

MODUL 12

MODEL-MODEL EVALUASI AKREDITASI, THEORY-DRIVEN EVALUATION MODEL, MODEL EVALUASI SEMU ATAU PSEUDOEVALUATION, EVALUASI EKSANTE (EX-ANTE)

Metode Pembelajaran	Estimasi waktu	Capaian Pembelajaran
-Presentasi virtual -Diskusi Virtual interaktif -Tanya jawab	120 Menit X 1	-Mahasiswa mampu memahami model-model evaluasi Akreditasi, Theory-driven Evaluation Model, Model Evaluasi Semu atau <i>pseudoevaluation</i> , Evaluasi Eksante (<i>EX-ante</i>).

Materi 12. Model-model Evaluasi Akreditasi, Theory-driven Evaluation Model, Model Evaluasi Semu atau *pseudoevaluation*, Evaluasi Eksante (*EX-ante*).

A. Akreditasi

Akreditasi adalah kegiatan penilaian kelayakan program dan/atau satuan pendidikan berdasarkan kriteria yang telah ditetapkan sebagaimana dinyatakan pada UU Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional, pasal 1 ayat (22). Akreditasi sekolah/madrasah adalah proses penilaian secara komprehensif terhadap kelayakan satuan atau program pendidikan, yang hasilnya diwujudkan dalam bentuk pengakuan dan peringkat kelayakan dalam bentuk yang diterbitkan oleh suatu lembaga yang mandiri dan profesional.

Kegiatan akreditasi diharapkan menjadi pendorong dan dapat menciptakan suasana kondusif bagi perkembangan pendidikan dan

memberikan arahan untuk melakukan penjaminan mutu sekolah/madrasah yang berkelanjutan, serta terus berusaha mencapai mutu yang diharapkan. Undang-undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional (Sisdiknas) mewajibkan akreditasi bagi seluruh sekolah dan madrasah sebagai bagian dari keseluruhan upaya penjaminan mutu Pendidikan.

Proses akreditasi dilakukan secara terbuka dengan tujuan untuk membantu dan memberdayakan program dan satuan pendidikan agar mampu mengembangkan sumber dayanya dalam mencapai tujuan pendidikan nasional. Proses akreditasi dilakukan oleh Lembaga independen yang dibentuk pemerintah melalui Peraturan Menteri Pendidikan dan Kebudayaan Nomor 13 tahun 2018 tentang Badan Akreditasi Nasional Sekolah/Madrasah dan Badan Akreditasi Nasional Pendidikan Anak Usia Dini dan Pendidikan Nonformal.

Peraturan Pemerintah No. 19 Tahun 2005 tentang Standar Nasional Pendidikan pasal 2 ayat (1) menyebutkan bahwa lingkup SNP meliputi: (1) standar isi; (2) standar proses; (3) standar kompetensi lulusan; (4) standar pendidik dan tenaga kependidikan; (5) standar sarana dan prasarana; (6) standar pengelolaan; (7) standar pembiayaan; dan (8) standar penilaian pendidikan.

Tujuan Akreditasi

Akreditasi sekolah/madrasah bertujuan untuk:

1. memberikan informasi tentang kelayakan sekolah/madrasah atau program yang dilaksanakannya berdasarkan SNP;
2. memberikan pengakuan peringkat kelayakan;
3. memetakan mutu pendidikan berdasarkan SNP; dan
4. memberikan pertanggungjawaban kepada pemangku kepentingan (*stakeholder*) sebagai bentuk akuntabilitas publik.

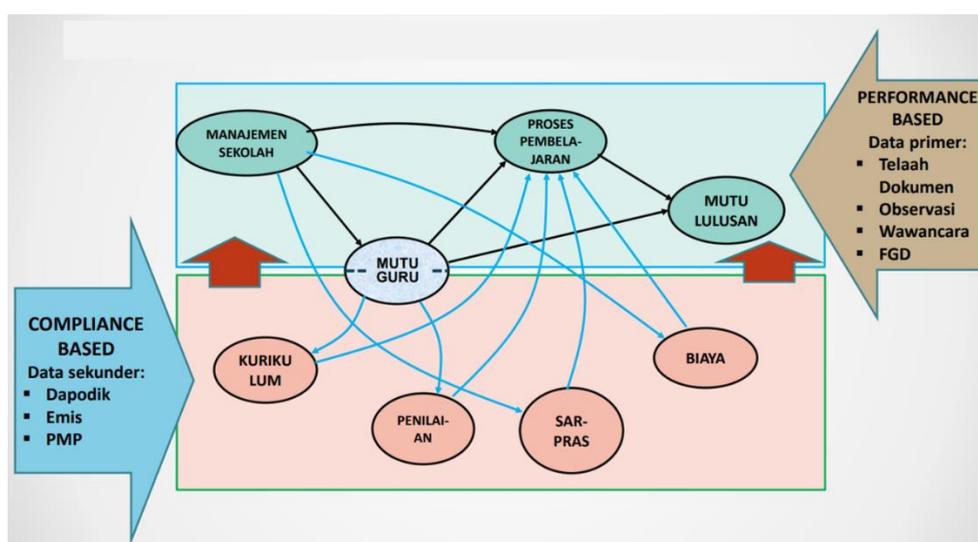
Manfaat Hasil Akreditasi

Hasil akreditasi sekolah/madrasah bermanfaat sebagai:

1. Acuan dalam upaya peningkatan mutu dan rencana pengembangan sekolah/madrasah;

2. Umpan balik dalam usaha pemberdayaan dan pengembangan kinerja warga sekolah/madrasah dalam rangka menerapkan visi, misi, tujuan, sasaran, strategi, dan program sekolah/madrasah;
3. Motivasi agar sekolah/madrasah terus meningkatkan mutu pendidikan secara bertahap, terencana, dan kompetitif baik di tingkat kabupaten/kota, provinsi, nasional bahkan regional dan internasional;
4. Bahan informasi bagi sekolah/madrasah untuk mendapatkan dukungan dari pemerintah, masyarakat, maupun sektor swasta dalam hal profesionalisme, moral, tenaga, dan dana; serta
5. Acuan bagi lembaga terkait dalam mempertimbangkan kewenangan sekolah/ madrasah sebagai penyelenggara ujian nasional.

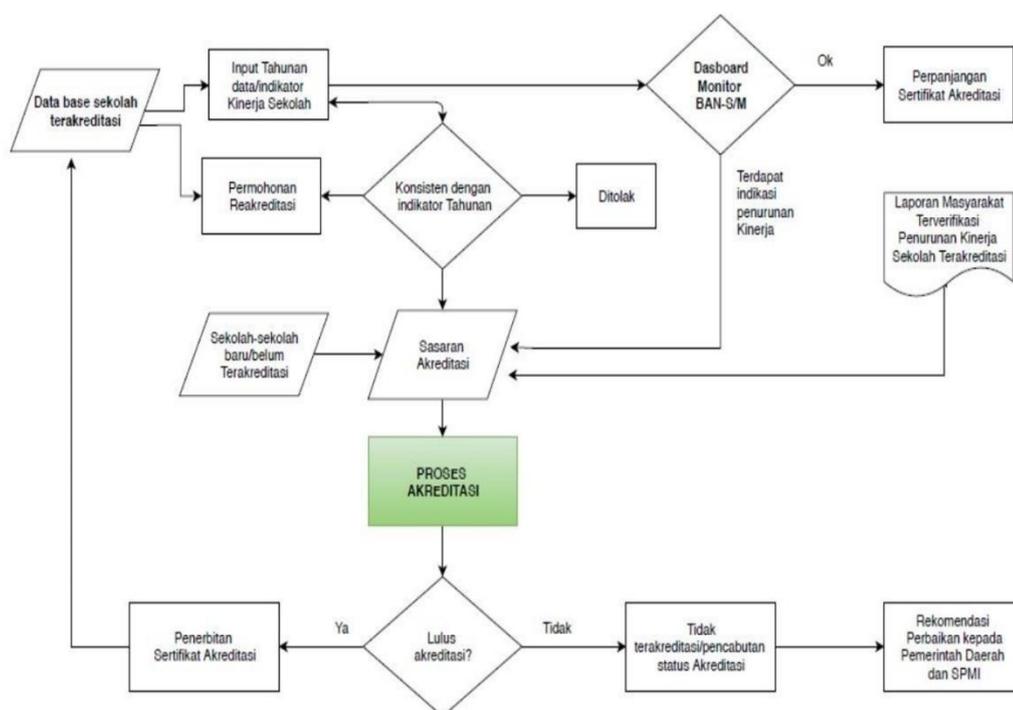
Mulai tahun 2020 paradigma akreditasi sudah berubah yang sebelumnya *compliance based* menjadi *performance based*. Jika dalam *compliance based* pengukuran berpedoman pada terpenuhinya standar melalui dokumen-dokumen yang ada di satuan Pendidikan. Dalam *performance based* dokumen tidak menjadi ukuran kualitas, pengukuran lebih titikberatkan pada empat komponen yang mengarah pada kinerja sekolah yaitu; manajemen sekolah, proses pembelajaran, mutu lulusan, dan guru yang posisinya beririsan dengan *compliance based*.



Gambar 20: Komponen Penilaian Akreditasi Berbasis Performance. Sumber: Badan Akreditasi Nasional Sekolah/Madrasah (2020)

Standar kurikulum, standar penilaian, standar sarana dan prasarana, dan standar pembiayaan posisinya sebagai *compliance based*, yang kontribusinya terhadap penilaian akreditasi persentasenya tidak lebih dari 20%. Instrument akreditasi *Performance based* dikembangkan berdasarkan standar Mutu Lulusan, Proses Pembelajaran, Manajemen Sekolah, dan Guru.

Instrument akreditasi yang dikembangkan berdasarkan empat standar kompetensi untuk mengukur kinerja sekolah, berimplikasi pada system kerja proses pelaksanaan akreditasi di sekolah/madrasah yang akan di berlakukan. Berikut flowchart sistem akreditasi yang mengakomodir *performance based*.



Gambar 21: Flowchart Sistem Akreditasi Sekolah/Madrasah

Mekanisme akreditasi dengan instrument baru yang berbasis kinerja ini dapat dilihat pada chart berikut.

