



MODEL-MODEL EVALUASI APLIKASI DAN KOMBINASINYA

Guna Mengembangkan
Model Evaluasi-Supervisi
Pembelajaran Berbasis
4Cs/MESp 4Cs

Dr. Eny Winaryati, M.Pd | Muhammad Munsarif, M.Kom
Dr. Mardiana, M.Pd.I | Dr. Suwahono, M.Pd

MONOGRAF

MODEL-MODEL EVALUASI, APLIKASI DAN KOMBINASINYA

(Guna Mengembangkan Model Evaluasi-Supervisi
Pembelajaran Berbasis 4Cs/MESp 4Cs)

Dr. Eny Winaryati, M.Pd.
Muhammad Munsarif, M.Kom.
Dr. Mardiana, M.Pd.I.
Dr. Suwahono, M.Pd.

www.penerbitbukumurah.com

KBM
INDONESIA
www.penerbitbukumurah.com



PENERBIT KBM INDONESIA adalah penerbit dengan misi memudahkan proses penerbitan buku-buku penulis di tanah air Indonesia. Serta menjadi media *sharing* proses penerbitan buku.

MONOGRAF

MODEL-MODEL EVALUASI, APLIKASI DAN KOMBINASINYA

(Guna Mengembangkan Model Evaluasi-Supervisi Pembelajaran
Berbasis 4Cs/MESp 4Cs)

*Copyright@2021 By Dr. Eny Winaryati, M.Pd., Dkk.
All right reserved*

Penulis: **Dr. Eny Winaryati, M.Pd., Muhammad Munsarif, M.Kom.**

Dr. Mardiana, M.Pd.I., Dr. Suwahono, M.Pd.

Perancang Sampul: **Danillstr**

Tata Letak: **Ainur Rochmah**

Editor Naskah: **Shofiyun Nahidloh, S.Ag., MHI.**

Diterbitkan oleh: **PENERBIT KBM INDONESIA**

Banguntapan, Bantul-Jogjakarta (Kantor I)

Balen, Bojonegoro-Jawa Timur, Indonesia (Kantor II)

081357517526 (Tlpn/WA)

Website: www.penerbitbukumurah.com

Email: karyabaktimakmur@gmail.com

Youtube: Penerbit Sastrabook

Instagram: [@penerbit.sastrabook](https://www.instagram.com/penerbit.sastrabook) | [@penerbitbukujogja](https://www.instagram.com/penerbitbukujogja)

Anggota **IKAPI** (Ikatan Penerbit Indonesia)

Isi diluar tanggung jawab penerbit

Cetakan Pertama, November 2021

15 x 23 cm, xviii + 122 hlm

ISBN: **978-623-5679-14-3**

Undang-Undang Hak Cipta No 19 Tahun 2002 di dalam pasal 72 menjelaskan:

1. Barangsiapa dengan sengaja dan tanpa hak melakukan perbuatan sebagaimana dimaksud dalam pasal 2 ayat (1) atau pasal 49 ayat (1) dan ayat (2) di pidana dengan pidana penjara masing-masing paling singkat (satu) bulan dan atau denda paling sedikit Rp 1.000.000,00 (satu juta rupiah) Atau pidana penjara paling lama 7 (tujuh) tahun dan/atau denda paling banyak Rp 5.000.000.000,00 (lima miliar rupiah).
2. Barangsiapa dengan sengaja menyiarkan, memamerkan, mengedarkan, atau menjual kepada umum suatu ciptaan atau barang hasil pelanggaran Hak Cipta sebagaimana dimaksud dalam ayat (1), dipidana dengan pidana penjara paling lama 5 (lima) tahun dan/atau denda paling banyak Rp. 500.000.000,00 (lima ratus juta rupiah).



KATA PENGANTAR

Alhamdulillah wa syukurillah penulis telah dapat menyelesaikan buku monograf ini. Monograf kali ini dikembangkan dari artikel yang tim penulis tulis dengan judul: *Conceptual Framework of Evaluation Model on 4 C'S-Based Learning Supervision*. Artikel ini telah dimuat pada jurnal *International Journal of Learning, Teaching and Educational Research (IJLTER)*, dan Model Evaluasi dalam Supervisi Pembelajaran IPA Berbasis Lima Domain Sains, yang dimuat pada jurnal *Penelitian dan Evaluasi Pendidikan*.

Monograf ini disusun berdasarkan keterbatasan literature tentang model-model evaluasi, untuk versi Indonesia. Pembahasan diawali dengan pengertian evaluasi meliputi tujuan, fungsi, sasaran, dan kemanfaatan. Cara melakukan kombinasi antar model, yang diawali dengan alasan melakukan kombinasi model evaluasi. Dilanjutkan dengan penjelasan tentang model-model evaluasi. Pembahasan diawali dari Organizational Elements Model (OEM) yang dikembangkan oleh Roger Kaufman, *Countenance Evaluation Model* dikembangkan oleh Stake, *The Discrepancy Evaluation Models (DEM)* yang dikembangkan oleh Provus.

Ide yang mendasari penulisan ini adalah spirit dari perlunya kemudahan user dalam melakukan implementasi dan melakukan modifikasi model evaluasi yang ada. Oleh karenanya, buku ini dilengkapi dengan model evaluasi pada supervise pembelajaran yang penulis kembangkan. Melalui contoh aplikasi ini, diharapkan memberi kemudahan kepada calon peneliti ketika akan mengkombinasikan beberapa model evaluasi dengan berdasarkan

rangkaian tahapan yang dikehendaki, yang disesuaikan dengan kebutuhan penelitian.

Besar harapan penulis, buku ini dapat memberi kemanfaatan bagi banyak pihak, serta menambah khasanah koleksi buku terkait dengan penggunaan suatu model evaluasi. Semoga buku monograf ini bermanfaat, dan terimakasih.

Semarang, 20 Oktober 2021

Tim Penulis





UCAPAN TERIMA KASIH

Buku Monograf ini disusun sebagai luaran dari hibah Kompetitif Nasional yang Tim Penulis dapatkan. Hibah ini merupakan kegiatan nasional yang secara rutin setiap tahun diadakan oleh Kementerian Riset, Teknologi, dan Pendidikan Tinggi Republik Indonesia (Kemenristekdikti), dan sekarang menjadi Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi.

Ucapan terimakasih Tim Penulis sampaikan kepada rektor Universitas Muhammadiyah Semarang (UNIMUS), yang telah memberi kesempatan untuk mengikuti pengajuan proposal hibah. Hal ini ditunjukkan dengan kontrak penelitian yang telah kami sepakati.

Apresiasi yang tinggi juga tim sampaikan pada pakar (ahli evaluasi, ahli pengukuran, ahli Pendidikan) dan praktisi (guru, kepala sekolah, pengawas sekolah, dinas Pendidikan) yang selalu hadir pada kegiatan Focus Group Discussion (FGD) yang kami adakan. Terimakasih atas waktu dan tenaganya dalam mencermati dan memberikan *input* yang sangat bermanfaat terkait dimensi dan indicator juga instrument 4 (empat) karakter keterampilan abad 21 yang tim susun.

Semarang, 23 Oktober 2021

Tim Penulis



www.penerbitbukumurah.com



DAFTAR ISI

KATA PENGANTAR.....	v
UCAPAN TERIMA KASIH.....	vii
DAFTAR ISI	ix
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR.....	xv
ABSTRAK	xvii
BAB 1	
PENGERTIAN PENGUKURAN, ASESMEN, DAN EVALUASI.....	1
1. Pengukuran	1
2. Assesmem	2
3. Evaluasi	4
BAB 2	
PENGERTIAN EVALUASI.....	5
1. Pendefinisian Evaluasi Secara Lebih Luas	7
2. Metamodel Evaluasi	12
BAB 3	
PENDEKATAN MODEL EVALUASI	15
1. Pendekatan Model Evaluasi	15
2. Model-Model Evaluasi.....	17
BAB 4	
KLASIFIKASI EVALUASI	19
1. Evaluasi yang proaktif (<i>Proactive evaluation</i>).....	20
2. Evaluasi yang klarifikatif (<i>Clarificative evaluation</i>).....	20
3. Evaluasi interaktif (<i>Interactive evaluation</i>).....	20

4. Evaluasi monitoring (<i>Monitoring evaluation</i>).....	21
5. Evaluasi dampak (<i>Impact evaluation</i>).....	21
BAB 5	
PENETAPAN MODEL EVALUASI.....	23
1. Penulis yang focus pada cabang evaluasi penggunaan:.....	28
2. Penulis yang fokus pada Cabang evaluasi Methode:.....	34
3. Penulis yang fokus pada Cabang evaluasi Nilai:.....	35
BAB 6	
OEM MODEL	39
BAB 7	
MODEL-MODEL EVALUASI.....	43
1. Model CIPP.....	43
2. <i>Measurement Evaluation Model</i>	46
3. <i>Congruence Model</i>	46
4. <i>Educational System Evaluation Model</i>	48
5. Alkin Model (<i>The UCLA Evaluation Model</i>).....	49
6. <i>Illuminative Model</i> (Malcolm Parlett).....	50
7. Model Evaluasi CSE (<i>Center for Study of Evaluation</i>).....	50
8. Model evaluasi Kirkpatrick.....	51
9. CIRO Model	54
BAB 8	
STAKE MODELS.....	57
1. Pendekatan <i>Stake model</i>	58
2. Alur Pendekatan Stake Models	60
BAB 9	
THE DESCREPANCY EVALUATION MODEL (DEM).....	65
1. Ruang Lingkup DEM.....	65
2. Tahapan DEM	67
3. Persyaratan DEM.....	71
BAB 10	
MESp 4Cs	73

BAB 11

METODE..... 79

- 1. Pendekatan yang digunakan 79
- 2. Cara Membangun Kerangka Konsep..... 82

BAB 12

PEMBAHASAN..... 85

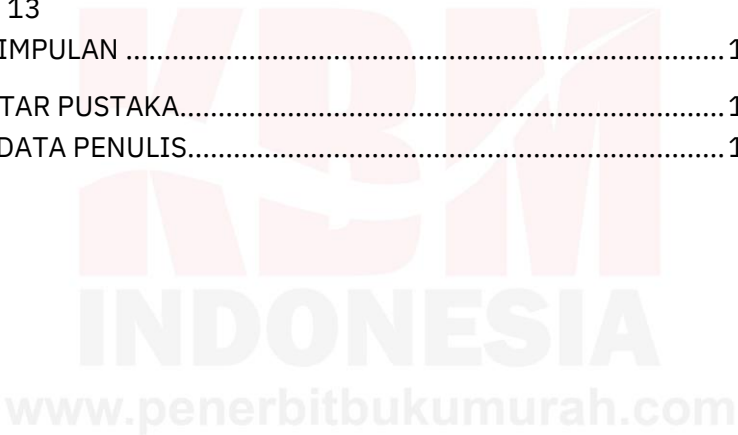
- 1. Kerangka Konseptual Awal Yang Dibangun..... 85
- 2. Landasan Model Evaluasi..... 87
- 3. Membangun Konten Evaluasi Supervisi Pembelajaran Berbasis 4 C's..... 88
- 4. Membangun kerangka MESp 4C's..... 91

BAB 13

KESIMPULAN 103

DAFTAR PUSTAKA..... 105

BIODATA PENULIS..... 119





www.penerbitbukumurah.com



DAFTAR TABEL

Tabel 1. Bentuk Evaluasi	13
Tabel 2. <i>The Five Levels of Results, the Levels of Planning and a Brief Description</i>	41
Tabel 3. Model Evaluasi Kirkpatrick's	52
Tabel 4. Tahap Evaluasi DEM	67
Tabel 5. Tahapan Supervisi Pembelajaran	74

INDONESIA
www.penerbitbukumurah.com



www.penerbitbukumurah.com



DAFTAR GAMBAR

Gambar 1.	Definisi Evaluasi (Sumber: Owen, 2004).....	8
Gambar 2.	Dampak Evaluasi terhadap Intervensi Sosial.....	9
Gambar 3.	Dampak Evaluasi Memberikan Informasi.....	10
Gambar 4.	Evaluasi Sebagai Pemanfaatan Pengetahuan	11
Gambar 5.	Pohon Teori Evaluasi.....	23
Gambar 6.	Measurement Evaluation Model.....	47
Gambar 7.	Model Evaluasi CSE.....	50
Gambar 8.	Implementasi Stake Model.....	59
Gambar 9.	Proses Deskripsi Data Model Stake	61
Gambar 10.	Stake's Representation of the Process of Judging the Merit of an Educational Program	63
Gambar 11.	Diagram Alir Tahap-Tahap Proses Evaluasi DEM ...	69
Gambar 12.	Alur Pelaksanaan Supervise Pembelajaran.....	75
Gambar 13.	Rangkaian tahap observasi- evaluasi-supervisi	76
Gambar 14.	Rancangan Konsep MESp 4Cs	77
Gambar 15.	Alur Kerangka Konseptual Model Evaluasi Supervise Pembelajaran	81
Gambar 16.	Kerangka Konsep MESp 4 Cs	84
Gambar 17.	Kerangka konsep membangun model evaluasi.....	87
Gambar 18.	Implementasi Stake Model.....	92
Gambar 19.	Proses diskripsi data contingencies dan congruence	93
Gambar 20.	Alur penyesuaian Tahap Pembelajaran- Supervisi-Stakes Model Evaluasi.....	93
Gambar 21.	Hubungan OEM dengan Beberapa Model Evaluasi.	95

Gambar 22. Diagram Alir Tahap Proses Evaluasi DEM..... 97
Gambar 23. Tahapan Proses Evaluasi 98
Gambar 24. Hubungan 6 Cell MESp 4C's dan Tahapan
Supervisi Pembelajaran 99
Gambar 25. Tahapan Proses Evaluasi Melalui MESp 4Cs100
Gambar 26. Tahapan Pengembangan MESp 4Cs101




Di era sekarang ini, guru dituntut meimplementasikan pembelajaran dan penilaian abad 21. Ada 4 karakter keterampilan yang harus dimiliki oleh guru yang ditransferkan kepada siswa, yaitu: (1) *Critical Thinking and Problem Solving*, (2) *Creativity and Inovation*, (3) *Collaboration*, dan (4) *Communication* (4C's). Harapannya agar siswa memiliki kesiapan menghadapi persoalan, tantangan dan kebutuhan masa depan yang semakin kompleks. Pelaksanaan dan ketercapaian 4C's diatas perlu untuk dinilai melalui aktivitas supervise. Harapannya baik siswa, guru maupun kepala sekolah akan memiliki komitmen yang sama dalam implementasi 4 C's ini. Supervisi yang dilakukan mengadopsi teori 360° Feedback, yang mendasarkan supervise dari segala arah yakni oleh: kepala sekolah, guru sejawat dan diri sendiri. Teori ini diadopsi berdasarkan data penelitian bahwa kepala sekolah tidak memberi tahu keseluruhan tentang strategi mengajar, kurang cepat memberi umpan balik dan tindak lanjut, pelaksanaan supervisi akademik masih kurang efektif, tidak terjadwal, kurang adanya ketersediaan waktu yang cukup untuk pemberian aktivitas reflektif.

Evaluasi dilakukan adalah untuk menghasilkan informasi, yang digunakan untuk mengambil suatu kebijakan/keputusan. Guna memperoleh informasi dapat dilakukan evaluasi dengan cara yang sederhana, sampai menggunakan tahapan-tahapan yang rinci dari berbagai strategi. Ada banyak cara/strategi untuk melakukan evaluasi, yakni melalui pendekatan suatu model evaluasi. Pendekatan model evaluasi yang dipilih disesuaikan dengan

karakteristik, fungsi, tujuan, sasaran, kondisi lapangan serta kemanfaatan dari evaluasi yang akan dilakukan.

Keterlaksanaan supervise ini perlu untuk dievaluasi terkait perencanaan, pelaksanaan dan output yang telah dilakukan oleh supervisor. Alat evaluasi yang digunakan mendasarkan kombinasi beberapa model evaluasi, yang lebih menitikberatkan pada teori kesenjangan. Model evaluasi yang digunakan adalah Stake Models, dikombinasi *The Discrepancy Evaluation Models* (DEM), dengan menggunakan Organizational Elements Model (OEM) sebagai alat penyetara. Kerangka model evaluasi yang akan digunakan adalah Model Evaluasi Supervisi Pembelajaran Berbasis 4 C's disebut MESp 4C's. Penelitian terkait supervise terhadap pelaksanaan 4C's oleh guru maupun evaluasi terhadap pelaksanaan supervise pembelajaran 4C's belum ada yang melakukan. Pada penelitian ini, baik guru maupun kepala sekolah terlibat dalam proses supervise maupun aktivitas evaluasi. Supervise menitikberatkan pada perbaikan pembelajaran, sedang evaluasi lebih menekankan dihasilkannya rekomendasi, sehingga akan dihasilkan perbaikan secara terus menerus.

Kata Kunci: Model-model evaluasi, teori dan aplikasinya, Evaluasi-Supervisi, MESp 4Cs



BAB 1

PENGERTIAN

PENGUKURAN, ASESMEN,

DAN EVALUASI

Menurut Giffin dan Nix (1991) pengukuran, asesmen, dan evaluasi adalah herarkhi. Pengukuran membandingkan hasil pengamatan dengan kriteria, asesmen menjelaskan dan menafsirkan hasil pengukuran, evaluasi adalah penetapan nilai atau implikasi suatu perilaku.