1. Pelaksanaan akreditasi yang menggunakan instrumen baru IASP 2020 perlu disosialisasikan kepada Sekolah/Madrasah yang menjadi sasaran agar sekolah mengerti perbedaan akreditasi berbasis kinerja yang menggunakan Instrumen Akreditasi Satuan

Pendidikan (IASP) dengan yang sebelumnya sehingga sekolah/madrasah dapat mempersiapkan diri sesuai dengan akreditasi tahun berjalan.



Gambar 22: Mekanisme Akreditasi Sekolah/Madrasah. Sumber: BAN S/M 2020

2. Asesmen kecukupan. Sekolah/madrasah yang akan diasesmen adalah yang telah memenuhi indikator pemenuhan mutlak (IPM) yang telah ditetapkan, telah mengisi DIA, dan telah mengunggah dokumen yang diperlukan dalam IASP 2020. Kegiatan asesmen ini merupakan rangkaian kesatuan dari kegiatan penilaian yang dilakukan oleh asesor, karena dalam asesmen ini asesor telah dapat melihat keterpenuhan indikator pemenuhan mutlak dan indikator pemenuhan relatif (IPR), membaca DIA yang diisi oleh sekolah/madrasah, dan juga dapat melakukan telaah dokumen yang diunggah oleh sekolah/madrasah sebelum melakukan visitasi.
3. Visitasi ke Sekolah/Madrasah. Visitasi adalah kegiatan verifikasi, validasi, dan klarifikasi data dan informasi yang telah diisi oleh sekolah/madrasah dalam Sispena-S/M melalui wawancara dan observasi terhadap kondisi objektif sekolah/madrasah. Visitasi dapat dilakukan melalui 2 (dua) cara yaitu secara langsung/luring

dan/atau secara daring. Visitasi secara langsung tidak mungkin dilakukan dalam waktu dekat karena adanya kebijakan pemerintah untuk mencegah penyebaran COVID-19, terutama terkait physical distancing dan pembatasan penggunaan transportasi dalam rangka pencegahan penyebaran COVID-19.

4. Validasi dan Verifikasi Hasil Visitasi. Laporan visitasi yang disampaikan oleh asesor perlu divalidasi dan diverifikasi, untuk menjamin proses dan hasil akreditasi yang kredibel dan dapat dipertanggungjawabkan.
5. Verifikasi Hasil Validasi dan Penyusunan Rekomendasi. Setelah validasi dan verifikasi hasil visitasi oleh asesor, BAN-S/M Provinsi melaksanakan verifikasi hasil validasi dan penyusunan rekomendasi. Kegiatan ini dilakukan agar penetapan hasil akreditasi benar-benar objektif sesuai dengan keadaan sekolah/madrasah.
6. Penetapan Hasil dan Rekomendasi Akreditasi. Hasil dan rekomendasi akreditasi sekolah/madrasah ditetapkan melalui rapat pleno BAN-S/M dalam bentuk surat keputusan. Surat Keputusan tersebut disusun sesuai dengan provinsinya masing-masing. Rekomendasi yang disusun oleh Tim Verifikasi BAN-S/M Provinsi dan 1 (satu) anggota BAN-S/M dilaporkan kepada pleno BAN-S/M untuk ditetapkan. Kemudian hasil akreditasi dan rekomendasi disampaikan kepada pihak terkait untuk dimanfaatkan dalam perencanaan perbaikan mutu pendidikan.
7. Pengumuman Hasil Akreditasi. Masyarakat perlu memperoleh informasi tentang status dan peringkat akreditasi sekolah/madrasah. Untuk itu, BAN-S/M dan BAN-S/M Provinsi perlu mengumumkan hasil akreditasi sekolah/madrasah kepada masyarakat melalui situs web BAN-S/M dan melakukan sosialisasi. Dalam kurun waktu 14 (empat belas) hari kerja setelah pengumuman sekolah/madrasah dan masyarakat diberi kesempatan untuk mengajukan keberatan/sanggahan atas hasil akreditasi kepada BAN-S/M Provinsi dan/atau BAN-S/M. Apabila sampai dengan 14 (empat belas) hari kerja setelah pengumuman tidak ada keberatan dari sekolah/madrasah dan/atau

masyarakat atas hasil akreditasi, maka hasil penetapan akreditasi dianggap final dan tidak dapat diganggu gugat.

8. Penerbitan Sertifikat Akreditasi dan Rekomendasi. Sertifikat diterbitkan dalam jangka waktu 14 (empat belas) hari setelah pengumuman hasil akreditasi. Apabila terdapat pengaduan/keberatan terhadap hasil akreditasi pada sekolah/madrasah tertentu, maka pemberian sertifikat dan rekomendasi kepada sekolah/madrasah tersebut menunggu sampai ada tindak lanjut dan keputusan dari BANS/M Provinsi.

B. Theory-driven Evaluation Model

Evaluasi model ini disusun berdasarkan asumsi bahwa intervensi program harus diekspresikan berdasarkan hubungan kausal atau teori program. Evaluasi berbasis teori dipelopori oleh Peter Rossi, Huey Tsyeh-Chen, C. Weiss, dan Stewart. Evaluasi program berdasarkan teori program dimulai dengan (1) mengembangkan teori yang baik dan memvalidasi teori tentang bagaimana program dengan tipe tertentu yang ada didalam satu tempat operasi yang sama dapat membuahkan hasil, (2) langkah awal seperti apa yang dapat untuk memperkirakan teori yang tepat dalam kontek evaluasi program tertentu. Kondisi terdahulu lebih kepada refleksi mengenai janji implisit yang melekat dengan evaluasi berbasis teori, karena keberadaan teori berarti bahwa tubuh perkembangan teori secara substansi telah menghasilkan dan menguji serangkaian konsep, hipotesis, dan prinsip-prinsip pragmatis, serta instrument yang dikaitkan dengan panduan penelitian.

Stufflebeam & Coryn *“The theory can then help a program evaluator decide what questions; indicators (that is, manifest variables); and linkages (assumed to be causal) between and among program elements should be used to evaluate a program covered by the theory.”*¹²⁴ Teori dapat membantu evaluator program menentukan pertanyaan-pertanyaan; indicator (yaitu, variable manifest); jaringan (diasumsikan sebagai penyebab) antara dan diantara element-element program yang

¹²⁴ Daniel L. Stufflebeam and Chris L. S. Coryn. *Evaluation Theory, Models, and Applications*, (San Fransisco: Jossey-Bass A Wiley Brand), P.158.

seharusnya digunakan untuk mengevaluasi program yang dilandasi teori.

Tujuan utama program evaluasi berbasis teori adalah untuk menentukan program mana yang menarik yang secara teori masuk akal, untuk memahami kenapa program tersebut berhasil atau gagal, dan untuk memberikan arah bagi perbaikan program. Dalam membuat rangkuman dan laporan menurut evaluasi berbasis teori, evaluator mencoba menghadirkan kesimpulan, pro dan kontra, mengenai teori program yang kuat, jangkauan operasinya harus sesuai dengan teori, produksinya sesuai hasil yang diharapkan, konfirmasinya hipotesis tentang hubungan sebab akibat, eksekusinya sesuai rencana, apakah ada modifikasi tujuan atau prosedur dan nilai gunaannya untuk disebarkan dan berkelanjutan.

Seperti halnya model evaluasi yang lain, evaluasi model *Theory-driven Evaluation* juga memiliki keunggulan dan kelemahan. Stufflebeam dan Coryn mengemukakan kekuatan dan kelemahan Model *Theory-driven evaluation* sebagai berikut.

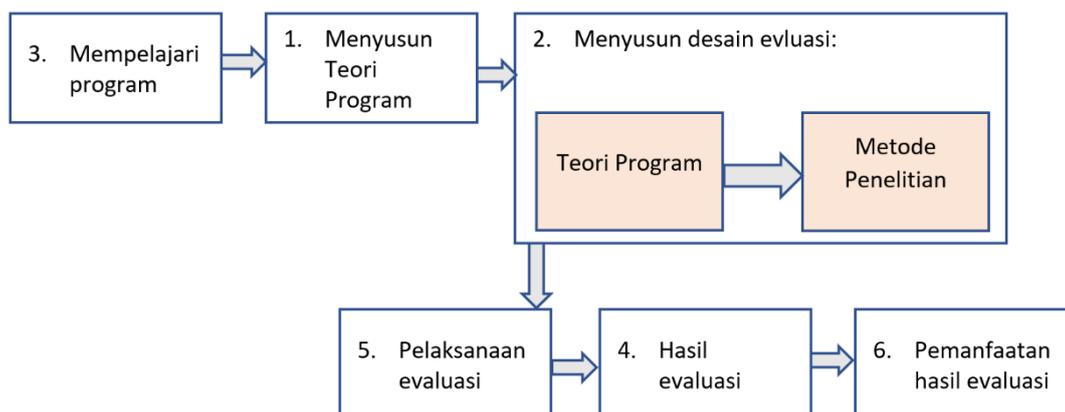
Kekuatan dalam suatu tugas evaluasi program. Sangat beralasan bagi evaluator untuk menguji keluasan rencana program yang mana dan operasinya didasarkan pada kecocokan teori atau model. Bisa menjadi sangat bermanfaat untuk menggunakan sedikit usaha menghubungkan program dan mencari variable kunci dan hubungan-hubungan. Usaha sederhana untuk membuat model program-label-bisa bermanfaat untuk mengidentifikasi ukuran variable, maka sebagai evaluator tidak perlu mengeluarkan terlalu banyak waktu dalam melakukan hal ini, dan selama sebagai model tidak perlu dipertimbangkan untuk sesuai atau melakukan validasi teori. Sayangnya metode yang dipublish tentang evaluasi yang memberikan kejelasan, arah yang berguna bagi pengembangan model yang logis dan skema-skema lainnya untuk menghadirkan dan menganalisis prosedur program, kejadian penting, dan hasil yang diinginkan yang saling mempengaruhi dan juga waktunya.

Kelemahan adalah teori logika program yang dapat dipertahankan tidak ada, para evaluator tidak perlu mengembangkannya. Faktanya, jika mereka mencoba melakukannya, mereka akan mendatangkan banyak ancaman terhadap keberhasilan evaluasi. Dibandingkan dengan mengevaluasi program dan logika yang mendasarinya, para evaluator dapat mengambil alih tanggungjawab staff program untuk desain program. Mungkin saja mereka melakukan pekerjaan kurang baik dalam konteks teori perkembangan, keterbatasan waktu dan sumber-sumber untuk mengembangkan dan menguji kesesuaian teori.

Mereka bisa saja mendatangkan konflik kepentingan yang diasosiasikan dengan keharusan mengevaluasi teori yang mereka kembangkan. Mereka dapat mengabaikan model program yang tidak tervalidasi sebagai teori, ketika tidak memenuhi hampir seluruh persyaratan teori. Mereka juga bisa menghentikan evaluasi dengan mengeluarkan terlalu banyak usaha untuk mengembangkan teori. Mereka bisa memfokuskan perhatiannya terhadap teori yang dikembangkan lebih awal didalam program dan kemudian menemukan bahwa program telah berkembang menjadi perusahaan yang sama sekali berbeda dari apa yang diteorisasikan pada permulaan, dimana teori awal dapat menjadi pemaksaan bagi program dan evaluasi program.

Sayangnya, tidak banyak program dalam Pendidikan dan ilmu social yang mendasari teori. Keinginan evaluator menggunakan evaluasi berbasis teori sering tidak menemukan evaluasi berbasis teori dapat digunakan dengan mudah untuk menunggu staf program melakukan penyusunan teori perkembangan dan langkah validasi sebelum melanjutkan evaluasi dan masih melakukan evaluasi dengan efektif, tepat waktu, dengan anggaran yang sesuai. Maka, dalam mengusulkan program kegiatan evaluasi berbasis teori, evaluator sering harus menjanjikan lebih banyak dibandingkan dengan apa yang mereka dapat lakukan.

Proses evaluasi berbasis teori menurut Wirawan dapat dilihat melalui gambar sebagai berikut.



Gambar 23: Proses Theory-Driven Evaluation

1. Mempelajari program. Evaluatur mempelajari program secara rinci dan bertemu dengan pimpinan, staf program dan para pemangku kepentingan yang berkaitan dengan program.
2. Menyusun teori program. Evaluatur menyusun logika program, perubahan yang ingin dicapai oleh program, layanan atau intervensi program, standar untuk mengukur akibat dan pengaruh program dan standar evaluasi yang akan dipergunakan.
3. Menyusun desain evaluasi. Desain terdiri dari teori program dan metode penelitian. Teori program menentukan bagaimana evaluasi akan dilaksanakan. Metode penelitian menentukan jenis informasi yang akan dijarah; sumber informasi yang diperlukan; pemilihan metode penelitian: kuantitatif, kualitatif atau mixed method dan instrument yang akan digunakan.
4. Pelaksanaan evaluasi. Proses mengumpulkan data dan informasi, tabulasi dan analisis data, menyusun laporan hasil evaluasi termasuk jika dilakukan evaluasi formatif dan sumatif.
5. Hasil evaluasi. Menyusun hasil evaluasi yang dapat berupa hasil evaluasi formatif dan hasil evaluasi sumatif. Dibahas dengan pimpinan dan staf program, kemudian dilakukan evaluasi meta untuk menentukan apakah nilai dan manfaat hasil evaluasi tinggi atau rendah.
6. Pemanfaatan hasil evaluasi. Membantu pimpinan dan staf program dalam memanfaatkan hasil evaluasi untuk mengambil keputusan mengenai program.

C. Evaluasi Semu

Salah satu model evaluasi yang sering mendapat citra negatif adalah Model Evaluasi Semu atau *pseudoevaluation*. Citra negatif berasal dari berbagai aspek proses evaluasi ini yang tidak memenuhi standar profesi evaluasi disamping hasilnya yang tidak objektif dan saintifik serta penuh manipulasi.¹²⁵

Pseudoevaluation sering bermotif politik atau keuntungan. Sebagai contoh, orang yang mencari kekuasaan bisa muncul dengan klaim tanpa dasar tentang pretasinya atau kesalahan lawan, atau mereka bisa saja menyembunyikan potensi informasi yang salah. Atau mungkin evaluator ingin menenangkan suasana, ingin mendapatkan dukungan dari pihak lain, atau ingin memperbaiki kapabilitas evaluasi pihak-pihak yang kurang teliti, pihak-pihak yang dievaluasi atau *stakeholder* lain yang bisa jadi bersepakat dengan evaluasi independent dan memperlemah hasil untuk memenangkan kepercayaan *evaluatees* (pihak yang dievaluasi) atau membantu mereka mendapatkan kekuasaan.

Evaluasi yang korup menjadi pertimbangan dalam *pseudoevaluation* karena mereka mengelabui melalui penelitian yang sengaja dilakukan dengan valid dan dapat digunakan oleh penguasa untuk menyesatkan konstituen atau untuk mendapatkan keuntungan dan menjaga keuntungan dari pihak lain dengan cara yang tidak *fair*, terutama bagi mereka yang kekuasaannya masih kecil. *Pseudoevaluation* juga digunakan karena mereka mengancam integritas profesi evaluasi. Sebaliknya, praktik baik dan konsistensi, dipahami secara luas oleh evaluator dan klien akan meningkatkan profesinya.

Stufflebeam dan Coryn dalam bukunya *Evaluation Theory, Models, and Applications*, (p.117) mengidentifikasi enam pendekatan *pseudoevaluation*: penelitian hubungan masyarakat; penelitian politik terkontrol; evaluasi percaloan; evaluasi dalih; evaluasi pemberdayaan

¹²⁵ Wirawan. *Evaluasi: Teori, Model, Metodologi, Standar, Aplikasi dan Profesi*, (Jakarta: PT Rajagrafindo Persada, 2016), h. 171.

dalam kesamaran; dan selektif, bisa jadi penyampaian balikan dari pelanggan yang bias.

Sesuai dengan namanya *pseudoevaluation* memiliki dua karakteristik yang khas yaitu; investigasi semu dan studi yang terinspirasi dari bidang Hubungan Masyarakat. Tipe pertama disebut evaluasi semu/samar, sebutan ini disematkan disebabkan oleh evaluasi ini dilakukan untuk kepentingan politik. Pelaksana evaluasi ini baik secara implisit maupun eksplisit menghadapi ancaman klien dalam menjalankan tugasnya. Tujuan klien memesan evaluasi semu ini adalah untuk membantu keberhasilan dalam mendapatkan, menjaga, atau meningkatkan pengaruh klien, kekuasaan atau uang. Dua pertanyaan utamanya dalam evaluasi ini adalah informasi apa yang menguntungkan dalam kondisi potensi konflik? Dan data seperti apa yang menguntungkan ketika harus berkonfrontasi? Metode yang khas dalam melakukan evaluasi semu diantaranya analisis dokumen, observasi partisipan, penelitian simulasi, investigasi privat, dan menjaga file-file pribadi. Secara umum klien ingin mendapat informasi yang secara teknis memungkinkan serta menjamin kemampuan mengontrol penyebaran informasinya.

Karakteristik kedua, terinspirasi dari bidang hubungan masyarakat. Kasus yang sama dalam pseudoevaluation adalah penelitian yang direncanakan, dilakukan, dan digunakan untuk tujuan pelayanan hubungan masyarakat. Dalam penelitian ini, tugas evaluator adalah mempropaganda informasi yang dibutuhkan. Tujuannya membantu klien menciptakan kesan positif bagi institusi, program, proses, dan sejenisnya. Pertanyaan yang dapat digunakan dalam penelitian model ini dirangkum dari para ahli public relation. Secara umum, penelitian public relation mencari informasi yang dapat membantu mengamankan dukungan masyarakat. Metode yang digunakan dalam penelitian public relation adalah survei, experiment, dan menggunakan konsultan ahli. Karakteristik yang dapat meresap dari penggunaan metode keragu-raguan evaluator public relation adalah upaya tendensius untuk memelihara gambaran yang bagus mengenai objek evaluasi.

D. Evaluasi Eksante (*EX-ante*)

Wirawan menyebutkan bahwa evaluasi eksante merupakan alat untuk memperbaiki kualitas program baru atau program yang diperbarui dan untuk menyediakan informasi berdasarkan pengambil keputusan mana yang dapat menentukan nilai dan manfaat suatu proposal program. Evaluasi *ex-ante* menyediakan informasi untuk keperluan:

1. Menentukan nilai dan manfaat apakah intervensi program sesuai dengan yang dibutuhkan oleh penerima manfaat dari program;
2. Sebagai dasar untuk memonitor dan evaluasi setelah program dimulai dan langkah yang diperlukan untuk monitoring dan evaluasi sesudah program dimulai serta manajemen yang tepat selama pelaksanaan program;

Menentukan apakah level pendanaan dan sumber-sumber yang diusulkan sesuai dengan hasil dan pengaruh yang diharapkan. Evaluasi *ex-ante* dilakukan sebelum proyek dimulai, "*relevansi*" diuji berdasarkan situasi yang aktual, tetapi dari sudut pandang yang berbeda, survey dapat dilakukan berdasarkan perkiraan dan prospek. Pada evaluasi pertengahan setelah proyek dimulai "*relevansi*" dan "*efficiency*" dapat dievaluasi berdasarkan situasi actual dan kinerja, tetapi "*efektifitas*" dan "*pengaruh yang kuat*" hanya bisa diuji menurut apa yang dinilai penting dan mungkin saat evaluasi dan tergantung pada tingkat pengaruh yang dihasilkan pada periode pertengahan proyek berjalan.