1. Pengukuran

Pengukuran adalah penetapan angka (kuantitatif) dengan cara yang sistematis untuk menyatakan keadaan individu. Keadaan individu berupa kemampuan kognitif, afektif dan psikomotor. Juga kemampuan mengendalikan emosi. Tujuannya adalah untuk menggambarkan karakteristik dari suatu objek atau yang diukur, sehingga dihasilkan suatu besaran dalam bentuk angka.

Kesalahan dalam pengukuran harus diperkecil terutama pada ilmu sosial, bidang humaniora dan pendidikan. Kesalahan dapat terjadi diantaranya disebabkan oleh alat ukur, cara mengukur, keadaan objek yang diukur. Alat ukur dapat berupa tugas rumah, kuis, ulangan, dan harus memiliki bukti kesahihan

dan kehandalan. Andal/reliabel adalah ketetapan, sedangkan sah/valid adalah ketepatan.

Andal/reliable adalah ajek, konsisten, stabil. Intinya adalah sejauh mana alat ukur tersebut memberikan hasil pengukuran yang relative sama untuk objek yang sama selama pada kondisi yang relative tidak berubah. Sah/valid adalah sejauh mana alat ukur yang digunakan mampu mengukur dengan cermat, sesuai fungsi ukurnya. Hasil ukur dari pengukuran, harus merupakan besaran yang dapat mencerminkan secara tepat keadaan yang sesungguhnya dari apa yang diukur, (Azwar, 2009).

Validitas suatu tes dapat dinyatakan dalam 3 (tiga) macam, yaitu: (1) Validitas isi (content validity); (2) validitas konstruk (construct validity), dan (3) validitas empiris/kriteria. Validitas isi terkait dengan konten atau materinya dari tes yang disusun, apakah sudah sesuai dengan tujuannya atau sesuai dengan yang harus dikuasi. Validitas konstruk merupakan validitas yang terkait dengan definisi konseptual yang telah ditetapkan. Jika akan mengukur bakat atau motivasi, maka kisi-kisi yang disusun secara konseptual diturunkan dari teori atau konsep yang ilmiah.

Validitas empiris/kriteria diperoleh dari hasil ujicoba secara empiris pada responden yang akan diteliti. Dikenal validitas kriteria internal dan eksternal. Validitas kriteria internal atau sering disebut validitas internal atau validitas butir tes. Validitas internal adalah validitas yang dapat diperlihatkan seberapa jauh hasil ukur butir tes tersebut. Validitas eksternal adalah penggunaan hasil ukur tes yang sudah baku, yang digunakan untuk mengukur. Hasil skor yang didapat dikorelasikan dengan skor hasil tes baku yang dijadikan kriteria,

2. Assesmem

Pengertian assessment dalam konteks pembelajaran adalah data/informasi dari hasil proses pembelajaran. Data yang didapat digunakan untuk mengetahui seberapa baik kinerja/unjuk kerja mahasiswa, dengan cara membandingkan capaian pembelajaran yang telah ditetapkan.

Asesmen mencakup semua cara yang digunakan untuk menilai unjuk kerja individu atau kelompok, meliputi pengamatan atau

laporan diri, dan berkaitan dengan semua proses pendidikan seperti karakteristik peserta didik, metode mengajar, kurikulum, fasilitas dan administrasi. Hasil assessment digunakan sebagai dasar penyusunan rancangan pembelajaran berikutnya, (Smith, M). Hasil assessment juga digunakan untuk solusi tentang apa yang dibutuhkan atau kesulitan dari seseorang, (McLoughlin & Lewis). Hasil assessment juga digunakan untuk mengetahui kekurangan dan kekuatan peserta didik, untuk perbaikan pembelajaran berikutnya.

Tujuan assessment adalah untuk meningkatkan kualitas pembelajaran atau pendidikan yang lebih baik. Kualitas ini meliputi: kualitas pembelajaran juga kualitas sistem penilaiannya, dimana keduanya saling kait-mengkait. Sehingga seorang guru akan melakukan perbaikan baik terkait pembelajarannya juga melakukan perbaikan terhadap system penilaian yang telah dilakukannya.

Bento, 1994, menyampaikan bahwa cara mengubah suatu lembaga melalui program penilaian, dengan kata lain cara melakukan perbaikan pembelajaran dengan cara melakukan penilaian. Menurut (Chittenden, 191) ada 4 (empat) hal kegiatan penilaian proses pembelajaran:

- a. Penelusuran, yaitu kegiatan yang dilakukan untuk menelusuri apakah proses pembelajaran telah berlangsung sesuai yang direncanakan atau tidak.
- b. Pengecekan, yaitu untuk mencari informasi apakah terdapat kekurangan pada peserta didik selama proses pembelajaran. Apakah telah dikuasai oleh peserta didik atau belum.
- c. Pencarian, yaitu untuk mencari dan menemukan penyebab kekuarangan yang muncul selama proses pembelajaran berlangsung. Sehingga dapat mengatasi kendala yang terjadi.
- d. Penyimpulan, yaitu menyimpulkan tentang tingkat pencapaian belajar yang telah dimiliki peserta didik. Penyimpulan dapat digunakan sebagai pelaporan kepada kepala sekolah, ortu dan peserta didik,

Fungsi assessment ada 2 (dua) yaitu (1) fungsi formatif dan (2) fungsi sumatif. Fungsi formatif adalah fungsi perbaikan dalam suatu proses pembelajaran. Suatu assessment yang digunakan sebagai dasar penyempurnaan proses pembelajaran. Fungsi

Sumatif adalah assessment yang digunakan untuk menentukan kelulusan peserta didik dalam suatu subjek yang ditetapkan.

3. Evaluasi

Evaluasi merupakan satu rangkaian kegiatan/usaha atau proses dalam menentukan nilai-nilai, meningkatkan kualitas, kinerja atau produktivitas suatu lembaga dalam melaksanakan programnya. Evaluasi menurut Giffin & Nix (1991) adalah *judgment* terhadap nilai atau implikasi dari suatu pengukuran. Tyler (1950) evaluasi adalah proses penentuan sejauh mana tujuan pendidikan telah dicapai. Evaluasi memuat informasi dan kebijakan tentang pelaksanaan dan keberhasilan suatu program yang selanjutnya digunakan untuk menentukan kebijakan berikutnya.

Berdasarkan pengertian diatas, maka evaluasi adalah proses pemberian nilai berdasarkan data kuantitatif atau diskripsi kualitatif hasil penilaian terhadap suatu hal atau objek yang berdasarkan pada acuan-acuan tertentu untuk menentukan tujuan tertentu, dan dihasilkan untuk keperluan pengambilan keputusan. Terkait evaluasi, secara lengkap akan dibahas dalam buku monograf ini.

PENGERTIAN EVALUASI

Evaluasi dapat didefinisikan sebagai penilaian tujuan prestasi melalui pengumpulan dan analisis data yang berguna untuk membuat keputusan pada nilai dari sebuah program (Hamm, 1985: 256-622; Stake, 1967: 2-4). Stufflebeam & Shinkfield (1985:159) menyatakan bahwa: *“Evaluation is the process of delineating, obtaining, and providing descriptive and judgmental information about the worth and merit of some object’s goals, design, implementation, and impact in order to guide decision making, serve needs for accountability, and promote understanding of the involved phenomena”*. Bahwasanya evaluasi adalah proses menggambarkan, memperoleh, dan memberikan informasi deskriptif dan penilaian tentang nilai dan kebaikan dari tujuan beberapa objek, desain, implementasi, dan dampak untuk memandu pengambilan keputusan, melayani kebutuhan akuntabilitas, dan mempromosikan pemahaman tentang fenomena yang terlibat. Stufflebeam (1973: 3-5), memandang evaluasi sebagai suatu proses memberikan informasi yang berarti dan berguna sebagai alternatif keputusan. Stufflebeam, menyatakan bahwa tujuan evaluasi adalah untuk menyajikan opsi bagi pengambil keputusan.

Komite studi nasional tentang evaluasi (*National study committee on evaluation*) dari UCLA (Stark & Thomas, 1994: 12),

menyatakan bahwa: *“Evaluation is the process of ascertaining the decision of concern, selecting appropriate information and collecting and analyzing information in order to report summary data useful to decision maker in selecting among alternatives”*. Evaluasi merupakan suatu proses atau kegiatan pemilihan, pengumpulan analisis dan penyajian informasi yang sesuai untuk mengetahui sejauh mana suatu tujuan program, prosedur produk atau strategi yang dijalankan telah tercapai, sehingga bermanfaat bagi pengambilan keputusan serta dapat menentukan beberapa alternatif keputusan untuk program selanjutnya.

Istilah yang sering digunakan dalam evaluasi adalah: tes (*test*), pengukuran (*measurement*), dan penilaian (*assessment*). Tes merupakan cara untuk menaksirkan besarnya kemampuan seseorang secara tidak langsung, yaitu melalui respon seseorang terhadap stimulus atau pertanyaan (Mardapi, 2008:2). Pengukuran adalah sebagai penetapan angka dengan cara yang sistematis untuk menyatakan keadaan individu (Allen & Yen, 1979). Penilaian menurut *The Task Group on Assessment and Testing* (TGAT), adalah semua cara yang digunakan untuk menilai unjuk kerja individu atau kelompok (Griffin & Nix, 1991:3).

Menurut Fitzpatrick, Sanders dan Worthen (2011: 7) evaluasi diartikan sebagai berikut : *“evaluation as the identification, clarification, and application of defensible criteria to determine an evaluation object's value (worth or merit) in relation to those criteria”*. Adalah merupakan kegiatan untuk menentukan suatu nilai objek (berharga atau pantas diterima) dengan melakukan identifikasi, klarifikasi dan aplikasi dari kriteria-kriteria yang ada. Evaluasi menurut Bennett (2006: 6) didefinisikan sebagai : *“evaluation involves learning about a new programme through gathering information, and that this information should be linked to decision making”*, evaluasi merupakan kegiatan mempelajari program melalui pengumpulan informasi, dan informasi yang diperoleh nantinya akan dijadikan landasan untuk membuat keputusan.

Selanjutnya Kosecoff dan Fink (1985: 20) mengatakan: *“Evaluation is a set of procedures to appraise a program's merit and to provide information about its goals, expectations, activities,*

outcomes, impact and cost". Disini seorang evaluator menggunakan prosedur tertentu untuk menilai jasa program dan menyediakan informasi tentang tujuan, harapan, aktivitas, hasil, dampak dan biaya. pengertian Evaluasi adalah " *evaluation is the systematic assessment of the worth or merit of some object*" (Stufflebeam & Shinkfield, 1985: 3), sementara Rossi (2004: 15) menyatakan bahwa: "*evaluation is a systematic method for collecting, analyzing, and using information to answer basic questions about project, policies and programs*". Maka dapat dipahami evaluasi program adalah kegiatan yang diadakan untuk seberapa tingkat keberhasilan dari suatu kegiatan yang telah direncanakan dengan cara yang sistematis, tujuannya untuk memberikan masukan kepada pembuat kebijakan. Menurut Weiss (1972: 4) "*the purpose of evaluation research is to measure the effects of a program against the goals it set out accomplish as a means of contributing to subsequent decision making about the program and improving future programming*", evaluasi adalah untuk mengukur dampak sebuah program, membandingkan dengan tujuan yang telah ditetapkan yang nantinya digunakan sebagai bahan pertimbangan dalam mengambil keputusan terhadap program tersebut dan meningkatkan program masa yang akan datang.

Evaluasi program adalah sebuah kegiatan yang menggunakan prosedur-prosedur tertentu dengan mengumpulkan data-data serta mengidentifikasinya, mengklarifikasi dan melihat bagaimana aplikasinya dari masing-masing criteria yang ada, dari informasi yang didapat inilah maka seorang evaluator akan dapat menilai suatu program, apakah program sudah berjalan sesuai dengan rencana, atau perlu adanya perbaikan-perbaikan, atau mungkin saja program itu dihentikan, tentu saja informasi ini juga sangat berguna bagi pembuat keputusan.

1. Pendefinisian Evaluasi Secara Lebih Luas

Pendefinisian secara lebih luas, adalah yang lebih memperluas definisi evaluasi di atas penilaian nilai. Jika pendefinisian didasarkan pandangan bahwa evaluasi dilihat sebagai kegiatan produksi pengetahuan, maka akan dibatasi oleh kebutuhan untuk memberikan informasi terkait program kepada audiens yang

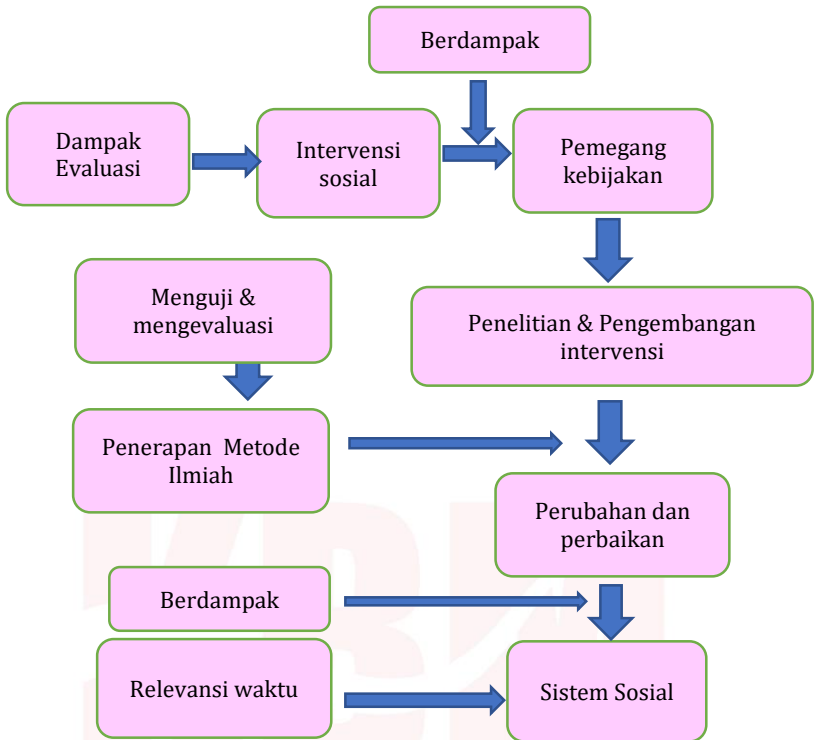
diidentifikasi yang memiliki kepentingan yang sah dalam program itu. Peran evaluasi yang diperluas ini digambarkan adanya definisi kecil yang kemudian diperluas.



Gambar 1. Definisi Evaluasi (Sumber: Owen, 2004)

Misalnya: memberikan informasi tentang nilai suatu program digambarkan pada gambar 1. Menunjukkan bahwa penilaian nilai program hanyalah bagian, dan merupakan bagian penting, dari apa yang dilakukan oleh evaluator. Elemen penting dari kedua definisi tersebut adalah bahwa pengetahuan yang dihasilkan harus diambil dari penyelidikan sistematis; yaitu, harus merupakan hasil penerapan kanon penelitian ilmu sosial. Kanon menurut KBBI adalah komposisi music yang dimainkan secara bersahutan-sahutan. Hubungan dengan pengertian diatas adalah bahwa hasil penerapan evaluasi menghasilkan nilai yang berdimensi banyak dan berdampak banyak. Penilaian perspektif nilai tampaknya didasarkan pada beberapa asumsi tentang penciptaan dan penggunaan intervensi program dalam lingkungan sosiopolitik.

- 1) Bahwa evaluasi dampak akan mengetahui apakah **intervensi** tersebut berhasil atau tidak, yang berhasil akan disebarluaskan ke seluruh sistem sosial, yang mengarah pada perbaikan luas dari masalah yang sedang ditangani. Hal ini menyiratkan bahwa pengembangan intervensi sosial yang dapat mempengaruhi pemegang kebijakan, serta konsisten dengan perspektif Penelitian dan Pengembangan tentang perubahan dan perbaikan. Al hasil akan berdampak pada sistem sosial yang akan selalu siap sampai dimana program dapat divalidasi secara ilmiah, dan bergantung pada penerapan metode ilmiah.

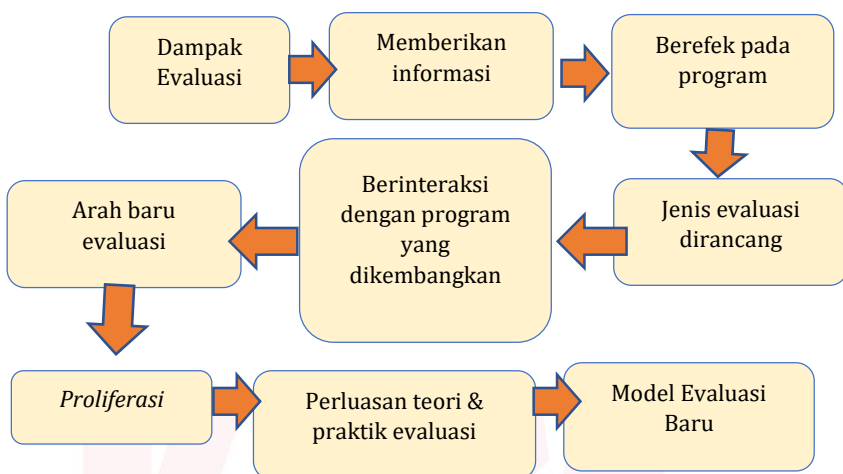


Gambar 2. Dampak Evaluasi teradap Itervensi Sosial

Hal ini membuat gagasan untuk menguji dan mengevaluasi intervensi yang dikembangkan secara terpusat, yang kemudian digunakan secara universal di seluruh sistem sosial, sebuah anakronisme. Menurut KBBI anakronisme adalah: hal ketidakcocokan dengan zaman tertentu, atau penempatan tokoh, peristiwa percakapan, dan unsur latar yang tidak sesuai menurut waktu di dalam karya sastra. Kesimpulannya adalah pengembangan suatu evaluasi harus didasarkan pada relevansinya dengan waktu.