Tabel 10: Perbedaan Sudut Pandang Evaluasi Bagi Tipe Evaluasi Proyek

	Ex-ante evaluation	Monitoring Note 2	Mid-term evaluation	Terminal evaluation	Ex-post evaluation
Confirmation of performance	- Note1	●	●	●	◇
Grasping of implementation on process	○	●	●	●	◇
<Five Evaluation Criteria>					
Relevance	●	-	●	●	◇
Effectiveness	○	-	◇	●	-
Efficiency	○	-	●	●	-
Impact	○	-	◇	◇	●
Sustainability	○	-	○	○	●

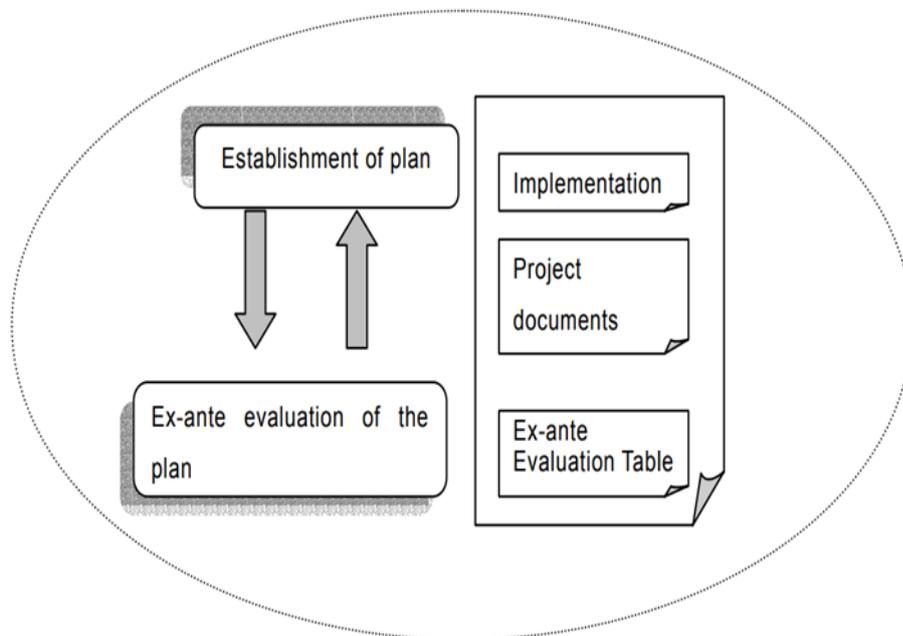
- : Penelitian berdasarkan situasi dan kinerja actual
- : Penelitian berdasarkan peramalan dan prospek
- ◇ : Penelitian berdasarkan apa yang dinilai
- : Penelitian utuh belum memungkinkan dilakukan sebelumnya, atau diselesaikan pada fase sebelumnya.

Catatan 1: Dalam konteks evaluasi *ex-ante*, berarti harus melakukan penelitian awal atau menetapkan indikator.

Catatan 2: Penelitian berbasis lima kriteria evaluasi normalnya melampaui scope monitoring, tetapi untuk manajemen dan operasinya, sangat penting agar tetap menjaga/mempertahankan sudut pandang lima kriteria evaluasi ini didalam pikiran.

Evaluasi *ex-ante* mengevaluasi rencana sebelum proyek dimulai. Hasil evaluasi digunakan untuk memperbaiki rencana dan untuk menentukan relevansi proyek. Maka dari itu, pertama-tama evaluasi akan menitikberatkan pada relevansi. Evaluasi relevansi bermaksud

menguji apakah kerjasama organisasi/perusahaan relevan dengan sudut pandang i) Kepentingannya-apakah kebutuhannya ada didaerah yang menjadi sasaran program, masyarakat dan warga setempat? ii) Prioritas – adalah kebijakan pengembangan dari pihak partner cocok dengan prioritas program? dan iii) kecukupan proyek sebagai sarana-kenapa proyek jenis ini dan dengan tujuan seperti ini dipilih? Apakah target wilayah dan target kelompok sesuai? – dan lain sebagainya. Ketika mengevaluasi kriteria evaluasi yang lain (efektifitas, pengaruh, keberlanjutan, dll), poin yang paling membedakan penelitian evaluasi pada mid-term evaluation dan evaluasi setelahnya adalah bahwa evaluasi didasarkan pada perkiraan dan prospek kinerja dan proses implementasi data sebelumnya. Dapat dikatakan, bahwa pengujian adalah tentang apakah pengaruh benar-benar dihasilkan ketika proyek diimplementasikan berdasarkan rencana, dan apakah proyek direncanakan dengan cara yang efeknya dapat di pahami dan diverifikasi.



Gambar 24: Model Penelitian Evaluasi Ex-ante JICA

Penelitian evaluasi *ex-ante* yang dilakukan JICA sebagai contoh, memiliki dua tujuan utama: i) Perencanaan proyek dan ii) perencanaan evaluasi. Pada evaluasi *ex-ante* yang dilakukan JICA, evaluasi direncanakan dan dilaksanakan menurut lima kriteria evaluasi. Hasil

evaluasi *ex-ante* dimanfaatkan untuk memperbaiki rencana, dan untuk menentukan relevansi proyek. Evaluasi *ex-ante* menyediakan informasi awal untuk melakukan monitoring dan evaluasi setelah proyek dimulai dan apakah langkah-langkah yang sangat diperlukan untuk menyesuaikan manajemen proyek secara keseluruhan dalam lingkaran proyek.

Hasil evaluasi *ex-ante* terdiri dari “dokumen proyek” yang meliputi kerangka catatan dan table evaluasi *ex-ante*. Factor-faktor penghambat dan rintangan diidentifikasi dalam evaluasi *ex-ante* direfleksikan dalam perencanaan proyek sebagai kebutuhan, atau ditulis dalam kolom tabel evaluasi tentang lima kriteria evaluasi *ex-ante* sebagai isu yang harus diingat setelah proyek dimulai. Dokumen ini tidak menjelaskan teori perencanaan. Karena teori perencanaan termasuk dalam garis dasar metode penelitian dan teori kerangka catatan serta bagaimana membangun kerangka catatan yang dibutuhkan, bagaimana menentukan indikator, target, dan metode analisis resiko.

RANGKUMAN

Secara umum model atau pendekatan evaluasi dapat diuraikan dalam enam belas bentuk yaitu; *Goal-based evaluation*, *Goal Free evaluation*, Evaluasi formatif dan sumatif, Evaluasi Responsif, Evaluasi CIPP, Evaluasi *adversary*, Evaluasi Ketimpangan, Model Evaluasi Sistem Analisis, Evaluasi Benchmarking, Evaluasi Kotak Hitam (*Black Box Evaluation Model*), Evaluasi Konosersip dan Kritikisme, *Utilization-focused Evaluation*, Akreditasi, *Theory-driven Evaluation Model*, Evaluasi Semu, Evaluasi Eksante (*EX-ante*).

Latihan 12

1. Jelaskan model evaluasi *goal-based evaluation*?
2. Jelaskan apa yang dimaksud dengan model evaluasi CIPP?
3. Jelaskan tujuan evaluasi akreditasi?
4. Mengapa evaluasi semu sering mendapat citra negatif?

5. Jelaskan perbedaan mendasar antara *Goal-based evaluation* dengan *Goal Free evaluation* model?

Pembahasan Latihan Soal 12

1. Evaluasi berbasis tujuan secara umum ingin mengukur apakah tujuan program atau proyek yang telah ditetapkan, dapat dicapai atau tidak. Model evaluasi ini memfokuskan pada mengumpulkan informasi yang bertujuan mengukur pencapaian tujuan kebijakan, program dan proyek untuk pertanggungjawaban dan pengambilan keputusan.
2. Model CIPP berangkat dari fisisofi *learning by doing*- yakni suatu usaha berkelanjutan untuk mengidentifikasi dan memperbaiki kesalahan-kesalahan yang terjadi dalam kegiatan evaluasi, untuk menginventarisir dan menguji prosedur-prosedur baru yang dibutuhkan, dan untuk memelihara dan menggabungkan praktik-praktik yang efektif. Didalam CIPP memuat aspek *Context, Input, Proses, dan Product* sebagai objek evaluasi.
3. Proses akreditasi dilakukan secara terbuka dengan tujuan untuk membantu dan memberdayakan program dan satuan pendidikan agar mampu mengembangkan sumber dayanya dalam mencapai tujuan pendidikan nasional. Akreditasi sekolah/madrasah bertujuan untuk:
 - Memberikan informasi tentang kelayakan sekolah/madrasah atau program yang dilaksanakannya berdasarkan SNP;
 - Memberikan pengakuan peringkat kelayakan;
 - Memetakan mutu pendidikan berdasarkan SNP; dan
 - Memberikan pertanggungjawaban kepada pemangku kepentingan (*stakeholder*) sebagai bentuk akuntabilitas publik.
4. *Pseudoevaluation* sering bermotif politik atau keuntungan. Evaluasi yang korup menjadi pertimbangan dalam *pseudoevaluation* karena mereka mengelabui melalui penelitian yang sengaja dilakukan dengan valid dan dapat digunakan oleh penguasa untuk menyesatkan konstituen atau untuk mendapatkan keuntungan dan menjaga keuntungan dari pihak lain dengan cara yang tidak *fair*, terutama bagi mereka yang kekuasaanya masih kecil.
5. *Goal-Based Evaluation*: Metode yang digunakan dalam evaluasi ini pada intinya adalah melibatkan tujuan operasional yang spesifik dan

mengumpulkan dan menganalisa informasi terkait untuk mengetahui seberapa baik setiap tujuan yang telah ditetapkan dicapai. *Goal-Free Evaluation*: sebagai ekspresi ketidakpuasan terhadap pendekatan *Goal-Based Evaluation* yang berpotensi invalid, karena tujuan pengembang program bisa jadi tidak sesuai, tidak realistis, tidak representative bagi kebutuhan konsumen atas penilaian, terutama maksud pengembang, atau terlalu sempit untuk mengcover pengaruh sampain yang mungkin terjadi. Model *Goal-Free Evaluation* berusaha mengabaikan tujuan program yang telah ditetapkan sebelumnya, karena akan berpengaruh terhadap hasil evaluasi yang kurang menunjukkan pengaruh sesungguhnya dari program yang dijalankan. Evaluasi bebas tujuan ingin mengetahui *the real effect of the program* terhadap sasaran program.