- 2) Hasil penting dari perkembangan ini berarti bahwa lebih banyak perhatian perlu diberikan pada kebutuhan informasi dari mereka yang memiliki tanggung jawab untuk merencanakan dan melaksanakan program sosial. Beberapa evaluator bergerak untuk memberikan informasi ini. Evaluasi dampak memiliki efek pada program yang disampaikan, jenis

evaluasi lainnya dirancang untuk berinteraksi dengan program saat itu terjadi.

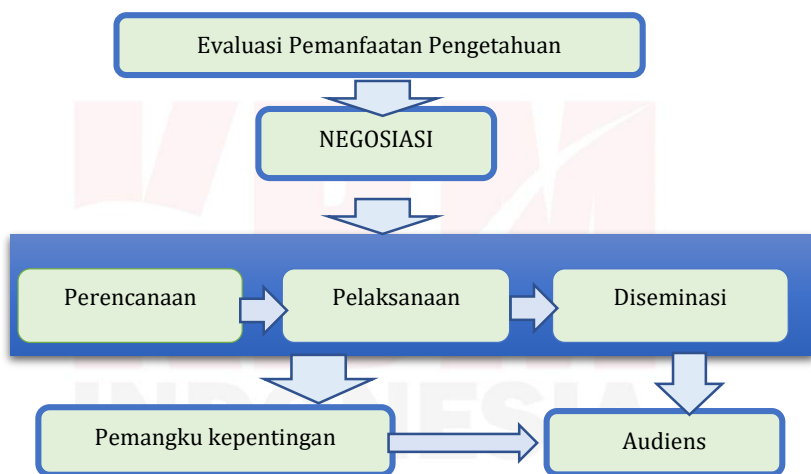


Gambar 3. Dampak Evaluasi Memberikan Informasi

Hasilnya telah menjadi proliferasi "model evaluasi," yang masing-masing umumnya dikaitkan dengan penilai. Hal ini memberi arah baru evaluasi (Stufflebeam, 2001). Terjadilah perluasan teori dan praktik evaluasi, dan adanya model baru yang sesuai dengan kebutuhan praktis untuk evaluasi di tempat kerja mereka. *Proliferasi* menurut KBBI adalah pembiakan yang subur atau fase sel saat mengalami pengulangan siklus sel tanpa hambatan.

- 3) Ada tantangan untuk memberikan beberapa struktur menyeluruh ke dunia teoretis yang penuh kompleksitas. Salah satu bahayanya adalah praktisi baru akan menggunakan salah satu model ini dan menggunakannya dalam situasi apa pun, dan terlepas dari kurangnya kesesuaian antara model yang diusulkan dan kebutuhan pengetahuan para pemangku kepentingan. Perkembangan metamodel evaluasi yang dirancang untuk memberikan solusi terkait beberapa kekhawatiran.

- 4) Setelah produk evaluasi dihasilkan maka ada hal mendasar adalah perlunya pengetahuan untuk tujuan yang mampu mendorong suatu pengambilan keputusan. Bagaimana caranya agar dapat mendorong penggunaan oleh pemangku kepentingan. Penggunaannya dapat ditingkatkan melalui negosiasi dengan pemangku kepentingan sebelum data evaluatif dikumpulkan (Owen, 1999). Keterlibatan pemangku kepentingan sekarang diakui sebagai hal yang esensial sebagai aspek praktik evaluasi dalam konteks penelitian kebijakan (Weiss, 1983).



Gambar 4. Evaluasi Sebagai Pemanfaatan Pengetahuan

Evaluasi pemanfaatan pengetahuan hasil evaluasi melibatkan tiga tahap:

- Negosiasi terkait rencana evaluasi dengan pemangku kepentingan.
- Mengakui implementasi desain evaluasi yang direncanakan.
- Diseminasi temuan evaluasi kepada audiens yang diidentifikasi selama negosiasi (Owen & Rogers, 1999)

Pentingnya pemanfaatan produk informasi perlu dipikirkan sejak awal evaluasi. Negosiasi telah didefinisikan dalam literatur resolusi konflik sebagai “diskusi antara dua pihak atau lebih dengan tujuan yang jelas untuk menyelesaikan perbedaan

kepentingan dan dengan demikian menghindari konflik sosial (Owen, 2004). Levinger dan Rubin (1994) mendefinisikan negosiasi sebagai "proses di mana dua atau lebih pihak yang saling bergantung" menggunakan tawaran memberi dan menerima dan tawaran balasan dalam upaya membangun penyelesaian yang dapat diterima bersama". Ini melibatkan penilai atau tim evaluasi dalam bekerja dengan pemangku kepentingan yang diidentifikasi untuk menetapkan apa saja parameter yang digunakan. Elemen utama dari negosiasi ini adalah untuk menentukan batasan yang diberikan dan yang ditetapkan seperti: faktor seperti anggaran, waktu pelaksanaan evaluasi program, dan pertimbangan etis lainnya.

2. Metamodel Evaluasi

Tahap negosiasi evaluasi, merupakan alat edukatif, untuk memberikan beberapa tatanan berbagai konsep dan model evaluasi yang ada. Model evaluasi yang akan dikembangkan dilakukan dengan cara memanfaatkan literatur yang ada, pertemuan dengan ahli teori kunci, dan informasi yang diperoleh pada konferensi evaluasi. Model kerangka kerja yang paling dekat yang dibutuhkan adalah model Context Input Process and Product (CIPP), yang dikembangkan oleh Daniel Stufflebeam (Stufflebeam, 1983). Kemudian dikombinasikannya dengan pengetahuan tentang berbagai pekerjaan evaluasi.

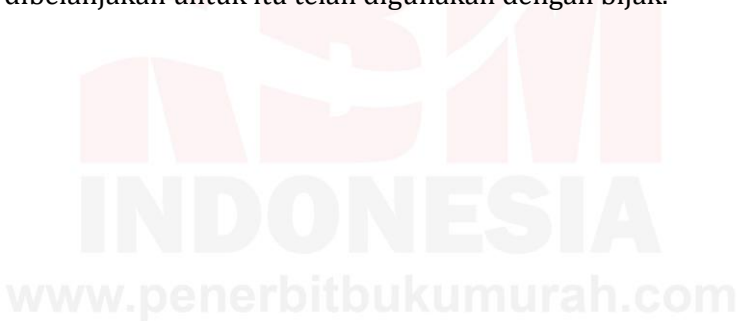
Tabel 1. Bentuk Evaluasi

Bentuk Evaluasi	Proaktif	Klarifikatif	Partisipatori/ Interaktif	Monitoring	Impak
Tujuan	Sintesis	Klarifikasi	Peningkatan	Memeriksa/ Menyempurnakan	Belajar/ Akuntabilitas
asumsi	Apa yang sudah diketahui harus mempengaruhi tindakan	Dasar pemikiran dan desain program perlu ditata	Mereka yang dekat dengan tindakan membutuhkan informasi untuk perubahan yang berkelanjutan	Program perlu dipantau untuk memastikan kualitas	Perlu tahu apa yang berhasil dan mengapa
imperatif	Pentingnya kerangka acuan eksternal	Pentingnya membuat intervensi eksplisit	Pentingnya keterlibatan penyedia	Pentingnya kontrol kualitas	Pentingnya kemampuan transfer: kontribusi untuk pengetahuan yang didanai

Tabel diatas memberikan gambaran umum tentang metamodel, yang meliputi: bentuk evaluasi. Bentuk proaktif, klarifikasi, partisipatif/Interaktif, pemantauan, dan dampak. Apa yang dilakukan evaluator dapat diringkas disini. Selain tujuan, masing-masing memiliki karakteristik, asumsi, dan imperatif. Selanjutnya, diasumsikan bahwa aplikasi dari bentuk tertentu (atau bentuk dalam beberapa kasus) untuk evaluasi akan mendorong pemanfaatan temuan dalam konteks organisasi tertentu.

Sebagai ilustrasi: terkait dengan bentuk dampak, di kolom sebelah kanan tabel. Evaluasi dampak dapat dipertahankan atas dasar kebutuhan untuk mengetahui apakah intervensi berhasil dan mengapa intervensi tersebut berhasil. Dalam istilah pemanfaatan pengetahuan, yang penting adalah bahwa evaluasi dapat mengarah pada dua jenis penggunaan.

- 1) Pertama, informasi tersebut dapat digunakan untuk memberikan keyakinan bahwa program yang sama atau serupa dapat dilaksanakan dengan peluang keberhasilan yang tinggi dalam konteks yang serupa.
- 2) Kedua, evaluasi dapat berkontribusi pada "pengetahuan yang didanai", akumulasi pengetahuan tentang kebijakan atau fenomena umum, misalnya, literatur tentang pelatihan yang efektif. Juga sering ada fokus akuntabilitas yang mendasari evaluasi dampak, untuk menunjukkan bahwa dana yang dibelanjakan untuk itu telah digunakan dengan bijak.





BAB 3

PENDEKATAN MODEL EVALUASI

1. Pendekatan Model Evaluasi

Ownston (2008) menyampaikan tentang strategi evaluasi dengan berdasarkan pada tujuan bahwa informasi kebutuhan audiens. Tujuan dirinci untuk menjembatani kesenjangan antara realita (lapangan) dengan standar, melalui pendekatan evaluasi tertentu. Menurut perkembangannya evaluasi digambarkan oleh Tyler, (1942) bahwa evaluasi berorientasi kepada tujuan. Scriven (1967: 39) mendesak evaluator untuk mengevaluasi dengan melihat maksud/manfaat dan berorientasi kepada *outcome* (hasil).

Weiss (1972: 5) menunjukkan cara model dapat dikembangkan dan diuji untuk menjelaskan bagaimana rantai peristiwa dalam evaluasi mengarah pada tujuan akhir untuk mencapai hasil. Stufflebeam (1973: 3-5) mengusulkan konteks, input, proses, dan produk (CIPP) model, yang menggambarkan empat jenis kegiatan evaluatif. Model CIPP meminta dari sebuah program: Apa yang perlu dilakukan? Bagaimana seharusnya itu dilakukan? Mengapa itu dilakukan? Apakah berhasil? Model CIPP berdamai dengan formatif dan sumatif Scriven.

Cronbach (1982: 51), seorang mahasiswa Tyler, menfokuskan pada proses pengambilan keputusan. Kontribusinya adalah menekankan konteks politik dalam pengambilan keputusan. Keputusan yang lebih mungkin dilakukan adalah dalam suasana politik yang hidup oleh kebijakan yang membentuk komunitas. Program evaluasi menyediakan sarana untuk merekomendasikan tindakan menentukan kelayakan.

Pada saat yang sama para peneliti sedang mengembangkan pendekatan yang berfokus pada bagaimana hasil evaluasi digunakan, yang lain memusatkan usaha mereka pada pengembangan metode yang digunakan. Pendekatan naturalis atau kualitatif, memberikan kebebasan evaluator untuk memilih metode yang digunakan untuk mengumpulkan, menganalisis, dan menafsirkan data mereka. Stake (1977: 372-390) terkenal dengan *evaluasi responsif* sebagai salah satu model. Stake menyampaikan bahwa pendekatan konvensional tidak cukup menerima kebutuhan dari klien evaluasi. Dia menganjurkan agar evaluator menghadiri kegiatan program yang sebenarnya, menanggapi kebutuhan audiens untuk informasi, dan menyajikan perspektif nilai yang berbeda ketika melaporkan keberhasilan dan kegagalan program. Evaluator menggunakan skema pengumpulan data agar tepat.

Dua pendekatan lain yang inters pada diskusi dalam metode naturalistik. Pertama, adalah *connoisseurship model* milik Eisner (1985), yang berakar dalam bidang kritik seni. Modelnya bergantung pada penilaian evaluator untuk menilai kualitas dari sebuah program pendidikan, seperti halnya kritikus seni menilai kompleksitas dari sebuah karya seni. Dua konsep adalah kunci untuk Eisner Model: *educational connoisseurship* dan *educational criticism*. *Educational connoisseurship* melibatkan apresiasi terhadap poin-poin penting dari program pendidikan, bakat yang berasal dari pengalaman evaluator dan latar belakang dalam domain tersebut. Pendidikan kritik bergantung pada kemampuan evaluator untuk verbalisasi fitur program, sehingga mereka yang tidak memiliki tingkat apresiasi sepenuhnya dapat memahami fitur program. Pendekatan kedua adalah *ethnographic evaluation*, pendukung yang percaya dapat menghasilkan lebih bermakna

gambar program pendidikan daripada menggunakan metode ilmiah tradisional (Ownston, 2008). Evaluator etnografi membenamkan diri dalam program yang sedang dipelajari untuk mengambil bagian dari *day-to-day* kegiatan dari individu-individu yang sedang dipelajari. Alat pengumpulan data mereka berupa catatan lapangan, informan kunci wawancara, kasus sejarah, dan survei. Tujuan mereka adalah untuk menghasilkan deskripsi yang kaya dari program dan untuk menyampaikan program penilaian mereka kepada stakeholder.

2. Model-Model Evaluasi

Model evaluasi ialah model desain evaluasi yang dibuat oleh ahli-ahli atau pakar-pakar evaluasi. Model evaluasi merupakan rancangan yang digunakan untuk melakukan evaluasi terhadap suatu program praktik pendidikan. Berbagai dimensi sistem sebagai suatu keseluruhan objek evaluasi: *Two generation of Evaluation Models (Gene V Class). The complete and detailed description of what constitutes of educational program is a concern of the educational system evaluation model.* Hasil yang dicapai, dan berbagai dimensi yang mempengaruhi juga menjadi objek evaluasi.

INDONESIA
www.penerbitbukumurah.com



KLASIFIKASI EVALUASI

Klasifikasi evaluasi dapat dilakukan berdasarkan pada: apa yang dievaluasi, tujuan evaluasi, fokus evaluasi, metode evaluasi, pendekatan evaluasi, lingkup atau tataran yang dievaluasi, dan orientasinya (Owen & Rogers, 1999: 52-63). Berdasarkan hal yang dievaluasi, evaluasi dapat dibagi ke dalam beberapa kelompok: 1) evaluasi kegiatan; 2) evaluasi program; 3) evaluasi kebijakan; 4) evaluasi pengelolaan keuangan; 5) evaluasi pengelolaan sumber daya manusia; 6) evaluasi terhadap sistem dan *governance*; 7) evaluasi terhadap struktur, mekanisme dan prosedur; 8) evaluasi efisiensi, efektivitas, kehematan, kelayakan.

Penggolongan evaluasi berdasarkan tujuan evaluasi meliputi: 1) evaluasi untuk tujuan tertentu, misalnya: untuk mempelajari fakta dan kemungkinan perbaikannya, untuk meningkatkan akuntabilitas, untuk meningkatkan kinerja.; 2) *goal free evaluation* atau evaluasi untuk mencari peluang perbaikan yang tidak ditetapkan terlebih dahulu. Berdasarkan fokus evaluasinya pekerjaan evaluasi dapat dibagi ke dalam lima kelompok: 1) *Input evaluation*; 2) *Process evaluation*; 3) *Output evaluation*; 4) *Outcomes evaluation*; 5) *Impact evaluation*. Berdasarkan pendekatannya, evaluasi dapat dibagi ke dalam: 1) Evaluasi semu; 2) Evaluasi formal; 3) Evaluasi keputusan teoretis. Berdasarkan orientasinya,

evaluasi dapat dikelompokkan ke dalam beberapa kategori sebagai berikut: 1) Evaluasi yang proaktif (*Proactive evaluation*); 2) Evaluasi yang klarifikatif (*Clarificative evaluation*); 3) Evaluasi interaktif (*Interactive evaluation*); 4) Evaluasi monitoring (*Monitoring evaluation*); 5) Evaluasi dampak (*Impact evaluation*); Berikut ini penjelasan masing-masing kategori:

1. Evaluasi yang proaktif (*Proactive evaluation*).

Evaluasi proaktif ini dapat dilakukan sebelum suatu kebijakan/program ditetapkan. Pendekatan-pendekatan kunci :

- a) Perkiraan kebutuhan;
- b) Review riset;
- c) Review: praktik-praktik terbaik (*best practices*).

Dalam pencarian bukti-bukti dapat digunakan teknik-teknik: DELPHI, *Customer Satisfaction Survey*, *Nominal Group Forum*, *Mapping Consep*, *Focus Group*.

2. Evaluasi yang klarifikatif (*Clarificative evaluation*).

Evaluasi klarifikatif ini berfokus pada klarifikasi struktur internal dan fungsi dari suatu program dan kebijakan. Pendekatan-pendekatan kunci:

- a) *Evaluability assessment*;
- b) *Logic development*;
- c) *Acreditation*.

Pendekatan ini cocok untuk program yang mulai dilaksanakan.

3. Evaluasi interaktif (*Interactive evaluation*).

Evaluasi interaktif ini dapat digunakan untuk memperoleh informasi atas implementasi program. Pendekatan penting yang bisa dipakai:

- a) respon evaluasi;
- b) Riset tindakan;
- c) Evaluasi pengembangan;
- d) Evaluasi pemberdayaan.

4. Evaluasi monitoring (*Monitoring evaluation*).

Evaluasi monitoring ini sangat tepat digunakan ketika program sudah dalam pelaksanaan. Evaluasi ini sudah melibatkan pengembangan sistem untuk pemantauan kemajuan program. Indikator kinerja kuantitatif sudah harus digunakan sebagai alat untuk mengorganisasikan data dalam evaluasi monitoring. Pendekatan-pendekatan utama yang dapat dipakai:

- a) *Component analysis*.
- b) Pengukuran kinerja (*Performance assessment*).
- c) *System analysis*.

5. Evaluasi dampak (*Impact evaluation*).

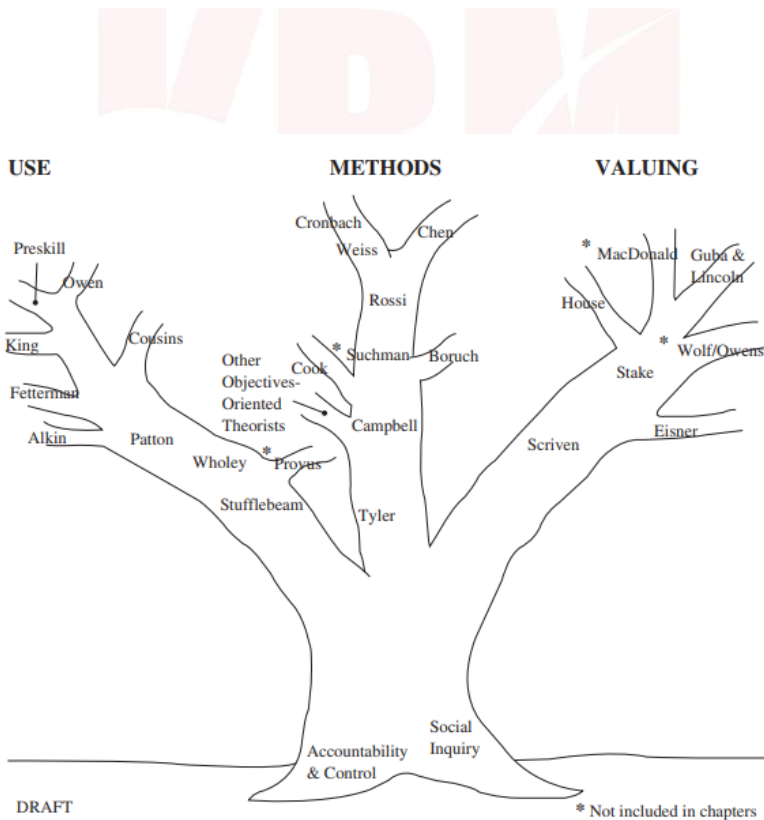
Evaluasi ini digunakan untuk menilai hasil dan dampak program yang sudah mapan. Evaluasi ini dapat digunakan untuk membuat keputusan tentang penghargaan, atau kemanfaatan program. Evaluasi ini disebut juga evaluasi sumatif (*Sumative evaluation*). Pendekatan yang dapat dipakai:

- a) Evaluasi yang berdasarkan sasaran.
- b) *Study process-outcome*.
- c) Evaluasi berdasarkan kebutuhan (*Needs-based evaluation*).
- d) *Goal-free evaluation* (termasuk mengevaluasi dampak yang bersifat positif dan negatif (*unintended impact*) dan tidak hanya *outcome* semata).
- e) *Performance audit*.



BAB 5

PENETAPAN MODEL EVALUASI



Sumber: (Alkin, & Christie, 2004).
Gambar 5. Pohon Teori Evaluasi

Gagasan Alkin untuk memetakan evaluasi yang diilustrasikan dengan pohon teori. Pohon teori ditempatkan pada pandangan bahwa tiga dimensi, digambarkan sebagai cabang, adalah prioritas tertinggi untuk setiap ahli teori, (Alkin, & Christie, 2004).

Pohon teori evaluasi digambarkan batang dan tiga cabang utama pohon. Batangnya dibangun di atas landasan akuntabilitas dan penyelidikan sosial yang sistematis. Kedua bidang ini telah mendukung pengembangan lapangan dengan cara yang berbeda. Kebutuhan dan keinginan untuk akuntabilitas menghadirkan kebutuhan untuk evaluasi. Pentingnya tindakan atau sumber daya yang digunakan dalam pelaksanaan program didukung oleh pemerintah. Sebagai akar evaluasi program, perlunya difikirkan akuntabilitas dengan cara seluas mungkin.

- a) Pengertian akuntabilitas adalah sebuah konsep etika yang berhubungan dengan kemampuan menjelaskan keputusan yang diambil dan aktivitas yang dilakukan. Akuntabilitas bukanlah kegiatan yang membatasi, melainkan dirancang untuk memperbaiki dan memperbaiki program dan masyarakat.
- b) Akar penyelidikan sosial dari pohon berasal dari perhatian untuk menggunakan serangkaian metode yang sistematis dan dapat dibenarkan untuk menentukan akuntabilitas. Sedangkan pertanggungjawaban memberikan alasan, terutama dari penyelidikan sosial bahwa model evaluasi telah diturunkan.

Ada 3 cabang utama yaitu: (a) Cabang Metode. (b) penilaian. (c) cabang kegunaan.

- a) Cabang utama pohon merupakan kelanjutan dari batang penyelidikan sosial. Hal ini mengindikasikan perlunya metode penelitian, dan konstruksi pengetahuan.
- b) Cabang menilai. Cabang penilaian dengan tegas menetapkan peran vital evaluator dalam penilaian. Penilaian menempatkan nilai sebagai data yang merupakan komponen paling penting sebagai kerja evaluator. Peran evaluator adalah memasukkan

secara sistematis, memfasilitasi penempatan nilai oleh orang lain.

- c) Cabang penggunaan. Orientasi terhadap evaluasi dan pengambilan keputusan, bahwa intinya pekerjaan yang dilakukan oleh para ahli teori di cabang ini mengungkapkan perhatian terhadap cara di mana informasi evaluasi akan digunakan dan berfokus pada mereka yang akan menggunakan informasi tersebut.

Penjelasan Alkin yang menempatkan pohon teori pada pandangan tiga dimensi, yang digambarkan sebagai cabang. Ada 3 (tiga) pengaruh posisi teoretis evaluasi yaitu: (1) antropologis/etnografis (Spindler dan Pelto menjadi dua yang paling menonjol), (2) keadilan sosial (House menjadi yang paling menonjol), dan (3) pemanfaatan evaluasi (Patton dan Alkin menjadi dua yang paling menonjol).

- a) Penekanan pada antropologi/etnografi pendidikan berpusat pada metode. Titik fokus seperti itu akan berarti bahwa elemen evaluasi yang paling penting adalah penggunaan metode etnografi. Intinya menjadi preferensi yang kuat untuk menggunakan metode kualitatif dalam evaluasi.
- b) Penekanan pada keadilan sosial akan menyiratkan bahwa tujuan evaluasi adalah untuk memberi nilai pada hasil evaluasi yang dalam beberapa hal mendukung cita-cita keadilan sosial. Keadilan sosial dilayani melalui penghapusan ketergantungan pada orang lain untuk terlibat dalam kegiatan yang mengarah pada perbaikan program; itu tidak berarti menilai suatu program berdasarkan apakah memenuhi kriteria atau tidak berasal dari seperangkat cita-cita yang adil secara sosial.
- c) Penekanan pada pemanfaatan menyiratkan adanya pemanfaatan temuan evaluasi. Dengan kata lain memberikan kontribusi di lapangan dan mendaftarkan manfaat dari kontribusi tersebut. Pemanfaatannya adalah penggunaan proses, yaitu bahwa proses melakukan evaluasi mengarah pada pemberdayaan. Observasi partisipan sebagai elemen kunci dari evaluasi pendidikan etnografi dan menekankan pentingnya menggunakan perspektif budaya untuk

menafsirkan dan menganalisis data. Evaluasi memberi kemanfaatan pengetahuan dalam evaluasi yang mengakui keunggulan negosiasi antara evaluator dan pemangku kepentingan sebagai tahap pertama evaluasi yang membentuk cara evaluasi selanjutnya akan dilakukan. sebagai gambaran pentingnya evaluasi bagi pembuat keputusan program (Alkin, 1985).

Pengetahuan yang dihasilkan dalam studi evaluasi kemungkinan untuk mengklasifikasikan pengetahuan yang dihasilkan oleh evaluasi, yang secara epistemologis dihasilkan klasifikasi evaluasi. Gaya atau bentuk evaluasi yang berbeda menghasilkan produk pengetahuan yang unik untuk setiap klaster. Pengembangan metamodel evaluasi dapat dikembangkan berdasarkan tujuan evaluasi, (Owen, 2004). Bentuk evaluasi yang didasarkan pada tujuan, dihasilkan klasifikasi sebagai berikut:

- (1) Evaluasi proaktif: dilakukan sebelum program dirancang. Tujuannya adalah untuk memberikan pengetahuan yang mengidentifikasi dan mendokumentasikan dimensi penting dari suatu program serta secara eksplisit mengarah pada pemangku kepentingan.
- (2) Evaluasi Partisipatif / Interaktif, yang berlangsung selama penyampaian suatu program. Tujuannya adalah untuk memberikan pengetahuan bagi keputusan yang berkaitan dengan perbaikan berkelanjutan dengan melibatkan penyedia program dalam proses evaluasi.
- (3) Evaluasi pemantauan, yang berlangsung selama umur program yang sudah mapan dan berkelanjutan. Tujuannya adalah untuk memberikan pengetahuan untuk memeriksa bahwa program tersebut “on track” dan untuk memberikan dasar penyempurnaan.
- (4) Evaluasi dampak, yang digunakan untuk menilai dampak dari suatu program yang telah ditetapkan. Tujuannya adalah untuk menentukan efek program dalam kaitannya dengan kriteria yang dipilih untuk menilai keberhasilannya.

Lebih mengutamakan interpretasi data melalui mata mereka yang terlibat dalam evaluasi daripada melalui mata peneliti (evaluator). Alkin menyampaikan fokus utama evaluasi adalah etnografi yang menjadi metodologi utama untuk melakukan evaluasi untuk motivasi. Perlunya tindakan pemberdayaan yang berfokus pada proses yang terlibat dalam evaluasi, yaitu dengan cara melatih individu untuk terlibat dalam kegiatan evaluasi, maka mereka menjadi berdaya.

Perbaikan sosial sebagai tujuan akhir evaluasi dan menyajikan sudut pandang didasarkan pada apa yang mereka sebut sebagai "filsafat realis akal sehat. Mempertahankan bahwa evaluasi yang akan dilakukan tergantung pada tujuan tertentu yang akan dilayani. Ada empat tujuan:

- a) Menilai prestasi dan nilai (selaras dengan valuing).
- b) Program dan peningkatan organisasi (kegunaan).
- c) Pengembangan pengetahuan (metode).
- d) Pengawasan dan kepatuhan (akuntabilitas).

Evaluasi pemanfaatan, dibagi menjadi 2 jenis. Scriven (2007) menekankan pada "Evaluasi Pemanfaatan" secara praktis; sedangkan Preskill dan Caracelli (1997); Weiss dkk. (2005); Patton (1997), dan Kanjanawasee, Sirichai (2009) berfokus pada definisi aplikasi dari kedua pendekatan, dan praktek. Akibatnya, definisi Pemanfaatan Evaluasi, disimpulkan sebagai penerapan temuan atau Hasil Evaluasi dalam meningkatkan praktik kerja yang mengarah pada perubahan proyek atau kebijakan pelaksanaan ke arah yang lebih baik serta melayani kebutuhan administrator proyek, pendukung, serta orang-orang terkait yang mengarah pada pengambilan keputusan, manajemen, atau pengembangan dalam peningkatan pelaksanaan atau salah satu kegiatan baik dalam pendekatan, maupun praktek.

Faktor-faktor yang digunakan dalam pengukuran Evaluasi Pemanfaatan, didasarkan pada Pemanfaatan Evaluasi dari pendekatan Alkin et al. (1979); Weiss (1980); Patton (1997), dan Kanjanawasee, Sirichai (2009) yang mencakup beberapa perbedaan sebagai: pendekatan yang menggunakan penggunaan

simbolik, penggunaan yang sah, dan penggunaan instrumental yang digabungkan sebagai Evaluasi Tindakan Pemanfaatan.

Peneliti mengklasifikasikan menjadi 4 faktor untuk mengukur Evaluasi Pemanfaatan secara lebih jelas dan konkrit seperti: 1) Penggunaan Konseptual, 2) Penggunaan Sah, 3) Penggunaan Simbolik, dan 4) Penggunaan instrumental. Berdasarkan pertimbangan dalam Model Pengukuran Evaluasi Pemanfaatan, dari 4 faktor ditemukan bahwa setiap faktor harus dipelajari indikator pada setiap faktor. Indikator harus mencakup berbagai masalah aplikasi di penilaian eksternal berdasarkan batasan penelitian. Rinciannya dapat disimpulkan menjadi total 4 faktor, dan 18 indikator, (Tochot, Junpeng & Makmee, 2012).

Pohon evaluasi telah memberikan penjelasan 3 (tiga) cabangnya yang meliputi: evaluasi penggunaan, evaluasi berdasarkan metode, dan evaluasi berdasarkan nilai. Berikut ini dijelaskan penulis terkait evaluasi diatas adalah sebagai berikut:

1. Penulis yang focus pada cabang evaluasi penggunaan:

a) Daniel Stufflebeam:

- Teori berorientasi keputusan Sangat penting untuk melakukan evaluasi yang dirancang khusus untuk membantu pemangku kepentingan program utama dalam pengambilan keputusan program.
- Model CIPP - fokus pada proses pengambilan keputusan: konteks-input-proses-produk (kelanjutan atau fokus ulang program).
- Desain evaluasi adalah PROSES bukan produk, ini adalah aliran informasi berkelanjutan bagi para pengambil keputusan
- Faktor kunci dari penggunaan evaluasi - interaksi reguler antara pemangku kepentingan (misalnya panel pemangku kepentingan yang representatif)
- Semakin banyak pemangku kepentingan yang terlibat, semakin banyak perspektif nilai yang diwakili, sehingga evaluasi yang lebih komprehensif.
- Standar Evaluasi Program:
 - utilitas (melayani kebutuhan info pengguna yang dituju)

- kelayakan (realistis, bijaksana, diplomatis, hemat)
- kepatutan (dilakukan secara legal, etis, dengan memperhatikan kesejahteraan mereka yang dievaluasi dan dipengaruhi oleh hasil)
- akurasi (mengungkapkan & menyampaikan info teknis yang memadai untuk menentukan nilai & nilai program)

b) Malcolm Provus:

- Evaluasi adalah proses berkelanjutan yang dirancang untuk membantu administrator program Pelayanan administrasi dalam pengelolaan pengembangan program melalui pengambilan keputusan yang baik.
- Tahapan evaluasi DISCREPANCY:
 - definisi - tentukan tujuan, proses, sumber daya
 - instalasi - mengidentifikasi perbedaan dalam implementasi prog
 - proses - menentukan perluasan pencapaian tujuan jangka pendek
 - produk - sejauh mana pencapaian tujuan akhir atau tujuan akhir
- Evaluasi perbedaan – penekanan yang kuat pada desain program, diskusi interaktif & keterlibatan staf program dalam tahap definisi

c) Joseph Wholey

- Evaluasi adalah pelayan manajemen, terutama pembuat kebijakan (fokus pada evaluasi program skala besar)
- Fokus pada pemanfaatan instrumen evaluasi jangka pendek.
- 4 tahap model "pembelian informasi berurutan":
 - 1) penilaian evaluabilitas - menilai biaya, ruang lingkup & kemungkinan penelitian, kegunaan bagi manajer prog
 - 2) evaluasi umpan balik cepat - fokus pada info yang masih ada & mudah dikumpulkan

- 3) pemantauan kinerja - mengukur kinerja prog, dibandingkan dengan kinerja sebelumnya / yang diharapkan
- 4) evaluasi intensif- perbandingan, kelompok kontrol untuk memperkirakan efektivitas prog, penyebab hasil yang diamati.

d) Michael Patton:

- Kewajiban evaluator adalah untuk memastikan bahwa pemanfaatan terjadi tidak hanya untuk pengambil keputusan tetapi untuk spektrum pemangku kepentingan yang lebih luas
- Evaluator harus mencari siapa yang mungkin menjadi "pengguna utama yang dituju".
- Diagram alir evaluasi yang berfokus pada pemanfaatan (UFE):
 - Identifikasi pengguna yang dituju
 - Komitmen pengguna terhadap fokus & pemanfaatan yang diinginkan
 - Keterlibatan dalam metode, desain, pengukuran
 - Keterlibatan dalam interpretasi aktif & direktif dari temuan & pengambilan keputusan
 - Membuat keputusan untuk diseminasi lebih lanjut
- Dua faktor kunci dalam pemanfaatan:
 1. Faktor pribadi (mengidentifikasi & melibatkan orang-orang kunci yang memiliki kepentingan dalam evaluasi & secara pribadi peduli dengan temuan
 2. Evaluator harus secara aktif melibatkan pengguna potensial untuk mendorong penerimaan
 3. Evaluator harus aktif-reaktif-adaptif
- Kekhususan pendekatan Patton:
 - perhatian dengan mengidentifikasi pengguna yang dituju
 - kesediaan untuk fleksibel dalam memodifikasi pertanyaan evaluasi yang telah ditentukan

sebelumnya, masalah untuk melayani kebutuhan pengguna dengan lebih baik

e) Marvin Alkin:

- Evaluasi berorientasi pengguna - model CSE (Pusat Studi Evaluasi)
- Evaluator tidak harus menilai agen. Mereka harus bekerja dengan pengguna utama pada awal proses evaluasi, untuk membangun sistem nilai untuk menilai data hasil

f) J. Bradley Cousins:

- Faktor pemanfaatan utama adalah untuk mendapatkan dukungan. Dan cara terbaik untuk menjamin ini adalah dengan melibatkan personel prog dalam evaluasi
- Evaluasi partisipatif:
 - melibatkan pengguna awal yang dituju
 - atur mereka dalam kelompok
 - partisipasi mereka harus terstruktur, berkelanjutan dan aktif
- Pengguna & evaluator utama adalah kolaborator dalam proses evaluasi
- Evaluasi partisipatif berusaha untuk memahami prog dengan maksud yang diungkapkan untuk menginformasikan & meningkatkan pelaksanaannya. Hal ini didasarkan pada kerangka pembelajaran organisasi
- Peran evaluator adalah untuk melatih & mengawasi personel lembaga dalam keterampilan evaluasi teknis. Evaluator menjadi konsultan yang bertanggung jawab atas kegiatan teknis (desain instrumen, analisis data, pelaporan)

g) Hallie Preskill:

- Fokus pada pembelajaran & pengembangan organisasi Menyesuaikan evaluasi agar sesuai dengan kebutuhan pengguna utama dan mendapatkan dukungan dari personel prog yang berpartisipasi dalam evaluasi

- Pembelajaran transformasional - proses di mana individu, tim, organisasi mengidentifikasi, memeriksa, dan memahami kebutuhan informasi untuk memenuhi tujuan mereka
- Evaluator dapat memfasilitasi TL dengan:
 - 1) mendekati organisasi menggunakan pendekatan klinis (fokus pada kebutuhan saat ini, pengaturan kontekstual, lingkungan historis organisasi, dialog & respons)
 - 2) mencakup batas-batas tradisional antara evaluator & staf program (penilai panduan & memfasilitasi proses pembelajaran)
 - 3) mendiagnosis kapasitas organisasi untuk belajar (mengamati gaya komunikasi, pengaturan dan kepemimpinan organisasi, dan tingkat penerimaan hasil).

h) Jean King:

- Interactive Evaluation Practice (IEP) - mendorong partisipasi & memperoleh penggunaan Menyesuaikan evaluasi secara kolaboratif dengan pengguna yang dituju utama untuk meningkatkan peluang pemanfaatan
- Faktor kunci keberhasilan pemanfaatan:
 - menerima struktur kekuasaan
 - berbagi makna pengalaman di antara peserta
 - relawan dan pemimpin
 - cukup waktu & sumber daya
 - tingkat kepercayaan interpersonal & organisasi yang tinggi
- Evaluator harus membangun dan menumbuhkan kepercayaan selama proses evaluasi
- Tiga peran evaluator:
 - 1) pengambil keputusan - dipaksa untuk membuat keputusan untuk memajukan proses
 - 2) aktor – aktor dalam acara teatrikal dari proses evaluasi. Terkadang sebagai pemimpin, manajer atau penasihat yang bijaksana

3) praktisi reflektif - membantu membimbing & menilai kemajuan evaluasi.

i) John M. Owen:

- Perbedaan antara konsultan dan evaluator
 - konsultan memberikan resep & rekomendasi
 - evaluator memberikan deskripsi & penilaian evaluatif
- Evaluator harus fleksibel dan adaptif terhadap kebutuhan dan kekhawatiran pemangku kepentingan (itulah yang membedakan Owen dari evaluator pengembangan organisasi lainnya)
- Keterlibatan pemangku kepentingan paling penting selama tahap perencanaan & negosiasi

j) David Fetterman:

- Evaluasi Pemberdayaan. Tujuannya adalah untuk mendorong penentuan nasib sendiri. Proses yang mendorong penentuan nasib sendiri di antara penerima evaluasi prog, termasuk pelatihan, fasilitasi, advokasi, penerangan & pembebasan
- Peran evaluator
 - Pelatih, fasilitator tambahan, memberikan pengetahuan dan alat kepada klien untuk penilaian diri dan akuntabilitas yang berkelanjutan.
 - Klien pada dasarnya melakukan evaluasi mereka sendiri.
- Dua bentuk evaluasi Pemberdayaan
 1. Evaluator mengajarkan peserta prog untuk melakukan evaluasi mereka sendiri
 2. Evaluator berfungsi sebagai coach untuk membantu orang lain melakukan evaluasi sendiri value & worth tidak statis. Evaluasi adalah proses berkelanjutan untuk mengakomodasi nilai & tujuan yang berbeda.
- Perbedaan antara evaluasi emovering dan partisipatif:
 - Empowering evaluation - memberdayakan mereka yang tertindas, sifatnya emansipatoris

- Evaluasi partisipatif- tingkatkan penggunaan melalui keterlibatan.

2. Penulis yang fokus pada Cabang evaluasi Metode:

a) Donald Campbell

- Terkait metode:
 - 1) Validitas internal
 - 2) Validitas eksternal
 - 3) Eksperimen tidak sempurna - desain kuasi-eksperimental
- Karakteristik utama dari kuasi-eksperimen - kurangnya pengacakan
- Metode kuantitatif - berguna untuk: evaluasi implementasi, identifikasi & interpretasi efek samping
- Evaluator adalah penjaga kebenaran

b) Edward Schuman

- Evaluasi merupakan bentuk penelitian evaluasi (common valuing) vs penelitian evaluatif (penggunaan metode ilmiah 7 teknik)
- Kriteria administratif bagian dari penilaian kelayakan melakukan studi

c) Thomas Cook

- Eksperimen semu - pengelolaan pra-tes & perbandingan kelompok tetapi tidak ada pengacakan
- Serangan langsung pada desain eksperimental - pilih yang sesuai untuk evaluasi tertentu & konteksnya, tidak sama ditetapkan untuk semua evaluasi
- Libatkan pemangku kepentingan untuk menentukan pertanyaan evaluasi

d) Robert Boruch

- Uji lapangan acak - metode terbaik untuk semua evaluasi (mereka berbeda dari eksperimen semu)
- Evaluasi adalah penelitian ilmiah konvensional

e) Peter Rossi

- Evaluasi berbasis teori - konstruksi teori program terperinci yang kemudian digunakan untuk memandu evaluasi
- Pekerjaan awal - fokus pada eksperimental & kuasi-eksperimental metode, validitas internal, validitas eksternal, dll

f) Lee J. Cronbach

- Aksen pada penggunaan pencerahan. Evaluasi bukanlah keputusan go/no go. Ini adalah bagian integral dari penelitian kebijakan yang berfokus pada komunitas pembentuk kebijakan dan memerlukan akomodasi politik potensial
- Juga menegaskan peran evaluator dalam menyediakan data bagi pengambil keputusan lokal untuk penggunaan instrumental
- Domain evaluasi: UTOS. unit(populasi)-perawatan-pengamatan(hasil)-pengaturan
- Faktor utama pemanfaatan - validitas eksternal. Fokus pada sejumlah masalah daripada mencapai absolut akurasi pada sejumlah kecil masalah
- Evaluator pada dasarnya adalah seorang pendidik
- Evaluator memberikan informasi kepada pembaca sehingga mereka dapat memperhitungkan ketika membuat penilaian mereka sendiri

3. Penulis yang fokus pada Cabang evaluasi Nilai:

a) Michael Scriven

- Peran evaluator adalah menilai dan dinilai. Dia melayani kepentingan publik.
- Buruk itu buruk, baik itu baik dan itu adalah peran evaluator untuk memutuskan yang mana
- Evaluator mengidentifikasi pesaing kritis - alternatif bersaing dan pertimbangkan mereka
- Penilaian harus diakhiri dengan penilaian sederhana baik/salah.

- Evaluator mengembangkan kriteria penilaian berdasarkan alternatif bersaing yang diperiksa
- Program dikatakan baik jika dapat memenuhi KEBUTUHAN (ditentukan oleh penilaian kebutuhan)
- Metode Modus Operandi (MO) - rantai sebab akibat yang khas
- Advokat evaluasi tanpa tujuan
- Teknik kuantitatif sangat penting

b) Elliot Eisner

- Teknik kualitatif sangat penting.
Evaluasi membutuhkan peta interpretatif yang canggih, untuk memisahkan apa yang sepele dari apa yang penting dan untuk memahami makna dari apa yang diketahui
- Keahlian - untuk memiliki pengetahuan tentang apa yang pernah dilihat, kemampuan untuk membedakan dan memahami hal-hal halus, pengalaman dalam observasi & interpretasi
- Kritik - terdiri dari:
 - Deskripsi kritis tentang peristiwa, reaksi, dll
 - Interpretasi kritis tentang apa yang penting (berdasarkan pengalaman evaluator dan masuk ke pikiran pemangku kepentingan)
- Peran kunci evaluator adalah menilai dan menilai (juga memutuskan mana pertanyaan untuk ditanyakan - apa yang penting)

c) Thomas Owen / Robert Wolf

- Model evaluasi lawan - para evaluator mewakili dua pandangan yang berlawanan (pendukung & lawan - seperti di pengadilan)
- Dua set evaluator yang berbeda menempatkan nilai pada data yang sama, dan set ini berada dalam posisi yang berlawanan.

d) Robert Stake

- Tidak ada nilai sebenarnya untuk apa pun (pengetahuan terikat konteks)

- Perspektif pemangku kepentingan merupakan elemen integral dalam evaluasi
- studi kasus adalah metode terbaik untuk mewakili keyakinan & nilai pemangku kepentingan & melaporkan hasil evaluasi
- Berbagai perspektif dari pemangku kepentingan yang berbeda perlu diwakili dalam evaluasi TETAPI evaluator memiliki peran utama dan dia membuat penilaian
- Peran evaluator - mendengar peserta yang berbeda, berunding, bernegosiasi TAPI secara non-demokratis memutuskan apa kepentingan mereka
- Peran evaluator - mencari konsensus TAPI jangan membuatnya setelah tidak ada
- Pemangku kepentingan adalah kelompok yang luas. Itu juga termasuk masyarakat umum

e) Barry MacDonald

- Evaluator sebagai negosiator perspektif. Kenali dan kumpulkan penilaian alternative, pendapat tentang program dan dari memfasilitasi negosiasi antara sudut pandang yang berbeda ini, nilai-nilai
- 3 (tiga) jenis evaluasi:
 - Birokratis - evaluator menerima nilai-nilai dari mereka yang memegang jabatan, dan memberikan informasi yang membantu mencapai tujuan
 - Otokratis - evaluator independen & bertindak sebagai penasihat ahli
 - Demokratis - evaluator menginformasikan masyarakat, berusaha untuk mewakili berbagai kepentingan, pendapat tentang program.

f) Ernest House

- Evaluator harus mewakili suara para pemangku kepentingan tersebut yang kurang terwakili, tidak didengar

- Evaluasi memberikan informasi bagi pengambil keputusan sehingga mereka dapat menentukan alokasi sumber daya yang adil dan sah
- Evaluasi bukanlah nilai-netral. Ia harus memenuhi keadilan sosial dengan mewakili suara mereka yang tidak berdaya
- Isu-isu kunci untuk menilai: benar, adil, adil
- Kesalahan etika dalam evaluasi: klientalisme, kontraktualisme, manajerialisme, metodologisme, pluralisme/elitisme, relativisme

g) Egon Guba & Yvonna Lincoln

- Pemangku kepentingan menempatkan nilai, karena tidak ada satu realitas tetapi realitas ganda berdasarkan persepsi dan interpretasi aktor yang terlibat dalam program
- Paradigma konstruktivis
- Peran evaluator - memunculkan perspektif yang berbeda dalam hubungannya satu sama lain dan informasi lainnya, mengatur proses negosiasi yang bertujuan untuk konsensus
- Prosedur evaluasi - interaksi pengumpulan dan analisis data yang berkelanjutan
- Peran tambahan dari evaluator:
 - Uji dan perbesar konstruksi dalam kelompok dengan memperkenalkan info tambahan baru
 - Memilah CC&I yang diselesaikan (Claims, Concerns & Issues)
 - Memprioritaskan CC&I yang belum terselesaikan
 - Mengumpulkan info terkait CC&I yang belum terpecahkan
 - Mempersiapkan agenda untuk negosiasi
 - Melakukan negosiasi
 - Pelaporan melalui studi kasus - konstruksi bersama sebagai produk.

Organizational Elements Model (OEM) dikembangkan oleh Roger Kaufman sebagai alat untuk mengidentifikasi elemen yang berbeda dalam suatu sistem. Sebuah sistem adalah "satu set komponen yang saling terkait yang bekerja sama untuk mencapai tujuan bersama" (Porter, 2005a). Berkenaan dengan OEM, Hannum & Kaufman (2005: 42,43) menawarkan kerangka untuk menghubungkan dan menyelaraskan upaya organisasi dengan nilai tambah kepada klien eksternal dan masyarakat.

Menganalisis hasil akhir keseluruhan organisasi dari perspektif sistemik juga membantu penulis mendekati masalah dengan cara yang efektif. OEM menyediakan kerangka sistemik, merancang dan menerapkan cara yang efektif untuk mencapai hasil akhir yang diinginkan. Kaufman membagi OEM menjadi lima unsur yang berinteraksi satu sama lain: Input, proses, dan produk, output, *outcome*. Produk dan output adalah hasil akhir yang terjadi dalam organisasi; *outcome* adalah hasil akhir yang luar organisasi (Chyung, 2005). Input: adalah bahan baku; proses adalah cara untuk melakukannya; produk adalah hasil sementara dalam proses; output adalah prestasi organisasi, sehingga produk yang dikirim kepada masyarakat; *outcome* adalah efek untuk masyarakat (Porter, 2005b). Intinya merupakan suatu kerangka

untuk berpikir, berkaitan dengan masalah yang dihadapi oleh organisasi. Mengorganisir dan menganalisis informasi yang dikumpulkan tentang organisasi dan masalah yang dihadapinya (Nickols, 2010: 1-5). Dia mengilustrasikan bahwa orang merupakan pusat dari segala sesuatu, oranglah yang merancang strategi, mengembangkan dan menggunakan teknologi, menimbulkan dan mempertahankan budaya organisasi.

OEM “menyediakan kerangka merancang dan menerapkan cara yang efektif untuk mencapai hasil akhir yang diinginkan” (Chyung, 2005). Lima unsur dalam OEM ini digunakan untuk menghubungkan sumber daya dalam organisasi untuk mengembangkan dan mencapai suatu tujuan. Model OEM juga memperhitungkan tiga jenis tingkat kebutuhan, tingkat mega (atau tingkat masyarakat), tingkat makro (atau tingkat organisasi), dan tingkat mikro (individu/ kelompok tingkat kecil). Kebutuhan ini memungkinkan organisasi mengidentifikasi kesenjangan yang berbeda untuk mencapai hasil akhir. OEM juga dapat digunakan mengidentifikasi celah dalam suatu organisasi dan menerapkan model di atas untuk menentukan kebutuhan organisasi mencapai tujuan. Chyung (2006) melakukan intervensi yang diimplementasikan berdasarkan karakteristik melalui model ARCS. (*A: Attention, R: Relevance, C: Confidence, and S: Satisfaction*). Model ARCS memberikan landasan teoritis memilih cara yang efektif melakukan penilaian pada input dan proses instruksional. Rangkaian kegiatan evaluasi berlandaskan efektivitas intervensi yakni intervensi sistematis dan sistemik untuk diterapkan pada program.

Kaufman (2006: 6-16), memberikan penegasan bahwa OEM adalah suatu tahapan bagi evaluator menentukan model evaluasi yang tepat untuk digunakan. Dalam penelitian “*Defining and Delivering Measurable Value: A Mega Thinking and Planning Primer*”, Koufman memberi gambaran bahwa *outcome* adalah suatu mega pemikiran dan perencanaan yang terfokus bukan pada organisasi tetapi pada masyarakat sekarang dan masa depan. Adanya migrasi kinerja dari individu, pertimbangan kepentingan masyarakat, dan kepentingan eksternal, dengan penekanan pada tanggung jawab, responsif, dan etika untuk menambah nilai bagi

semua. *Outcome* berada pada perencanaan mega, dengan pemikiran strategis. *Six Critical Success Faktor* (CSF) digunakan untuk setiap keberhasilan organisasi. CSF digunakan untuk proses perencanaan. CSF mencakup: 1) menilai kebutuhan (mengidentifikasi masalah berdasarkan kebutuhan; 2) menganalisis kebutuhan (menentukan persyaratan rinci dan mengidentifikasi solusi, tetapi belum pilih, alternatif solusi); 3) memilih solusi dari antara alternatif; 4) menerapkan; 5) mengevaluasi; 6) melakukan perbaikan terus-menerus (di setiap langkah).