Test Formatif Pilihan Berganda

1. Proses evaluasi berbasis tujuan dilaksanakan dengan tahapan sebagai berikut!
 - a. Mengidentifikasi tujuan, Merumuskan tujuan menjadi indikator, Mengembangkan metode dan instrument, Memastikan program telah mencapai tujuan, Menjaring dan menganalisis data program, Kesimpulan.
 - b. Meneliti dan mengidentifikasi pengaruh program yang relevan tanpa merujuk goal dan objektif yang ada di rencana program.
 - c. Menentukan pengaruh apa yang muncul dan secara logis disebabkan oleh program dan intervensi program.
 - d. Mengidentifikasi apa yang terjadi tanpa merujuk pada tujuan dan sasaran program.
2. Evaluasi *ex-ante* adalah....
 - a. Kegiatan evaluasi yang dilakukan dengan tujuan menilai program yang dijalankan sudah sesuai atau belum dengan standar yang telah ditetapkan. Serta mengukur sudah sampai ditingkat mana pencapaian program.
 - b. Evaluasi yang di gunakan untuk kepentingan pihak tertentu dalam rangka mendapatkan keuntungan pribadi atau kelompok

- dengan memanipulasi hasil evaluasi atau melakukan rasionalisasi.
- c. Kegiatan mengevaluasi rencana sebelum proyek dimulai. Hasil evaluasi digunakan untuk memperbaiki rencana dan untuk menentukan relevansi proyek.
 - d. Evaluasi yang dimaksudkan untuk mendorong kesadaran mutu yang terdapat beberapa proses atau objek dan untuk memahami signifikansinya.
3. Dalam evaluasi akreditasi, yang dimaksud dengan kegiatan visitasi adalah....!
- a. Proses pengumpulan data dan informasi yang dilakukan asesor melalui observasi, telaah dokumen, wawancara, dan angket.
 - b. Penetapan Hasil dan Rekomendasi Akreditasi sekolah/madrasah yang dilakukan melalui rapat pleno BAN-S/M dalam bentuk surat keputusan.
 - c. Laporan visitasi yang disampaikan oleh asesor perlu divalidasi dan diverifikasi, untuk menjamin proses dan hasil akreditasi yang kredibel dan dapat dipertanggungjawabkan.
 - d. Kegiatan penilaian yang dilakukan oleh asesor untuk menentukan keterpenuhan indikator pemenuhan mutlak dan indikator pemenuhan relatif (IPR) dan kelayakan satuan Pendidikan untuk diakreditasi.
4. Evaluasi konosersip memiliki keunggulan jika dibandingkan dengan model evaluasi yang lain, diantara keunggulan tersebut adalah...!
- a. Memerlukan waktu dan biaya lebih banyak dibanding model evaluasi lain.
 - b. Evaluator berfungsi sebagai penjaring data dan menganalisis data.
 - c. Hasil evaluasi dimintakan pendapat saintifik kepada para pakar.
 - d. Evaluator bukan sebagai penentu akhir dari hasil evaluasi.
5. Model evaluasi yang erat kaitannya dengan upaya memahami pikiran customer/calon pembeli sebuah produk adalah...!
- a. Ex-ante evaluation model
 - b. Theory-driven Evaluation Model
 - c. Benchmarking evaluation

d. Black box evaluation model

Kunci jawaban dan pedoman penskoran

1. A
2. C
3. A
4. C
5. D

Bobot

Latihan soal essay 75 % dengan bobot per soal 15.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{50}{75} \times 100 = \mathbf{66,67}$$

Tes Formatif Pilihan Berganda 25% dengan bobot per soal 5.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{20}{25} \times 100 = \mathbf{80}$$

$$66,67 \times 75\% = \mathbf{50}$$

$$80 \times 25\% = \mathbf{20}$$

$$50 + 20 = \mathbf{70}$$

$$\text{Nilai akhir} = \mathbf{70}$$

MODUL 13

DELAPAN KONSEP STANDAR KEGUNAAN (UTILITY STANDARDS) DAN TUJUH STANDARD KESOPANAN (PROPRIETY STANDARDS). EMPAT KONSEP; STANDAR KELAYAKAN (FEASIBILITY STANDARDS)

Metode Pembelajaran	Estimasi waktu	Capaian Pembelajaran
-Presentasi virtual -Diskusi Virtual interaktif -Tanya jawab	120 menit X 1	Mahasiswa mampu memahami Delapan konsep Standar Kegunaan (<i>Utility Standards</i>) dan Tujuh Standard Kesopanan (<i>Propriety Standards</i>). Empat konsep; Standar Kelayakan (<i>Feasibility Standards</i>).

Materi 13. Delapan konsep Standar Kegunaan (*Utility Standards*) dan Tujuh Standard Kesopanan (*Propriety Standards*). Empat konsep; Standar Kelayakan (*Feasibility Standards*).

A. Standar Evaluasi

Salah satu Standar Evaluasi Program dikembangkan oleh *Joint Committee* diinisiasi oleh tiga organisasi: *The American Educational Research association* (AERA), *the American Psychological Association* (APA), dan *the National Council on Measurement in Education*. Representasi dari ketiga organisasi ini digabungkan oleh anggota yang berasal dari duabelas organisasi profesional lainnya (Asosiasi Administrator Sekolah Amerika, Asosiasi penilaian konseling, dan Asosiasi Pendidikan Nasional) mereka didapuk untuk mengembangkan serangkaian standar yang akan menjadi panduan evaluasi Pendidikan dan program pelatihan, proyek, dan materi-materi yang dapat digunakan di berbagai tempat. Meskipun standar tersebut belum diadopsi sebagai standar resmi oleh organisasi tersebut di atas, paling tidak telah memberikan suatu kerangka komprehensif yang dapat digunakan untuk menguji kualitas evaluasi yang akan dilakukan.

Standar evaluasi yang dimaksud dapat dirangkum menurut lima sifat evaluasi sebagai berikut:

- *Utility*: Standar ini dimaksudkan untuk memastikan bahwa evaluasi akan memberikan informasi yang dibutuhkan oleh calon pengguna. Artinya, evaluasi perlu mempertimbangkan bagaimana manfaat dan kesesuaian evaluasi bagi calon penggunanya.
- *Feasibility*: Standar ini dimaksudkan untuk memastikan bahwa evaluasi akan dilakukan dengan realistis, hati-hati, diplomatis, dan cermat. Pertimbangan ini, berkaitan dengan tingkat evaluasi yang dapat diimplementasikan dengan sukses di lokasi tertentu.
- *Propriety*: Standar ini dimaksudkan untuk memastikan bahwa evaluasi akan dilakukan secara legal, etis, dan dengan memperhatikan keselamatan mereka-mereka yang terlibat dalam evaluasi, serta mereka yang terdampak hasil evaluasi. Ini berkenaan dengan bagaimana evaluasi dilakukan dengan mempertimbangkan manusia, etika, moral, kepatutan, keabsahan, dan professional.
- *Accuracy*: Standar ini dibuat untuk memastikan bahwa evaluasi akan mengungkapkan dan menyampaikan informasi yang memadai secara Teknik tentang nilai dan manfaat program yang telah dievaluasi. Standar ini berusaha menjawab bagaimana evaluasi ini dapat diandalkan baik proses dan hasilnya, ketepatannya, dapat dipercaya dan menunjukkan hasil yang sesungguhnya tanpa ada rekayasa.
- *Meta-evaluation*: Standar ini diberlakukan untuk menjamin dan mengontrol kualitas evaluasi itu sendiri agar sesuai target yang direncanakan.

Penjelasan lebih detail mengenai masing-masing standar dapat dilihat dalam uraian seperti berikut. Diadopsi dari Donna M. Mertens & Amy T. Wilson, *Program Evaluation Theory and Practice second edition: A Comprehensive Guide* (2019), pp.26-27.

B. *Utility standards*

Standar ini diajukan untuk memastikan bahwa evaluasi akan menyediakan informasi yang dibutuhkan oleh calon pengguna.

U1 Evaluator Credibility: Evaluasi harus dilakukan oleh orang yang memenuhi syarat yaitu yang sudah mapan dan mampu menjaga kredibilitasnya dalam konteks evaluasi.

U2 Attention to Stakeholders: Evaluasi harus mencurahkan perhatiannya terhadap perbedaan individu dan kelompok yang berinvestasi dalam program dan dipengaruhi oleh evaluasinya.

U3 Negotiated Purposes: tujuan evaluasi harus diidentifikasi dan dinegosiasikan secara berkelanjutan, berdasarkan pada kebutuhan *stakeholder* dan calon pengguna.

U4 Explicit Values: Evaluasi harus mengklarifikasi dan menentukan nilai individu dan nilai budaya yang mendasari tujuan, proses, dan penilaian.

U5 Relevant Information: Informasi evaluasi harus menyampaikan identifikasi dan kebutuhan calon pengguna yang bersifat mendesak.

U6 Meaningful Processes and Products: Evaluasi harus mengonstruksi aktifitas, deskripsi, penemuan, dan keputusan dengan cara yang membangkitkan semangat partisipan untuk menemukan kembali, menginterpretasikan ulang, atau merevisi perilaku dan pemahaman mereka.

U7 Timely and Appropriate Communication and Reporting: Evaluasi harus dilakukan secara berkelanjutan untuk memenuhi kebutuhan akan informasi bagi berbagai audien.

U8 Concern for Influence and Consequences: Evaluasi harus menaikan tanggungjawab dan adaptif penggunaan, serta berhati-hati terhadap konsekuensi negatif yang tidak diharapkan dan kekeliruan dalam penggunaan.

C. *Feasibility Standards*

Standar ini dimaksudkan untuk memastikan bahwa evaluasi akan dilakukan dengan reliabel, hati-hati, diplomatis, dan cermat. Pertimbangan ini, berkaitan dengan tingkat evaluasi yang dapat diimplementasikan dengan sukses di lokasi tertentu.

F1 Project Management: Evaluasi harus menggunakan manajemen proyek strategis yang efektif.

F2 Practical Procedures: Evaluasi harus menggunakan prosedur yang praktis dan responsive terhadap standar operasi program. Hal ini dimaksudkan untuk meminimalisir terjadinya penyimpangan ketika informasi yang dibutuhkan telah diperoleh.

F3 Contextual Viability: Evaluasi harus menghargai, memonitor, dan menyeimbangkan budaya dan kepentingan politik dan kebutuhan individu serta kelompok.

F4 Resource Use: Evaluasi harus menggunakan sumber daya secara efisien dan memperhatikan konteks, kebutuhan, dan harapan klien serta pihak lain.

D. *Propriety Standards*

Standar ini dimaksudkan untuk memastikan bahwa evaluasi akan dilakukan secara legal, etis, dan dengan memperhatikan keselamatan mereka-mereka yang terlibat dalam evaluasi, serta mereka yang terdampak hasil evaluasi. Ini berkenaan dengan bagaimana evaluasi dilakukan dengan mempertimbangkan manusia, etika, moral, kepatutan, keabsahan, dan profesional.

P1 Responsive and Inclusive Orientation: Evaluasi harus responsive terhadap para pemangku kepentingan dan komunitasnya.

P2 Formal Agreements: Kesepakatan dalam evaluasi harus di negosiasikan sehingga ada kejelasan tanggungjawab dan kewajiban dan untuk memperhitungkan segala kebutuhan, harapan, dan konteks budaya klien serta stakeholder lainnya.