Tabel 2. The Five Levels of Results, the Levels of Planning and a Brief Description

Name of the Organizational Element	Nama of the Level of Planning and Focus	Brief Description	Type of Planning
<i>Outcomes</i>	<i>Mega</i>	<i>Result and their consequences for external clients and society (shared vision)</i>	<i>Strategic</i>
<i>Outputs</i>	<i>Micro</i>	<i>The results an organization can or does deliver outside organization.</i>	<i>Tactical</i>
<i>Products</i>	<i>Micro</i>	<i>The building block result that are produced within the organization.</i>	<i>Operational.</i>
<i>Processes</i>	<i>Process</i>	<i>The ways, means, activities, procedures, methods used internally.</i>	
<i>Inputs</i>	<i>Input</i>	<i>The human, physical, financial resources an organization can or does use.</i>	

Sumber: Kaufman, & Thomas (1980).



MODEL-MODEL EVALUASI

1. Model CIPP.

Model ini dikembangkan oleh Komite Studi Evaluasi yang dibentuk oleh Komite Penasehat Riset (Research Advisory Committee) Phi Delta Kappa, yang diketahuai oleh Stufflebeam. Pada waktu memimpin tim ini ia bekerja di Ohio State University. Organisasi ini menugaskan kepada Stufflebeam dan kawan-kawannya, untuk berkarya mengembangkan pemikiran tentang evaluasi pendidikan (Hutahaean, 2005: 14-15).

Model CIPP untuk evaluasi adalah suatu kerangka kerja yang komprehensif untuk membimbing evaluasi program, proyek, personil, produk, institusi, dan sistem. Model ini diperkenalkan oleh Daniel Stufflebeam pada tahun 1966 untuk memandu evaluasi yang diamanatkan proyek, didanai pemerintah federal AS karena proyek-proyek ini muncul tidak bisa memenuhi persyaratan untuk dikendalikan, memanipulasi variabel-percobaan, yang kemudian dianggap sebagai standar emas untuk evaluasi program. Sejak itu, model ini telah banyak dilakukan dan dikembangkan lebih lanjut. Mereka menerapkan atau mengontrak orang lain untuk menerapkan model ini termasuk pejabat pemerintah, pejabat yayasan, staf program dan proyek, personil bantuan internasional, administrator sekolah, dokter, pemimpin militer, dan evaluator.

Model ini dikonfigurasi untuk digunakan dalam evaluasi internal yang dilakukan oleh evaluator organisasi, dalam evaluasi diri yang dilakukan oleh tim proyek atau penyedia layanan individu, dan dalam evaluasi eksternal kontrak. Telah digunakan di seluruh...

Model CIPP terbentuk dari empat dimensi evaluasi yaitu evaluasi *Context*, *Input*, *Process*, dan *Product*. Keempat evaluasi ini merupakan suatu rangkaian keutuhan, walaupun dalam pelaksanaannya seseorang dapat saja hanya melakukan satu jenis atau kombinasi dari dua atau lebih dimensi evaluasi, namun sesungguhnya kekuatan model terletak dari rangkaian kegiatan keempat jenis evaluasi itu sendiri (Daryanto. 2005). Dalam pandangan Stufflebeam, bahwa tujuan yang paling penting adalah bukan membuktikan, tetapi memperbaiki. "*the CIPP approach is based on the view that the most important purpose of evaluation is not prove but to improve*" (Madaus, criven, Stufflebeam, 1993: 118).

Model evaluasi CIPP (*Context*, *Input*, *Process*, *Product*) yang dikembangkan oleh Stufflebeam, dan mengidentifikasi empat tipe keputusan yang diambil, masing-masing tipe keputusan berhubungan dengan tipe atau jenis evaluasi, yaitu:

- a. *Context evaluation as a means of servicing planning decisions (to determine objectives); 2)input evaluations these structuring decisions (to design instruction); 3)process evaluations to guide implementing (to use or improve instruction); and 4)product evaluations to serve recycling decisions (to judge products, outputs, or outcomes).* (Stufflebeam & Shinkfield, 1985: 169),
Orientasi pada evaluasi konteks adalah untuk mengidentifikasi kekuatan dan kelemahan obyek, seperti institusi program, populasi, atau perorangan dan untuk menetapkan arah perbaikan. Tujuan dari ini adalah untuk menilai keseluruhan obyek, mengidentifikasi kekurangan, mengidentifikasi kekuatan atau kelebihan yang nantinya dapat digunakan untuk memperbaiki kekurangan-kekurangan.
- b. Evaluasi input untuk membantu mengambil keputusan, menentukan sumber-sumber yang ada, rencana dan strategi untuk mencapai tujuan, serta bagaimana prosedur kerja untuk mencapainya. Komponen-komponen input mencakup sumber

daya manusia, materi program dan rancangan aplikasinya, sarana dan peralatan pendukung, dana dan berbagai prosedur serta aturan yang diperlukan.

Evaluasi inputs dilaksanakan untuk menilai alternative pendekatan, rencana tindak, rencana staf, dan pembiayaan bagi kelangsungan program dalam memenuhi kebutuhan kelompok sasaran serta mencapai tujuan yang ditetapkan. Evaluasi ini berguna bagi pembuat kebijakan untuk memilih rancangan, bentuk pembiayaan, alokasi sumber daya, pelaksana dan jadwal kegiatan yang apling sesuai bagi kelangsungan program

- c. Evaluasi proses untuk membantu mengimplementasikan keputusan, sampai sejauhmana rencana telah diterapkan dan apa yang harus direvisi. Komponen-komponen proses meliputi proses pembelajaran dan pelaksanaan program, proses pengelolaan program, dan hambatan manakah yang dijumpai selama pelaksanaan program berlangsung dan perlu diatasi.

Evaluasi process ditujukan untuk menilai implementasi dari erencana yang telah ditetapkan guna membantu para pelaksana dalam menjalankan kegiatan dan kemudian akan dapat membantu kelompok pengguna lainnya untuk mengetahui kinerja program dan memperkirakan hasilnya.

- d. Evaluasi produk digunakan untuk membantu membuat keputusan selanjutnya, baik mengenai hasil yang telah dicapai maupun apa yang dilakukan setelah program itu berjalan, dengan kata lain evaluasi ini merupakan penilaian yang dilakukan untuk mengukur keberhasilan dalam pencapaian tujuan. Adapun kelebihan dan keterbatasan model ini (Aman, 2009) adalah lebih komprehensif dan keterbatasannya adalah bila diterapkan pada program pembelajaran, mempunyai tingkat keterlaksanaan yang kurang tinggi.

Evaluasi product dilakukan dengan tujuan untuk mengidentifikasi dan menilai hasil yang dicapai, yang diharapkan dan yang tidak diharapkan, jangka pendek dan jangka panjang baik bagi pelaksana kegiatan agar dapat menfokuskan diri dalam mencapai sasaran program maupun

bagi pengguna lainnya dalam menghimpun upaya untuk memenuhi kebutuhan kelompok sasaran. **Stufflebeam, 2003)**

2. *Measurement Evaluation Model.*

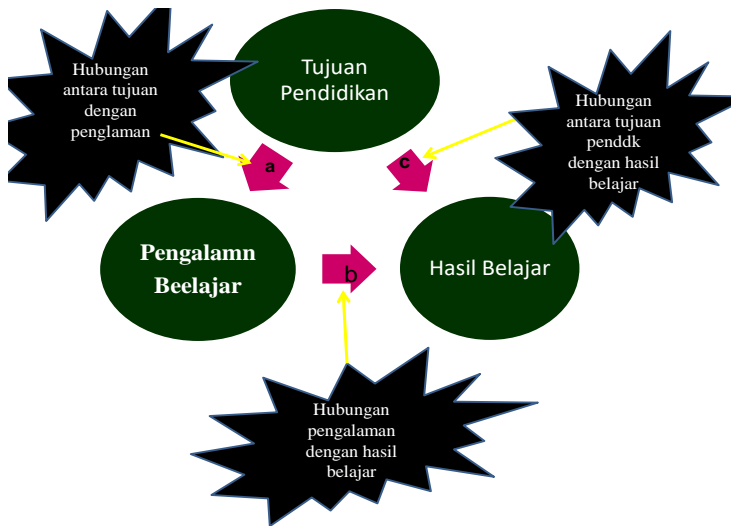
Model ini menitik beratkan peranan kegiatan pengukuran dalam melaksanakan proses evaluasi. Adanya pengaburan batas antara pengukuran dan evaluasi. Menurut R. Thorndike: *“If anything exists, it exists in quantity, and if it exists in quantity it can be measured”*. Pengukuran dipandang sebagai kegiatan menentukan besarnya suatu (*atribute*) tertentu yang dimiliki oleh obek, orang maupun peristiwa, dalam bentuk unit ukuran tertentu. Pengukuran digunakan untuk berbagai aspek tingkah laku dengan tujuan untuk melihat perbedaan-perbedaan individual atau kelompok, yang hasilnya diperlukan dalam rangka seleksi, bimbingan, dan perencanaan pendidikan bagi para siswa di sekolah.

Ruang lingkup objek kegiatan evaluasi adalah tingkah laku, terutama tingkah laku siswa. Aspek tingkah laku yang dinilai meliputi: kemampuan hasil belajar, kemampuan pembawaan (intelgensi, bakat), minat, sikap, dan aspek-aspek kepribadian siswa. Menitikberatkan pada pengukuran terhadap hasil belajar yang dicapai siswa pada masing-masing bidang pelajaran dengan menggunakan tes. Hasil belajar disini adalah bidang pengetahuan (kognitif) yang mencakup berbagai tingkat kemampuan, yang evaluasinya menggunakan prosedur yang distandarisasikan.

3. *Congruence Model.*

Pendidikan adalah sebagai proses yakni: tujuan pendidikan, pengalaman belajar, penilaian hasil belajar. Menurut model dengan tokohnya Raph W, Tyler, John B, Carrol, Lee J. Cronbach ini, evaluasi adalah sebagai kegiatan untuk melihat sejauh mana tujuan-tujuan pendidikan telah diapai siswa dalam bentuk hasil belajar. Dengan kata lain, evaluasi pada dasarnya ingin memperoleh gambaran mengenai efektivitas dari sistem pendidikan dalam mencapai tujuannya. Pada akhirnya dapat dicapai baik individu maupun kelompok, dapat diambil keputusan tentang tindakan-tindakan apa

yang perlu diambil sehubungan dengan sistem pendidikan dan siswa yang bersangkutan. Bagi kepentingan sistem pendidikan, hasil evaluasi digunakan sebagai umpan balik untuk kebutuhan memperbaiki bagian-bagian sistem yang masih “lemah” Memberikan info bagi pihak luar sejauh mana tujuan-tujuan yang telah dicapai oleh sistem pendidikan tsb. Adalah usaha untuk memeriksa kesesuaian (*congruence*) antara tujuan-tujuan pendidikan yang diinginkan dan hasil belajar yang telah dicapai. Termasuk didalamnya perubahan tingkah laku anak didik, hasilnya digunakan untuk penyempurnaan sistem bimbingan siswa, memberikan informasi kepada pihak luar.



Gambar 6. Measurement Evaluation Model

Langkah pokok model *Congruence*.

- 1) Merumuskan dan mempertegas tujuan-tujuan pengajaran. Masing-masing tujuan diperjelas rumusannya, sehingga memberikan arah yang tegas dalam perencanaan evaluasi yang akan dilakukan.
- 2) Menetapkan “*test situation*” yang diperlukan. Menetapkan jenis-jenis situasi yang akan memungkinkan para siswa untuk memperhatikan tingkah laku yang akan dievaluasi.

- 3) Menyusun alat evaluasi Menyusun alat evaluasi yang cocok untuk digunakan dalam menilai jenis tingkah laku yang tergambar dalam tujuan
- 4) Menggunakan hasil evaluasi Mengolah hasil sedemikian rupa agar dapat memenuhi tujuan diadakanya evaluasi.
- 5) Melakukan evaluasi, mana yang perlu disempurnakan, bagian mana yang belum dicapai oleh siswa, untuk dilakukan pembimbimbangan yang lebh terarah.

4. *Educational System Evaluation Model*

Tokoh model evaluasi ini adalah Daniel L, Stufflebeam, Michael Scriven, Robert E. Stake, Malcolm M Provus. Model evaluasi ini bertitik tolak bahwa suatu sistem pendidikan dipengaruhi oleh berbagai faktor, karakteristik anak didik maupun lingk sekitarnya, tujuan sistem dan peralatan yang dipakai, serta prosedur dan mekanisme pelaksanaan sistem. Membandingkan *performance* dari berbagai dimensi sistem yang sedang dikembangkan dengan sejumlah kriteria tertentu. Malcolm M. Provus dalam pembahasannya mengenai The Discrepancy Evaluation Model mengemukakan bahwa *The Discrepancy Evaluation Models (Malcolm M Provus). There can be non evaluation without discrepancy information; the can be non discrepancy without standards or criteria.* Perlu ditetapkan dengan tegas kriteria yang akan dijadikan ukuran untuk mengevaluasi *performeance* dari masing-masing dimensi. Kegiatan evaluasi harus sampai pada *judgment* baik buruknya, efektif tidaknya sistem pendidikan. Robert E Stake dalam bukunya *The Countenance of Educational Evaluation*, bahwa *from relative judgment as well as from absolute judgment, we obtain an overal or composite rating of merit, a rating to be used an making an educatonal dicision.* Dalam mangadakan *judgment* dapat menggunakan standar mutlak/relatif dalam bentuk perbandingan dengan sistem pendidikan yang lain.

Beberapa hal penting terkait dengan model ini adalah bahwa berbagai dimensi sistem sebagai suatu keseluruhan objek evaluasi. Gene V Class dalam tulisannya yang berjudul: *Two generation of Evaluation Models*, menegaskan bahwa "*The complete and detailed*

description of what constitutes of educational program is a concern of the educational system evaluation model".

5. Alkin Model (The UCLA Evaluation Model)

Marvin Alkin merupakan orang yang fokus pada evaluasi dan pengambilan keputusan. Model awal Alkin (1972b) memiliki banyak kesamaan dengan model CIPP Stufflebeam. Perbedaan utama antara karya Alkin dan Stufflebeam adalah pengakuan Alkin bahwa proses dan produk memiliki dimensi sumatif dan formatif.

Marvin Alkin (1991) menempatkan pandangannya dimana evaluasi sebagai "berorientasi pengguna" dalam rangka untuk meningkatkan kemanfaatannya. Alkin lebih focus pada pengguna sebagai kelanjutan evaluasi yang berfokus pada pemanfaatan. Alkin sangat menolak keterlibatan evaluator sebagai penilai suatu agen. Alkin lebih suka bekerja dengan pengguna utama di awal proses evaluasi untuk menetapkan sistem nilai untuk menilai data hasil potensial.

Model evaluasi menurut Alkin adalah proses pelayanan terhadap pengambilan keputusan dengan memilih informasi yang sesuai, mengumpulkan dan analisis informasi untuk mendapat data ringkas yang berguna dalam rangka pengambil keputusan guna memilih alternatif terbaik. Ada lima jenis kegiatan: 1) penilaian sistem, 2) perencanaan program, 3) implementasi program, 4) perbaikan/peningkatan program dan 5) sertifikasi program. Penjelasannya sebagai berikut:

1. *System assessment* (penilaian sistem) yaitu: menyuguhkan informasi mengenai keadaan atau posisi sistem.
2. *Program Planning* (perencanaan program) yaitu: membantu pemilihan program tertentu yang mungkin akan berhasil bagi pemenuhan kebutuhan program.
3. *Program implementation* (implementasi program) yaitu: menyiapkan informasi apakah program sudah disosialisasikan kepada kelompok tertentu sesuai yang direncanakan.
4. *Program Improvement* (perbaikan/peningkatan program) yaitu: menyuguhkan informasi bagaimana program

berfungsi, bagaimana program berlangsung, apakah tujuan tercapai dan kendala-kendala yang muncul.

5. *Program Certification* (sertifikasi program) yaitu: memberikan informasi tentang arti, nilai dan manfaat suatu program.

6. *Illuminative Model* (Malcolm Parlett).

Model ini dikaitkan dengan pendekatan dalam **bidang antropologi**. Menekankan pada evaluasi kualitatif dan “terbuka”. Sistem pendidikan yang dinilai tidak ditinjau sebagai sesuatu yang terpisah melainkan sebagai bagian keseluruhan sistem pendidikan di sekolah dalam hubungan dengan suatu *learning milieu* (lingkungan material, psikososial yang ikut mempengaruhi pengalaman belajar siswa). Hasil evaluasi lebih bersifat deskripsi dan interpretasi, bukan pengukuran dan prediksi. Bersemboyankan “*the judgment is the evaluation*”.