P3 Human Rights and Respect: Evaluasi harus menjaga hak-hak manusia dan hak undang-undang, dan harus menghormati martabat dan interaksi para partisipan dan stakeholder lainnya.

P4 Clarity and Balance: Evaluasi harus lengkap, mudah dipahami, dan *fair* dalam menyampaikan kebutuhan stakeholder dan tujuannya.

P5 Transparency and Disclosure: Evaluasi harus mendeskripsikan penemuan secara lengkap, keterbatasan, dan kesimpulan hasil bagi semua pemangku kepentingan, kecuali kalau ingin melakukan pelanggaran aturan dan standar kesopanan.

P6 Conflicts of Interests: Evaluasi harus mengidentifikasi, membatasi, dan jika perlu memainkan peran mediasi dalam situasi ketika terjadi konflik kepentingan yang mungkin membutuhkan proses perdamaian dan hasilnya.

P7 Fiscal Responsibility: Evaluasi harus mempertanggungjawabkan seluruh sumber daya yang digunakan, mengikuti proses dan prosedur *fiscal* yang berlaku dan memastikan bahwa klien memahami tentang sumber *fiscal* yang digunakan.

RANGKUMAN

Standar *utility* diajukan untuk memastikan bahwa evaluasi akan menyediakan informasi yang dibutuhkan oleh calon pengguna.

Standar *Feasibility* dimaksudkan untuk memastikan bahwa evaluasi akan dilakukan dengan realistis, hati-hati, diplomatis, dan cermat. Pertimbangan ini, berkaitan dengan tingkat evaluasi yang dapat diimplementasikan dengan sukses di lokasi tertentu.

Standar *propriety* dimaksudkan untuk memastikan bahwa evaluasi akan dilakukan secara legal, etis, dan dengan memperhatikan keselamatan mereka-mereka yang terlibat dalam evaluasi, serta mereka yang terdampak hasil evaluasi. Ini berkenaan dengan bagaimana evaluasi dilakukan dengan mempertimbangkan manusia, etika, moral, kepatutan, keabsahan, dan professional.

Latihan 13

1. Jelaskan mengapa kegiatan evaluasi harus berpedoman pada standar?
2. Sebutkan standar apa saja yang harus dipedomani dalam melaksanakan evaluasi?
3. Salah satu standar dalam *propriety* adalah *Conflicts of Interests*, yang artinya....?
4. Jelaskan pertimbangan penggunaan standar *Propriety* dalam evaluasi?
5. Jelaskan tujuan penggunaan standar *Meta-evaluasi* dalam penelitian evaluasi?

Pembahasan Latihan Soal 13

1. Serangkaian standar akan menjadi panduan evaluasi Pendidikan dan program pelatihan, proyek, dan materi-materi yang dapat digunakan di berbagai tempat serta untuk menguji kualitas evaluasi yang akan dilakukan.
2. *Utility, Feasibility, Propriety, Accuracy, Meta-evaluation.*
3. Standar evaluasi harus mengidentifikasi, membatasi, dan jika perlu memainkan peran mediasi dalam situasi ketika terjadi konflik kepentingan yang mungkin membutuhkan proses perdamaian dan hasilnya.
4. Standar *propriety* dimaksudkan untuk memastikan bahwa evaluasi akan dilakukan secara legal, etis, dan dengan memperhatikan keselamatan mereka-mereka yang terlibat dalam evaluasi, serta mereka yang terdampak hasil evaluasi. Ini berkenaan dengan bagaimana evaluasi dilakukan dengan mempertimbangkan manusia, etika, moral, kepatutan, keabsahan, dan professional.
5. Standar *Meta-evaluasi* ini diberlakukan untuk menjamin dan mengontrol kualitas evaluasi itu sendiri agar sesuai target yang direncanakan.

Test Formatif Pilihan Berganda

1. Evaluasi harus dilakukan oleh orang yang memenuhi syarat yaitu yang sudah mapan dan mampu menjaga kredibilitasnya dalam kontek evaluasi. Penjelasan tersebut merujuk pada standar....!
 - a. *Attention to Stakeholders*

- b. *Evaluator Credibility*
 - c. *Explicit Values*
 - d. *Relevant Information*
2. Berikut adalah indikator-indikator yang ada dalam *Meta-Evaluation*, kecuali...!
- a. Evaluasi harus responsive terhadap kebutuhan calon pengguna.
 - b. Evaluasi harus berdasarkan dokumentasi yang akurat dan mencukupi.
 - c. Evaluasi harus mengidentifikasi dan menerapkan standar mutu yang sesuai.
 - d. Evaluator menghormati martabat, kesejahteraan, nilai diri setiap individu dan mengakui pengaruh budaya yang dibawanya serta diantara kelompok yang lain.
3. Pernyataan yang merupakan standar *Utility* dalam evaluasi adalah...!
- a. Evaluasi harus menggunakan metode pemilihan informasi, pengumpulan, dan penyimpanan yang masuk akal.
 - b. Evaluasi harus menggunakan desain dan analisis yang secara Teknik memadai dan cocok dengan tujuan evaluasi.
 - c. Informasi evaluasi harus menyampaikan identifikasi dan kebutuhan calon pengguna yang bersifat mendesak.
 - d. Komunikasi evaluasi harus sesuai kenyataan dan bidangnya dan bebas dari miskonsepsi, distorsi, dan kesalahan.
4. Standar *Propriety* dalam evaluasi dapat diidentifikasi melalui pernyataan berikut!
- a. Informasi evaluasi harus tepat, dapat diandalkan, dan konsisten.
 - b. Informasi evaluasi harus memiliki validitas yang cukup dan scopenya untuk tujuan evaluasi.
 - c. Evaluasi harus didasarkan pada perjanjian formal dan renegotiasi, memerhatikan konteks, kebutuhan, dan harapan klien serta pihak lain.
 - d. Evaluasi harus mendokumentasikan sasaran evaluasi dan konteksnya dengan detail-detail yang sesuai dan jangkauan yang sesuai untuk tujuan evaluasi.

5. Pernyataan yang merupakan standar *Feasibility* dalam evaluasi adalah...!
- Practical Procedures: Evaluasi harus menggunakan prosedur yang praktis dan responsive terhadap standar operasi program.
 - Clarity and Balance*: Evaluasi harus lengkap, mudah dipahami, dan *fair* dalam menyampaikan kebutuhan stakeholder dan tujuannya.
 - Clarity and Balance*: Evaluasi harus lengkap, mudah dipahami, dan *fair* dalam menyampaikan kebutuhan stakeholder dan tujuannya.
 - Explicit Values: Evaluasi harus mengklarifikasi dan menentukan nilai individu dan nilai budaya yang mendasari tujuan, proses, dan penilaian.

Kunci jawaban dan pedoman penskoran

- B
- D
- C
- C
- A

Bobot

Latihan soal essay 75 % dengan bobot per soal 15.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{50}{75} \times 100 = \mathbf{66,67}$$

Tes Formatif Pilihan Berganda 25% dengan bobot per soal 5.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{20}{25} \times 100 = \mathbf{80}$$

$$66,67 \times 75\% = \mathbf{50}$$

$$80 \times 25\% = \mathbf{20}$$

$$50 + 20 = \mathbf{70}$$

$$\text{Nilai akhir} = \mathbf{70}$$

MODUL 14

DELAPAN STANDAR KETEPATAN (ACCURACY STANDARDS), TIGA STANDAR META EVALUASI DAN STANDAR ETIKA DALAM EVALUASI

Metode Pembelajaran	Estimasi waktu	Capaian Pembelajaran
-Presentasi virtual -Diskusi Virtual interaktif -Tanya jawab	120 menit X 1	-Mahasiswa mampu memahami Delapan konsep Delapan Standar Ketepatan (<i>Accuracy Standards</i>), tiga standar <i>Meta Evaluasi</i> dan Standar etika dalam evaluasi.

Materi 14. Delapan Standar Ketepatan (*Accuracy Standards*), tiga standar *Meta Evaluasi* dan standar etika dalam evaluasi.

A. *Accuracy standards*

Standar ini dibuat untuk memastikan bahwa evaluasi akan mengungkapkan dan menyampaikan informasi yang memadai secara teknik tentang nilai dan manfaat program yang telah dievaluasi. Standar ini berusaha menjawab bagaimana evaluasi ini dapat diandalkan baik proses dan hasilnya, ketepatannya, dapat dipercaya dan menunjukkan hasil yang sesungguhnya tanpa ada rekayasa.

A1 Trustworthy Conclusions and Decisions: Kesimpulan dan keputusan evaluasi harus dapat dipercaya secara budaya dan kontek dimana hal tersebut akan memiliki konsekuensi.

A2 Valid Information: Informasi evaluasi harus memiliki validitas yang cukup dan scopenya untuk tujuan evaluasi.

A3 Reliable Information: informasi evaluasi harus tepat, dapat diandalkan, dan konsisten.

A4 Explicit Evaluand and Context Descriptions: Evaluasi harus mendokumentasikan sasaran evaluasi dan konteksnya dengan detail-detail yang sesuai dan jangkauan yang sesuai untuk tujuan evaluasi.

A5 Sound Qualitative and Quantitative Methods: Evaluasi harus menggunakan metode pemilihan informasi, pengumpulan, dan penyimpanan yang masuk akal.

A6 Sound Designs and Analyses: Evaluasi harus menggunakan desain dan analisis yang secara Teknik memadai dan cocok dengan tujuan evaluasi.

A7 Explicit Evaluation Reasoning: Alasan-alasan pokok evaluasi mulai dari informasi dan analisis sampai kepada temuan, interpretasi, kesimpulan, dan pemberian makna harus terdokumentasikan secara jelas tanpa ada kerusakan dan penghilangan

A8 Valid Communication and Reporting: Komunikasi evaluasi harus sesuai kenyataan dan bidangnya dan bebas dari miskonsepsi, distorsi, dan kesalahan.

B. Meta-Evaluation

M1 Purposes: Evaluasi Meta harus responsive terhadap kebutuhan calon pengguna.

M2 Standards of Quality: evaluasi Meta harus mengidentifikasi dan menerapkan standar mutu yang sesuai.

M3 Documentation: Evaluasi Meta harus berdasarkan dokumentasi yang akurat dan mencukupi.