7. Model Evaluasi CSE (*Center for Study of Evaluation*)

Evaluasi ini terdiri atas empat tahap: “1) *needs assessment*, 2) *program planning*, 3) *formatif evaluation*, 4) *summative evaluation*”. (Fernandes, 1984: 11).



Gambar 7. Model Evaluasi CSE

Pada *needs assessment*, memberikan informasi tentang keadaan atau posisi program, berkenaan dengan tujuan atau kebutuhan yang dikehendaki, *program planning* membantu memilih program tertentu yang mungkin akan berhasil untuk memenuhi kebutuhan program. Pada *formative evaluation*, menyiapkan informasi, berbagi informasi pada kelompok tertentu tentang bagaimana program bekerja atau berjalan, atau hal-hal baru yang muncul tidak terduga untuk perbaikan program, pada

summative evaluation memberikan informasi tentang nilai atau manfaat program yang telah dilaksanakan.

8. Model evaluasi Kirkpatrick.

Model ini banyak digunakan untuk mengevaluasi program pelatihan, dikembangkan oleh Kirkpatrick dengan empat level yaitu: “*reaction, learning, behavior, results*” (Kirkpatrick, 2006: 21).

- 1) *Reaction*, mengevaluasi terhadap reaksi peserta pelatihan dengan mengukur kepuasan peserta. Program dirasakan lebih efektif bila proses pelatihan dirasa menyenangkan dan memuaskan bagi peserta pelatihan, sehingga peserta termotivasi untuk belajar dan berlatih. Sebaliknya apabila peserta merasa tidak puas maka peserta tidak akan termotivasi untuk mengikuti pelatihan lebih lanjut.
- 2) *Learning*, didefinisikan sebagai perubahan sikap, perbaikan pengetahuan, dan peningkatan ketrampilan peserta setelah selesai mengikuti program.
- 3) *Behavior*, perubahan sikap peserta pelatihan setelah peserta kembali ketempat kerja masing-masing. Bagaimana peserta dapat mentransfer pengetahuan, sikap dan ketrampilan yang diperoleh selama pelatihan untuk diterapkan di tempat kerja.
- 4) *Results*, difokuskan pada hasil akhir yang meliputi peningkatan produktivitas, peningkatan kualitas, penurunan biaya, penurunan kuantitas kecelakaan kerja, dan kenaikan keuntungan.

Adapun kelebihan dan keterbatasan model ini (Aman, 2009) adalah lebih komprehensif karena mencakup aspek kognitif, skill dan afektif, objek evaluasi tidak hanya hasil belajar semata tetapi mencakup proses, *output* maupun *outcomes* lebih mudah diterapkan, keterbatasannya adalah kurang memperhatikan *input* dan sulit untuk mengukur *impact*.

Model Kirkpatrick dalam konteks reformasi kebijakan, merupakan titik awal yang baik. Pendekatan yang dikembangkan oleh Kirkpatrick ini, dikenal sebagai Model Kirkpatrick, telah

digunakan dalam beberapa cara yang berbeda oleh berbagai organisasi, baik dalam adaptasi atau bentuk asli. Model ini menitikberatkan pengumpulan data pada empat tingkat yang berbeda, seperti yang ditunjukkan pada table berikut ini.

Tabel 3. Model Evaluasi Kirkpatrick's

Evaluasi tingkat lanjut	Results level	Dampak pengembangan kapasitas pada organisasi kepolisian dan tujuannya
	Performance level	Dampak pengembangan kapasitas pada kinerja/perilaku di tempat kerja
Tingkat dasar evaluasi	Training level	Dampak pengembangan kapasitas dalam hal apa yang dipelajari oleh petugas peserta pelatihan
	Reaction level	Dampak peningkatan kapasitas dalam hal kepuasan petugas peserta pelatihan

Penjelasan secara rinci tabel diatas adalah:

- 1) Tingkat yang lebih rendah (reaksi dan pelatihan) memiliki arti penting bagi mereka yang terlibat dalam desain dan pengembangan materi pelatihan dan pengembangan, serta pengembangan instruktur. Evaluasi tingkat pertama (tingkat reaksi/ Reaction level) dapat memberikan data yang sangat berharga tentang masalah yang muncul selama program pengembangan kapasitas itu sendiri dan, terkadang, wawasan tentang penyebab jika program tersebut kurang efektif.
- 2) Sedangkan tingkat pelatihan (Training level) berusaha mengukur apakah pengetahuan dan keterampilan yang dibutuhkan yang terkandung dalam tujuan program benar-benar telah dipelajari. Hal ini biasanya dicapai melalui tes atau penilaian formal yang menggunakan pengukuran yang objektif dan terukur.

Tingkat reaksi dan pembelajaran relatif mudah diatur tetapi tidak memberikan indikator yang signifikan dari ujian akhir program pengembangan kapasitas, dilihat dalam konteks

reformasi kebijakan: yaitu, dampak nyata pada kinerja tempat kerja. Dengan demikian, Kirkpatrick memasukkan dua level lebih lanjut.

- 3) Level kinerja/performance mencoba untuk mengukur kinerja pekerjaan kebijakan melalui berbagai alat evaluasi selama periode waktu tertentu. Berhubungan erat dengan tingkat hasil (result evaluation) yang berupaya mengukur pengaruh program pengembangan kapasitas terhadap kinerja keseluruhan organisasi kepolisian. Seperti yang Anda bayangkan, ini tidak mudah dicapai karena sulitnya mengendalikan banyak variabel yang mempengaruhi hasil kinerja yang diketahui. Namun, terlalu sering evaluasi penggunaan dalam organisasi kepolisian tetap terkunci secara impoten pada tingkat reaksi langsung (yaitu melalui penggunaan yang disebut 'lembar bahagia'). Dalam proses reformasi kepolisian, penekanannya harus pada penggunaan evaluasi untuk memverifikasi dampak pada tujuan organisasi (tingkat hasil) atau setidaknya untuk mengidentifikasi beberapa perubahan signifikan dalam kinerja/sikap staf yang relevan (tingkat prestasi kerja).

Kekuatan model Kirkpatrick, terletak pada potensinya sebagai alat diagnostik dalam memantau kemajuan dalam tujuan reformasi secara keseluruhan. Misalnya, di mana sebuah organisasi telah merancang dan mengimplementasikan sebuah program untuk mengatasi kegagalan petugas dalam menghormati hak-hak orang yang ditahan di bawah hukum setempat. Berdasarkan dua level yang terakhir ada kemungkinan dapat dipastikan melalui wawancara dengan orang-orang yang ditangkap. Hasil deteksi pada tingkat kinerja pekerjaan (yaitu persentase dari orang yang ditangkap melaporkan bahwa mereka tidak diberitahu tentang hak-hak mereka oleh petugas terlatih). Hal ini mengindikasikan bahwa kinerja beberapa petugas telah gagal sesuai yang diinginkan. Hasil eksplorasi di level bawah adalah langkah awal untuk mengisolasi penyebab kegagalan dalam memenuhi kebutuhan organisasi melalui peningkatan kapasitas. Analisis hasil

evaluasi dilakukan secara lebih rinci sampai dihasilkannya informasi peningkatan kapasitas.

9. CIRO Model

Pendekatan evaluasi lain yang cocok untuk adaptasi adalah yang dijelaskan dalam karya P. Warr, M. Bird dan N. Rackham. Pendekatan ini dikenal dengan nama 'CIRO'. Model ini didasarkan pada empat kategori pengukuran tetapi berbeda dari model Kirkpatrick dalam beberapa hal. Ini membayangkan empat kategori pengambilan data: (1) Context evaluation; (2) Input evaluation; (3) Reaction evaluation; (4) Outcome evaluation.

- (1) Context evaluation (evaluasi konteks) berusaha mengukur konteks di mana program pengembangan kapasitas berlangsung. Ini meneliti cara kebutuhan kinerja diidentifikasi, tujuan pembelajaran ditetapkan, dan cara tujuan menghubungkan dan mendukung kompetensi yang diperlukan. Selain itu, harus mempertimbangkan bagaimana komponen program ini mencerminkan budaya dan struktur organisasi. Jenis evaluasi ini menegaskan jika pengembangan kapasitas diperlukan.
- (2) Input evaluation (evaluasi masukan) mencoba mengukur sejumlah masukan untuk program pembangunan kapasitas, dengan maksud untuk membantu para manajer dalam proses mengidentifikasi masukan mana yang paling hemat biaya. Untuk itu, ini berfokus pada sumber daya yang dibutuhkan untuk memenuhi kebutuhan kinerja (misalnya staf, fasilitas, peralatan, katering, anggaran), konten dan metode penyampaian yang memungkinkan tercapainya peningkatan kapasitas, petugas yang berpartisipasi, dan hasil dari program sebelumnya yang serupa.
- (3) Reaction evaluation (evaluasi reaksi) mencoba mengukur bagaimana petugas peserta pelatihan bereaksi terhadap program tersebut. Melawan apa yang dimaksudkan oleh program, jenis evaluasi ini mengacu pada pendapat subjektif peserta tentang pembangunan kapasitas dan bagaimana hal itu dapat ditingkatkan.

(4) Outcome evaluation (evaluasi hasil) harus mengukur hasil pelatihan dan pengembangan terhadap tolok ukur tujuan program. Penulis membedakan empat tingkat evaluasi hasil yang memiliki kesejajaran yang kuat dengan model Kirkpatrick yaitu: hasil pembelajaran peserta pelatihan (yaitu perubahan dalam pengetahuan dan keterampilan mereka), hasil di tempat kerja (yaitu perubahan dalam kinerja pekerjaan yang sebenarnya), hasil untuk area organisasi yang relevan (yaitu departemen atau unit spesialis), hasil bagi organisasi secara keseluruhan. Seperti dalam model Kirkpatrick, ukuran hasil terakhir inilah yang mewakili tantangan terbesar karena tuntutan untuk membuktikan bahwa pengembangan kapasitas, dibandingkan dengan faktor-faktor lain, menghasilkan perubahan nyata dalam organisasi. Tentu saja banyak tergantung pada sifat tujuan pembelajaran. Mereka yang mengakibatkan hasil yang nyata, dapat diamati dan terukur, seperti pengurangan biaya operasi (misalnya pengurangan biaya bahan bakar untuk kendaraan polisi), peningkatan layanan kepolisian (misalnya panduan pencegahan kejahatan) dan peningkatan efisiensi kerja (misalnya metode patroli terstruktur) akan siap membantu pendekatan evaluasi ini.



BAB 8

STAKE MODELS

Stake membagi objek evaluasi mejadi 3 kategori, yaitu: *Antecedents*, *Transaction*, *Outcomes*. *Antecedents* yaitu sumber/model/input yang ada pada sistem itu untuk dikembangkan, seperti tenaga, keuangan, karakteristik siswa, dan tujuan yang ingin dicapai, kondisi yang ada sebelum instruksi yang mungkin berhubungan dengan hasil. *Transaction*, mencakup rencana kegiatan maupun proses pelaksanaannya, termasuk urutan kegiatannya, penjadwalan, bentuk interaksi guru-murid, cara menilai hasil belajar, dsb. Keterlibatan secara berurutan atau pertemuan yang dinamis merupakan proses instruksi. *Outcomes* merupakan hasil yang dicapai oleh para siswa, reaksi guru terhadap sistem tersebut, dan efek samping dari sistem yang bersangkutan, serta dampak dari pengalaman instruksional (Popham, 1993: 5-15; Stake, 1977: 372-390; Fernandes, 1984: 8-10). Selain digunakan untuk evaluasi pembelajaran, juga bisa untuk pelatihan.

1. Pendekatan *Stake model*

Pendekatan *Stake model* (*Countenance Evaluation Model*) adalah: *contingencies* dan *congruence*. *Contingencies* adalah hubungan logis antara 3 dimensi *antecedent, transaction, outcomes*, khususnya antara tujuan, strategi, dan hasil baik dalam rencana maupun di lapangan. *Congruence* adalah kesesuaian antara yang diharapkan (kriteria) dan yang terjadi/dihasilkan baik rencana, proses dan hasil. Menurut Stake, evaluasi apapun adalah mendeskripsikan secara lengkap untuk penilaian program. Stake menciptakan kerangka evaluasi untuk membantu penilai dalam mengumpulkan, mengorganisir, dan menafsirkan data kuantitatif dan kualitatif untuk mencapai tujuan. Skemanya menitikberatkan suatu perhatian pada perbedaan antara tindakan deskriptif dan pertimbangan yang sesuai dengan fase program pendidikan: *antecedent, transaction, dan outcome* (Popham, 1993: 5-15; Stake, 2000: 350-351).

Stake lebih lanjut membagi untuk mendeskriptifkan apa yang akan terjadi atau kejadian sebenarnya yang diamati. Tindakan pertimbangan dibagi berdasarkan apakah mereka mengacu pada standar yang digunakan dalam pembuatan *judgments* atau penilaian yang sebenarnya (Popham, 1993: 5-15). Evaluator mencatat apa yang dimaksudkan oleh pengembang, apa yang dirasakan pengamat, apa yang pelanggan harapkan, dan nilai apakah yang ditempatkan pada setiap aspek (Stake, 1977: 372-390). Gambar 3 menggambarkan tata letak laporan dan data yang akan dikumpulkan oleh evaluator.

<u>Description Matrix</u>		<u>Judgment Matrix</u>	
<i>Intens</i>	<i>Observations</i>	<i>Standards</i>	<i>Judgement</i>
R a t i o n a l	C o n t i n g e n c i e s	C o n t i n g e n c i e s	A N T E C E D E N T
	C o n t i n g e n c i e s	C o n t i n g e n c i e s	T R A N S A C T I O N
	C o n t i n g e n c i e s	C o n t i n g e n c i e s	O U T C O M E S

Sumber Stake's, 2000:351.

Gambar 8. Implementasi Stake Model

Stake's model berguna karena memberikan bukti yang luas untuk mendukung keberhasilan suatu kegiatan dengan mendokumentasikan hubungan antara semua komponen program kegiatan. Evaluator memulai dengan melihat maksud untuk kondisi dan perilaku, termasuk apa yang diinginkan, diantisipasi, dan efeknya. Evaluator berikutnya akan menganalisis informasi dalam matriks deskriptif dengan melihat kesesuaian antara maksud dan pengamatan (Stake, 1977: 372-390). Evaluator akan menerapkan standar data deskriptif untuk membentuk penilaian. Stake menyatakan bahwa dalam setiap tahap evaluasi ada data deskriptif yang mencocokkan antara *intent* dengan *observation* sedangkan penilaian (*judgment*) membandingkan secara absolut antara data deskriptif dari setiap tahap dengan standar (Stake, 2006: 6).

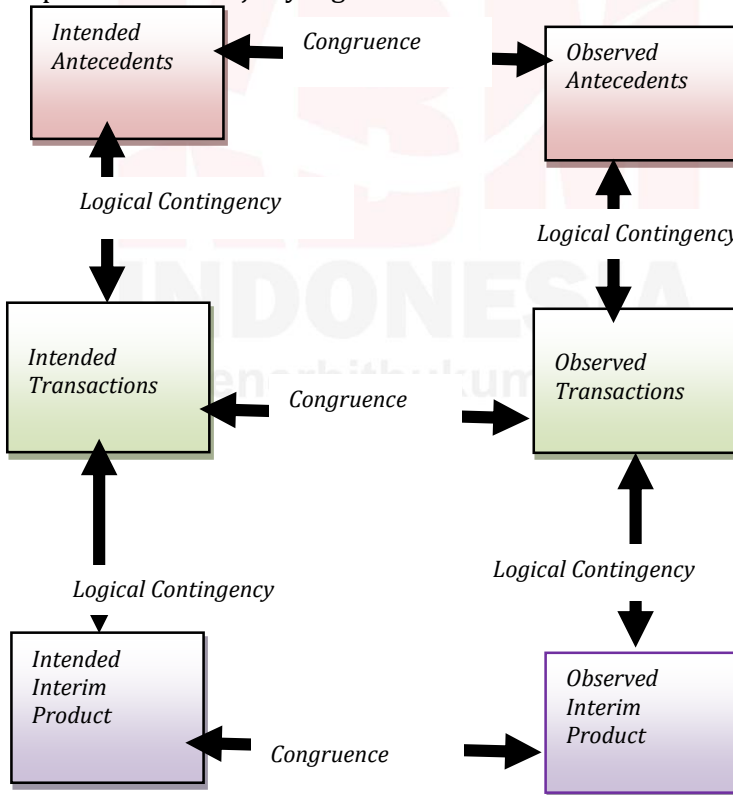
Evaluators juga mempelajari hubungan, antara *anteseden*, *transaction*, dan *outcomes*. *Anteseden* adalah kondisi sebelum program dijalankan; *transaction*, adalah program yang dinamis yang terjadi selama operasi program; *outcomes* adalah efek program, baik yang disengaja maupun yang tidak disengaja. *Outcomes* tergantung pada transaksi dan pada kondisi anteseden. Evaluasi formatif lebih cenderung dalam kontingensi, evaluasi sumatif lebih cenderung pada kongruensi (Stake, 1977: 372-390). Data matriks Stake melibatkan evaluator terus menerus dalam deskripsi dan pertimbangan di awal, selama, dan pada akhir dari program instruksional (Popham, 1993: 5-15). Matriks membagi suatu interaksi unit yang kompleks ini dapat dikelola dan diidentifikasi titik-titik mana dari instrumen yang harus digunakan untuk mengumpulkan data.