C. Standar Etika dalam Evaluasi

Selain standar evaluasi yang dikeluarkan oleh *Joint Committee*, terdapat Standar etika dalam evaluasi yang dikeluarkan *American Evaluation Association*. Berikut uraian lima prinsip standar etika:

- *Systematic inquiry.* Evaluators melaksanakan penelitian evaluasi berbasis data dengan seksama, metodologis, dan kontek yang relevan.

- *Competence*. Evaluators memiliki kemampuan profesional untuk melayani para pemangku kepentingan.
- *Integrity/honesty*. Evaluators bersikap jujur dan transparan untuk memastikan integritas hasil evaluasinya.
- *Respect for people*. Evaluators menghormati martabat, kesejahteraan, nilai diri setiap individu dan mengakui pengaruh budaya yang dibawanya serta diantara kelompok yang lain.
- *Common good and equity*. Evaluators berusaha keras untuk memberi sumbangsih terhadap kebaikan umum dan kemajuan yang wajar bagi masyarakat.¹²⁶

Standar evaluasi dan prinsip-prinsip etika bagi evaluator merupakan alat yang sangat berguna dalam mengembangkan dan melaksanakan penelitian evaluasi. Dengan berpegang pada pedoman standar dan etika, evaluator dibimbing untuk bekerja dalam koridor ilmiah yang dapat dipertanggungjawabkan hasilnya. Pada akhirnya nama baik evaluator sebagai peneliti akan tetap terjaga integritasnya.

¹²⁶ <https://www.eval.org/About/Guiding-Principles>

RANGKUMAN

Standar *accuracy* dibuat untuk memastikan bahwa evaluasi akan mengungkapkan dan menyampaikan informasi yang memadai secara teknik tentang nilai dan manfaat program yang telah dievaluasi. Standar ini berusaha menjawab bagaimana evaluasi ini dapat diandalkan baik proses dan hasilnya, ketepatannya, dapat dipercaya dan menunjukkan hasil yang sesungguhnya tanpa ada rekayasa.

Meta-evaluation: Standar ini diberlakukan untuk menjamin dan mengontrol kualitas evaluasi itu sendiri agar sesuai target yang direncanakan.

Standar evaluasi dan prinsip-prinsip etika bagi evaluator merupakan alat yang sangat berguna dalam mengembangkan dan melaksanakan penelitian evaluasi. Dengan berpegang pada pedoman standar dan etika, evaluator dibimbing untuk bekerja dalam koridor ilmiah yang dapat dipertanggungjawabkan hasilnya.

Latihan 14

1. Jelaskan mengapa kegiatan evaluasi harus berpedoman pada standar?
2. Sebutkan standar apa saja yang harus dipedomani dalam melaksanakan evaluasi?
3. Jelaskan alasan penelitian evaluasi harus berpedoman pada standar Etika?
4. Jelaskan pertimbangan penggunaan standar Propriety dalam evaluasi?
5. Jelaskan tujuan penggunaan standar *Meta-evaluasi* dalam penelitian evaluasi?

Pembahasan Latihan Soal 14

1. Serangkaian standar akan menjadi panduan evaluasi Pendidikan dan program pelatihan, proyek, dan materi-materi yang dapat digunakan di berbagai tempat serta untuk menguji kualitas evaluasi yang akan dilakukan.
2. *Utility, Feasibility, Propriety, Accuracy, Meta-evaluation.*
3. Standar evaluasi dan prinsip-prinsip etika bagi evaluator merupakan alat yang sangat berguna dalam mengembangkan dan melaksanakan penelitian evaluasi. Dengan berpegang pada pedoman standar dan etika, evaluator dibimbing untuk bekerja dalam koridor ilmiah yang dapat dipertanggungjawabkan hasilnya. Pada akhirnya nama baik evaluator sebagai peneliti akan tetap terjaga integritasnya.
4. Standar propriety dimaksudkan untuk memastikan bahwa evaluasi akan dilakukan secara legal, etis, dan dengan memperhatikan keselamatan mereka-mereka yang terlibat dalam evaluasi, serta mereka yang terdampak hasil evaluasi. Ini berkenaan dengan bagaimana evaluasi dilakukan dengan mempertimbangkan manusia, etika, moral, kepatutan, keabsahan, dan professional.
5. Standar *Meta-evaluasi* ini diberlakukan untuk menjamin dan mengontrol kualitas evaluasi itu sendiri agar sesuai target yang direncanakan.

Test Formatif Pilihan Berganda

1. Salah satu standar etika dalam evaluasi adalah *Systematic inquiry* yang berarti...!
 - a. Evaluators memiliki kemampuan professional untuk melayani para pemangku kepentingan.
 - b. Evaluators melaksanakan penelitian evaluasi berbasis data dengan seksama, metodologis, dan kontek yang relevan.
 - c. Evaluators berusaha keras untuk memberi sumbangsih terhadap kebaikan umum dan kemajuan yang wajar bagi masyarakat.
 - d. Evaluators menghormati martabat, kesejahteraan, nilai diri setiap individu dan mengakui pengaruh budaya yang dibawanya serta diantara kelompok yang lain.

2. Berikut adalah indicator-indikator yang ada dalam *Meta- Evaluation*, kecuali...!
 - a. Evaluasi harus responsive terhadap kebutuhan calon pengguna.
 - b. Evaluasi harus berdasarkan dokumentasi yang akurat dan mencukupi.
 - c. Evaluasi harus mengidentifikasi dan menerapkan standar mutu yang sesuai.
 - d. Evaluasi menghormati martabat, kesejahteraan, nilai diri setiap individu dan mengakui pengaruh budaya yang dibawanya serta diantara kelompok yang lain.
3. Pernyataan yang merupakan standar *Utility* dalam evaluasi adalah...!
 - a. Evaluasi harus menggunakan metode pemilihan informasi, pengumpulan, dan penyimpanan yang masuk akal.
 - b. Evaluasi harus menggunakan desain dan analisis yang secara Teknik memadai dan cocok dengan tujuan evaluasi.
 - c. Informasi evaluasi harus menyampaikan identifikasi dan kebutuhan calon pengguna yang bersifat mendesak.
 - d. Komunikasi evaluasi harus sesuai kenyataan dan bidangnya dan bebas dari miskonsepsi, distorsi, dan kesalahan.
4. Standar *Propriety* dalam evaluasi dapat diidentifikasi melalui pernyataan berikut!
 - a. Informasi evaluasi harus tepat, dapat diandalkan, dan konsisten.
 - b. Informasi evaluasi harus memiliki validitas yang cukup dan scopenya untuk tujuan evaluasi.
 - c. Evaluasi harus didasarkan pada perjanjian formal dan renegosiasi, memerhatikan konteks, kebutuhan, dan harapan klien serta pihak lain.
 - d. Evaluasi harus mendokumentasikan sasaran evaluasi dan konteksnya dengan detail-detail yang sesuai dan jangkauan yang sesuai untuk tujuan evaluasi.
5. Pernyataan yang merupakan standar *Feasibility* dalam evaluasi adalah...!
 - a. Practical Procedures: Evaluasi harus menggunakan prosedur yang praktis dan responsive terhadap standar operasi program.

- b. *Clarity and Balance*: Evaluasi harus lengkap, mudah dipahami, dan *fair* dalam menyampaikan kebutuhan stakeholder dan tujuannya.
- c. *Clarity and Balance*: Evaluasi harus lengkap, mudah dipahami, dan *fair* dalam menyampaikan kebutuhan stakeholder dan tujuannya.
- d. *Explicit Values*: Evaluasi harus mengklarifikasi dan menentukan nilai individu dan nilai budaya yang mendasari tujuan, proses, dan penilaian.

Kunci jawaban dan pedoman penskoran

1. B
2. D
3. C
4. C
5. A

Bobot

Latihan soal essay 75 % dengan bobot per soal 15.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{50}{75} \times 100 = \mathbf{66,67}$$

Tes Formatif Pilihan Berganda 25% dengan bobot per soal 5.

$$\frac{\text{Skor Perolehan}}{\text{Skor Maksimal}} \times 100 \quad \text{Contoh} \quad \frac{20}{25} \times 100 = \mathbf{80}$$

$$66,67 \times 75\% = \mathbf{50}$$

$$80 \times 25\% = \mathbf{20}$$

$$50 + 20 = \mathbf{70}$$

$$\text{Nilai akhir} = \mathbf{70}$$

GLOSARIUM

Audiens

Ialah pihak yang secara langsung atau tidak langsung berurusan dengan evaluasi. Mereka yang mempunyai perhatian dan minat terhadap evaluasi dan memperoleh hasilnya (sponsor, klien, partisipan, stakeholder).

Client

Ialah orang atau lembaga yang memesan atau meminta kepada evaluator untuk melakukan evaluasi.

Data kualitatif

Ialah data dalam keterangan tentang kejadian, transkrip wawancara, dan dokumen tertulis.

Data kuantitatif

Ialah data berupa angka atau berupa jumlah yang dapat diukur.

Evaluator

Seseorang yang mendapat mandate untuk melakukan evaluasi.

External Evaluator

Seseorang yang melaksanakan evaluasi tetapi bukan pegawai dalam organisasi/institusi yang dievaluasi.

Formative Evaluation

Evaluasi yang dilaksanakan selama perencanaan atau implementasi program atau produk yang didesain untuk memberi umpan balik untuk memperbaiki objek yang dievaluasi.

Gender Analysis

Berbagai metode yang digunakan untuk mengenali perbedaan partisipasi, akses, dan hak perempuan dibandingkan dengan hak pria dalam konteks sosial, ekonomi, dan sistem hukum.

Instrument

Ialah alat ukur yang digunakan dalam evaluasi, seperti: tes, kuesioner, observasi, wawancara, laporan, checklist, dan lain-lain.

Internal Evaluator

Seseorang yang melaksanakan evaluasi dan dipekerjakan oleh organisasi yang rumahnya akan dievaluasi.

Meta- Evaluation

adalah mengevaluasi suatu evaluasi, yang dapat dilakukan bersama kegiatan evaluasi biasa atau rutin untuk perbaikan sehingga evaluasi akan bertambah baik. Evaluasi meta dapat dilakukan ketika sedang mengevaluasi atau sesudah evaluasi selesai, dilaksanakan untuk mengetahui apa yang telah dilakukan.

Program

Ialah segala sesuatu yang dilakukan oleh seseorang, baik yang nyata (*tangible*) maupun yang abstrak (*intangible*), dengan harapan akan mendatangkan hasil atau pengaruh.

Sponsor

Ialah orang atau lembaga yang meminta evaluasi dan membiayai proses evaluasi tersebut.

Stakeholder

Pihak yang berkepentingan dan langsung dipengaruhi oleh hasil evaluasi.

Summative Evaluation

Evaluasi dilakukan pada akhir program atau setelah programnya selesai.

DAFTAR PUSTAKA

- Arikunto, Suharsimi dan Cepi Safruddin Abdul Jabar. 2009. *Evaluasi Program Pendidikan*. Jakarta: Bumi Aksara.
- Arikunto, Suharsimi. 2013. *Dasar-dasar Evaluasi Pendidikan*. Jakarta: PT Bumi Aksara.
- Creswell, John W. 2010. *Research Design: Pendekatan Kualitatif, Kuantitatif, dan Mixed Method*. Yogyakarta: Pustaka Pelajar.
- Madaus, George F and Kellaghan, Thomas. 2002. *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, Edited by Thomas Kellaghan, George F. Madaus, dan Daniel L. Stufflebeam. New York: Kluwer Academic Publishers.
- McMilan, James H. 2008. *Educational Research: Fundamentals for the Consumer*. (USA: Pearson Education, Inc.
- Mergaert, Lut and Minto, Rachel. 2015. *Ex Ante and Ex Post Evaluations: Two Sides of the Same Coin? The Case of Gender Mainstreaming in EU Research Policy*. Symposium on Policy Evaluation in the EU EJRR 1.
- Mertens, Donna M. 2005. *Research and Evaluation in Educational and Psychology: Integrating Diversity with Quantitative, Qualitative, And Mixed Methods*, 2nd Edition. California: Sage Publications, Inc.
- Mertens, Donna M & Wilson, Amy T. 2019. *Program Evaluation Theory and Practice Second edition: A Comprehensive Guide*. New York: The Guilford Press A Division of Guilford Publications, Inc.
- Ralph W. Tyler. 1989. *Educational Evaluation*. Boston: Kluwer Academic Publishers.
- Scriven, Michael. 2003. *International Handbook of Educational Evaluation*, T. Kellaghan, D.L. Stufflebeam (eds.). Dordrecht: Kluwer Academic Publishers.
- Stake, Robert E. 1973. *Program Evaluation Particulariy Responsive Evaluation* paper presented at a conference on "New Trends in Evaluation", Goteborg, Sweden.
- Stufflebeam, Daniel L & Shinkfield, Anthony J. 1988. *Systematic Evaluation A Self- Instructional Guide to Theory and Practice*. Boston: Kluwer NIJhoff Publishing a member of the Kluwer Academic Publishers Group.
- Stufflebeam, Daniel L. 2002. *Evaluation Models Viewpoints on Educational and Human Services Evaluation Second Edition, Foundational Models for Century Program Evaluation*. New York: Kluwer Academic Publishers.

- Stufflebeam, Daniel L and Coryn, Chris L. S. *Evaluation Theory, Models, and Applications Second Edition*. San Fransisco: Published by Jossey-Bass A Wiley Brand.
- Sugiyono. 2005. *Metode Penelitian Administrasi*. Bandung: Alfabeta.
- Sukardi. 2015. *Evaluasi Program Pendidikan dan Pelatihan*. Jakarta: PT Bumi Aksara.
- Suriasumantri, Jujun S. 2017. *Filsafat Ilmu: Sebuah Pengantar Populer keterkaitan Ilmu, Agama, dan Seni*. Jakarta: Pustaka Sinar Harapan.
- Wirawan. 2016. *Evaluasi Teori, Model, Metodologi, Standar, Aplikasi, dan Profesi*. Jakarta: Grafindo Persada.
- Yarbrough, Donald B. 2017. *Developing the Program Evaluation Utility Standards: Scholarly Foundations and Collaborative Processes*. Canadian Journal of Program (Special Issue), 284–304 doi: 10.3138/cjpe.349
- Youker, Brandon W, Allyssa Ingraham, dan Nicholas Bayer. 2014. An assessment of goal- free evaluation: Case studies of four goal-free evaluations. Volume 46, October 2014, Pages 10-16. <https://doi.org/10.1016/j.evalprogplan.2014.05.002>
- Committee on Scientific Principles for Education Research. (2002). *SCIENTIFIC RESEARCH IN EDUCATION*.
<http://www.nap.edu/catalog/10236.html>
<https://www.sfivet.swiss/identifying-object-evaluation>
<https://www.eval.org/About/Guiding-Principles>

Tim Penulis

Musringudin lahir di Kebumen sebuah kota kecil di daerah selatan Jawa Tengah. Penulis adalah dosen di Pascasarjana UHAMKA Prgram Studi Administrasi Pendidikan, mengampu mata kuliah Evaluasi Program Pendidikan dan Filsafat Sain & Teknologi. Pada jenjang S1 dipercaya mengajar mata kuliah Administrasi & Supervisi Pendidikan dengan sub-materi Kepemimpinan. Minat penulis dalam bidang Pendidikan diantaranya manajemen Pendidikan, kepemimpinan Pendidikan, termasuk penelitian dan evaluasi program pendidikan.

Pendidikan formal yang pernah ditempuh adalah S1 di jurusan Pendidikan Bahasa Inggris di salah satu Universitas Swasta di Jakarta. Kemudian, jenjang S2 dan S3 di selesaikan pada Universitas Negeri Jakarta Program Studi Manajemen Pendidikan. Pada tahun 2012, penulis mendapat kesempatan bergabung dengan sekolah Demokrasi Tangerang Selatan sebuah NGO bergerak dibidang Pendidikan demokrasi yang dikelola oleh Komunitas Indonesia untuk Demokrasi (KID).

Karir profesionalnya dimulai sejak tahun 2002 sebagai guru di sekolah swasta di Jakarta. Pada tahun 2004 diberi tugas tambahan sebagai kepala SMA, dan pada tahun 2013 di percaya untuk menjadi koordinator pendidikan di Lembaga Pendidikan yang mengelola unit PAUD, SD, SMP, dan SMA. Penulis juga mendapat amanah sebagai anggota Badan Akreditasi S/M Provinsi DKI Jakarta periode 2019 – 2022.

Prof. Dr. Abd. Rahman A. Ghani, SH., M.Pd, kelahiran Tangerang, 1961. Studi yang pernah didapatkan S1 di Universitas Muhammadiyah Jakarta, S2 Universitas Negeri Yogyakarta, hingga S3 di Universitas Negeri Jakarta Penelitian dan Evaluasi Pendidikan.

Mengawali pekerjaan pada bidang Pendidikan pada tahun 24 Januari 2000/17 Syawal 1420 di UHAMKA:

Menjadi Direktur Sekolah Pascasarjana pada periode 2013 – 2019 Dan sekarang diamanahi menjadi Wakil Rektor I bidang Akademik UHAMKA Periode 2019-sekarang.

Kiprah akademik yang didapat sesuai keilmuan hingga gelar Profesor dan Kepakaran dalam keilmuan Pendidikan bidang Metodologi Penelitian.

Adapun keaktifan pada masyarakat:

1. Pimpinan Daerah Muhammadiyah Jakarta Selatan sebagai Wakil Ketua (2015-2020)
2. Pimpinan Wilayah Muhammadiyah sebagai Konsultan Ahli (2015-2020) Aktif sebagai kader Persyarikatan sejak tahun 1996
3. Pengurus Majelis Ulama Indonesia (MUI) Kec. Ciledug sebagai Dewan Pertimbangan MUI (2021-2026)
4. Pengurus Himpunan Evaluasi Pendidikan Indonesia (HEPI) sebagai Dewan Penasehat (2021-2024)
5. Pengurus Badan Kerjasama Perguruan Tinggi Islam Swasta Se-Indonesia sebagai Bidang Organisasi Kelembagaan (2019-2023)
6. Pengelola School Education Journal PGSD UNIMED sebagai Reviewer.

Selain aktif dalam pada Lembaga masyarakat, karya tulis yang telah diterbitkan diantaranya: Mengurai Simpul Pendidikan, Metodologi Penelitian Tindakan Sekolah, Pendidikan Holistik Konsep dan Implementasi dalam Pendidikan, Membangun Sumber Daya Manusia Yang Agamis Dan Profesional, Pendidikan Pancasila Dan Kewarganegaraan Ikhtiar Membangun Indonesia Berkemajuan, Penilaian Autentik dalam Pembelajaran.

Dr. Dwi Priyono, M.Ed received his Doctoral degree in Educational Management from Universitas Negeri Jakarta and is currently teaching Educational Programmes Evaluation at Universitas Muhammadiyah Prof. Dr. HAMKA. Previously he obtains his bachelor degree on Elementary Education and Kindergarten. He graduated his master Degree from Western Washington University (WWU), USA. His professional interests focus on leadership and management expertise.

His current projects include the “Development of Relational Leadership Model for Kindergarten Principal”. In total, he has various published journals and books regarding Kindergarten studies such as: “The Challenge of ECCE Inclusive Education for Society 5.0 in Southeast Asia” for SEAMEO Journals, research on “Staffs' Performance in South Sumatera Department of Education and Culture”, and books titled “Relational Leadership for Kindergarten Principals”, “Early Childhood Care Education Management”.

He started his career from becoming a lecture for Kindergarten School Teachers at a Teachers' Education School at Palembang City, South Sumatera. His achievements were also to become the Head of Education Office for South Sumatera Province, Head of Cultural and Tourism Office for South Sumatera Province, and also had been the Head of Education Office for Bangka District. From 2017 until now, he serves as the Director of SEAMEO CECCEP, an active member of PGRI (Teachers Association of Republic of Indonesia).

Buku Evaluasi Program Pendidikan ini memberikan wawasan kepada mahasiswa tentang konsep Evaluasi Program sebagai riset dan menjadi dasar pengetahuan untuk menyusun proposal penelitian Evaluasi.

- Dr. Musringudin, M. Pd
- Prof. Dr. Abd. Rahman A. Ghani, M. Pd
- Dr. Dwi Priyono, M. Pd

Untuk akses **Buku Digital**,
Scan **QR CODE**



Media Sains Indonesia
Melong Asih Regency B.40, Cijerah
Kota Bandung - Jawa Barat
Email : penerbit@medsan.co.id
Website : www.medsan.co.id



ISBN 978-623-362-446-6 (PDF)