2. Alur Pendekatan Stake Models

Wood, (2001: 18-27) dalam penelitiannya *Stake's Countenance Model: Evaluating an environmental Education Professional Development Course*, adalah sebuah kerangka kerja evaluasi yang berurutan mengatur pertanyaan penelitian menurut dampaknya terhadap fitur yang dimaksudkan. Sepuluh fitur dianalisis dengan menggunakan delapan instrumen pengumpulan data kuantitatif dan kualitatif. Model ini dikelola dan dianalisis kesesuaiannya antara hal yang diharapkan dengan hal yang sebenarnya terjadi. Proses analisis dilakukan dengan melihat *congruence* (kesesuaian) antara maksud dan pengamatan. Diamati juga interaksi antara variabel evaluasi kontingensi yang disebut *contingencies* (hubungan logis) yakni hubungan antara *antesedent*, *transaction*, dan variabel *outcomes*. *Outcome* tergantung pada *transaction* dan pada kondisi *antesedent*. Kemungkinan analisis secara simultan dari beberapa aspek, evaluasi menemukan interaksi yang signifikan untuk perbaikan. Efektivitas dari model stake untuk evaluasi formatif karena difasilitasi pemeriksaan menyeluruh dari data kualitatif dan kuantitatif dalam semua tahap pelaksanaan program. Gambar alur evaluasi tertera pada gambar di bawah.

Poin-poin pokok Stake (Stufflebeam (1984: 212) meliputi:

- 1) Evaluasi akan membuka audiens dengan melihat, mencermati dan mendeskripsikan apa yang mereka kerjakan?
- 2) Evaluasi akan mendeskripsikan program untuk dihubungkan dengan *antecedent*, *traksaction* dan *outcomes*.
- 3) Efek samping dan keuntungan insidental dan juga maksud mempelajari *autcome*.
- 4) Evaluasi akan menjauhi kesimpulan sumatif yang final tetapi bermanfaat mengumpulkan, menganalisis, merefleksikan, dan menilai bermacam-macam berkenaan dengan orang yang menjadi objek untuk dievaluasi.
- 5) Eksperimen dan standar tes sering tidak tepat atau tidak cukup untuk menemukan tujuan evaluasi. Termasuk juga dalam menggantikan tambahan variasi metode, termasuk pendekatan subjek yang lemah.



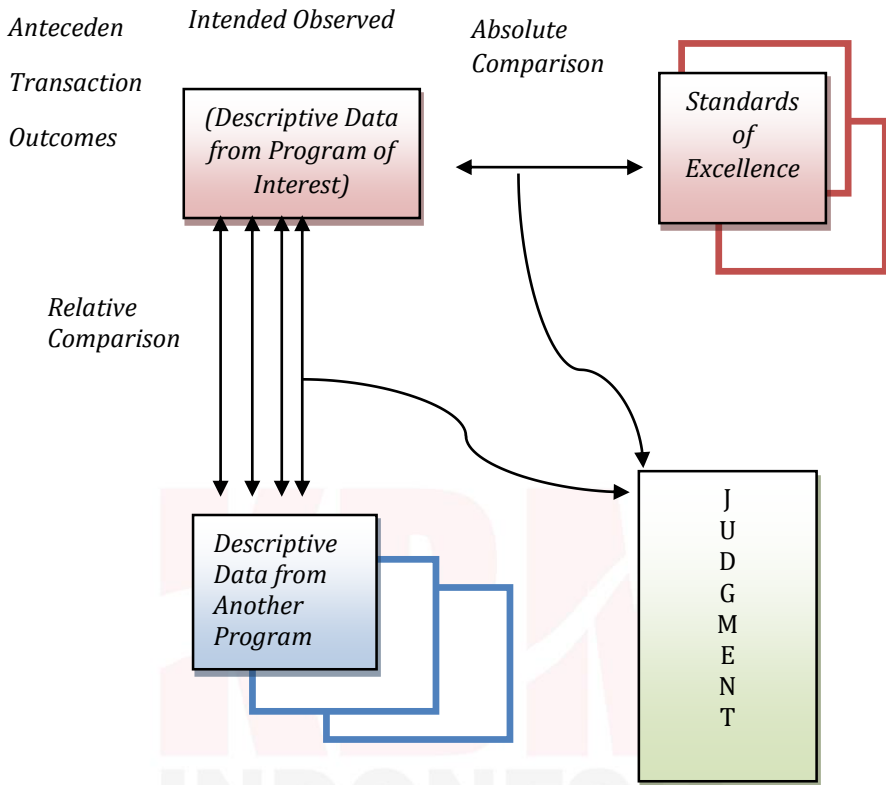
Sumber: Stake, 1977: 372-390

Gambar 9. Proses Deskripsi Data Model Stake

Representasi operasi yang terlibat dalam mencapai penilaian evaluasi muncul (gambar 4). Stake belum melihat setiap proses evaluasi sebagai bentuk yang tertib, berikut ini lebih melekat dalam proses yang digambarkan dalam gambar (Stufflebean, 1984: 223-225):

- 1) Evaluator mengumpulkan dan menganalisis informasi deskriptif (dan menjelaskan alasan program).
- 2) Evaluator mengidentifikasi standar absolut (keyakinan-keyakinan formal dan informal oleh kelompok-kelompok acuan yang relevan dari standar keunggulan).
- 3) Evaluator mengumpulkan data deskriptif dari program lain dan berasal dari standar relatif terhadap program yang menarik.
- 4) Evaluator menilai seberapa jauh program kepentingan memenuhi standar mutlak dan relatif.
- 5) Baik tunggal atau dalam kolaborasi di atas orang lain, para hakim evaluator program, yaitu, memutuskan standar yang diindahkan. Secara spesifik, ia memberikan hal yang penting, untuk setiap set standar.

INDONESIA
www.penerbitbukumurah.com



Sumber. Stake, 1977: 372-390.

Gambar 10. Stake's Representation of the Process of Judging the Merit of an Educational Program



THE DESCREPANCY EVALUATION MODEL (DEM)

1. Ruang Lingkup DEM

The Discrepancy Evaluation Models (DEM), yang dirancang oleh Malcolm Provus pada tahun 1969 adalah cara yang efektif untuk mengevaluasi program-program akademik. Provus mendefinisikan sebagai proses evaluasi untuk menyepakati standar program, dan menggunakan perbedaan antara hasil dan standar untuk mengevaluasi program. Intinya adalah mengatur aspek dari program, dan menggunakan informasi perbedaan untuk mengidentifikasi kelemahan program. Dengan demikian, DEM dapat digunakan sebagai penilaian formatif untuk menentukan apakah akan merevisi, meningkatkan, mempertahankan atau mengakhiri program. Model ini ideal untuk menemukan masalah dengan perbedaan. Hal ini juga dapat digunakan untuk merancang program dari awal sampai akhir, yang disebut evaluasi sumatif. Modelnya terutama set pemecahan masalah, prosedur berusaha untuk mengidentifikasi kelemahan (menurut standar yang dipilih) dan melakukan perbaikan dengan pemutusan berkenaan dengan opsi tersebut yang terakhir (Provus, 1969: 5-8).

DEM disebut dengan evaluasi kesenjangan program. Kesenjangan program adalah sebagai suatu keadaan antara yang diharapkan dalam rencana dengan yang dihasilkan dalam pelaksanaan program. Evaluasi kesenjangan dimaksudkan untuk mengetahui tingkat kesesuaian antara standard yang sudah ditentukan dalam program dengan penampilan aktual dari program tersebut. Standar adalah: kriteria yang telah dikembangkan dan ditetapkan dengan hasil yang efektif. Penampilan adalah: sumber, prosedur, manajemen dan hasil nyata yang tampak ketika program dilaksanakan.

Provus mendefinisikan evaluasi sebagai alat untuk membuat pertimbangan (*judgement*) atas kekurangan dan kelebihan suatu objek berdasarkan standar dan kinerja. Model ini juga dianggap menggunakan pendekatan formatif dan berorientasi pada analisis sistem. Standar dapat diukur dengan menjawab pertanyaan bagaimana program berjalan. Sementara pencapaiannya adalah apakah yang sebenarnya terjadi. Evaluator hanya boleh membantu dengan membentuk dan menjelaskan peranan standar dan pencapaian (Provus, 1969: 2-5; Steinmetz, 2000: 127-135).

Menurut Provus evaluasi adalah untuk membangun dan afirmatif, tidak untuk menghakimi. *The Discrepancy Evaluation Models (DEM)* (Provus, 1971: 15-18) adalah suatu model evaluasi program yang menekankan pentingnya pemahaman sistem sebelum evaluasi. Kapan saja kita sedang mencoba untuk mengevaluasi sesuatu, ditekankan bahwa harus mempunyai pemahaman yang tepat dan jelas atas hal yang dievaluasi, untuk menetapkan standar.

DEM menawarkan pendekatan, pragmatis sistematis untuk berbagai kebutuhan evaluasi. Dari kegiatan sehari-hari individu guru dapat mengevaluasi program pendidikan. DEM dapat dimanfaatkan untuk struktur pengumpulan informasi penting, untuk informasi pengambilan keputusan. Sebuah fitur utama dari DEM adalah penekanan pada **evaluasi diri** dan perbaikan program secara sistematis.

2. Tahapan DEM

Evaluasi pada tingkat yang paling sederhana dapat dilihat sebagai perbandingan kinerja terhadap standar. Ketika evaluasi dipandang sebagai suatu proses untuk pengembangan program, stabilisasi, dan penilaian, seperti yang terjadi dalam DEM, ada lima perbandingan yang relevan seperti untuk kenyamanan, perbandingan dari masing-masing tingkat kinerja dengan standar yang sesuai menunjuk pada tahap evaluasi yang relevan. Hubungan ini ditunjukkan pada tabel 2 di bawah ini.

Tabel 4. Tahap Evaluasi DEM

<i>Stage</i>	<i>Performance</i>	<i>Standard</i>
I	<i>Program Design</i>	<i>Design Criteria</i>
	<i>Input Dimension</i>	
	<i>Process Dimension</i>	
	<i>Output Dimension</i>	
II	<i>Program Operation</i>	<i>Program Design</i>
		<i>Input Dimension</i>
		<i>Process Dimension</i>
III	<i>Program Interim Products</i>	<i>Program Design</i>
		<i>Process Dimension</i>
		<i>Output Dimension</i>
IV	<i>Program Terminal Products</i>	<i>Program Design</i>
		<i>Output Dimension</i>
V	<i>Program Cost</i>	<i>Cost of Other Program with Same Product</i>

(Sumber: Provus 1969: 10)

Pada semua tahap, beberapa indikator kinerja yang diperoleh dibandingkan dengan standar yang berfungsi sebagai "kriteria kinerja". Pada tahap I, deskripsi desain program akan diperoleh informasi "*kinerja*". Kinerja ini dibandingkan dengan standar "kriteria desain". Perbedaan antara kinerja dan standar yang bertanggung jawab sebagai pengelolaan program. Pada tahap II standar untuk perbandingan adalah program desain yang diperoleh dalam informasi program tahap I. Kinerja terdiri dari pengamatan dari lapangan mengenai instalasi program. Informasi

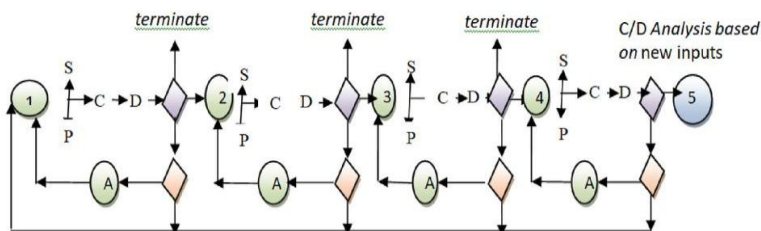
perbedaan dapat digunakan oleh pengelola program untuk mendefinisikan kembali program atau prosedur perubahan instalasi. Pada tahap III standar adalah bagian dari desain program yang menggambarkan hubungan antara proses program dan produk sementara. Informasi perbedaan digunakan baik untuk mendefinisikan kembali proses dan hubungan proses untuk produk sementara atau untuk kontrol yang lebih baik proses yang digunakan di lapangan. Pada tahap IV standar adalah bagian dari rancangan program yang mengacu pada tujuan terminal. Program informasi kinerja terdiri dari langkah-langkah kriteria yang digunakan untuk memperkirakan efek terminal proyek. Pada titik waktu ini, para pengambil keputusan memiliki lebih dari satu proyek dengan hasil yang serupa yang tersedia bagi mereka untuk dianalisis, mereka dapat memilih untuk melakukan analisis biaya-manfaat untuk menentukan efisiensi program. "Kriteria Desain" merupakan asumsi dasar sebagai standar bagi semua kriteria, mengandung tiga elemen dasar. Unsur-unsur dasar dari program apapun adalah Input, Proses, dan Output.

DEM menggunakan tahap dan kategori konten untuk memungkinkan perbandingan (Provus, 1969: 5-14; Kaufman & Thomas, 1980: 137). Tahapan itu adalah:

- 1) Tahap 1 *definition stage* (program didefinisikan). Untuk desain program dengan terlebih dahulu mendefinisikan input yang diperlukan, proses, dan output, kemudian mengevaluasi kelengkapan dan konsistensi internal dari desain, staf program mengorganisir a) gambaran tujuan, proses, atau aktivitas dan kemudian b) menggambarkan sumber daya yang diperlukan. Harapan atau standar ini adalah dasar dimana evaluasi berkelanjutan tergantung.
- 2) Tahap 2 *installation stage* (instalasi program). Untuk menilai tingkat instalasi program terhadap standar awal. Desain/definisi program menjadi standar baku untuk diperbandingkan dengan penilaian operasi awal program. Gagasannya adalah untuk menentukan sama dan sebangun, sudah atau tidaknya program telah diterapkan sebagaimana desainnya.

- 3) Tahap 3 *product stage* (proses program/*interim products*). Untuk menilai hubungan antara variabel yang akan berubah dan proses yang digunakan untuk mempengaruhi perubahan. Menggambarkan hubungan antara proses program dan produk sementara. Evaluasi ditandai dengan pengumpulan data untuk menjaga keterlaksanaan program. Gagasannya adalah untuk memperhatikan kemajuan kemudian menentukan dampak awal, pengaruh, atau efek.
- 4) Tahap 4 *product stage* (*product program/terminal products*). Untuk menilai desain program sampai tujuannya (tujuan, hasil). pengumpulan data dan analisa yang membantu ke arah penentuan tingkat capaian sasaran dari *outcome*. Dalam tahap 4 ini pertanyaannya adalah "Apakah sasaran program telah dicapai?" Harapannya adalah untuk merencanakan *follow up* jangka panjang pemahaman atas dampak.
- 5) Tahap 5 *cost-benefit analysis* (program cost), yakni melakukan analisis biaya-manfaat untuk menentukan efisiensi program. Menunjukkan peluang untuk membandingkan hasil dengan yang dicapai oleh pendekatan lain yang serupa.

Pada model ini, proses evaluasi pada langkah-langkah dan isi kategori sebagai cara memfasilitasi perbandingan pencapaian program dengan standar, sementara pada waktu yang sama mengidentifikasi standar yang digunakan untuk perbandingan di masa depan. Argumentasi Provus, bahwa semua program memiliki daur hidup (*life cycle*). Karena program terdiri atas langkah-langkah pengembangan, aktivitas evaluasi yang banyak diartikan adanya integrasi pada masing-masing komponennya, (Provus, 1969: 13).



(Sumber: Provus 1969: 13)

Gambar 11. Diagram Alir Tahap-Tahap Proses Evaluasi DEM

Dalam grafik S = standar, P = Program kinerja, C = membandingkan, D = perbedaan informasi, dan A = perubahan dalam kinerja program atau standar. Tahap V merupakan pilihan biaya-manfaat yang tersedia untuk evaluator hanya setelah empat tahap pertama telah dinegosiasikan.

Perhatikan bahwa penggunaan informasi perbedaan selalu mengarah pada keputusan untuk a) melanjutkan ke tahap berikutnya; b) mendaur ulang tanggung setelah terjadi perubahan dalam standar program atau operasi; c) mendaur ulang ke tahap pertama, atau d) menghentikan proyek. Dari sudut manajer program, perbedaan informasi memungkinkan dia untuk menunjukkan suatu kelemahan dalam program ini adalah untuk salah satu dari dua tujuan: untuk mengubah operasi program, atau untuk mengubah spesifikasi tempat program beroperasi. Seorang pengawas sekolah atau dewan pendidikan akan peduli dengan laju pergerakan proyek melalui tahap evaluasi dengan informasi perbedaan pada setiap tahap yang diberikan. Umumnya, lama waktu yang dibutuhkan untuk sampai ke tahap dua, tiga, dan empat, semakin besar biaya jika proyek gagal. Proyek lebih cepat bergerak ke stadium lanjut, umumnya dengan resiko kegagalan yang sedikit.

Pada masing-masing dari empat tahap, standar pasti dibandingkan dengan kinerja program yang sebenarnya untuk menentukan apakah ada kejanggalan. Penggunaan perbedaan informasi selalu mengarah ke salah satu dari empat pilihan. Lanjutkan ke tahap evaluasi berikutnya jika tidak ada perbedaan. Jika ada perbedaan, maka mendaur ulang melalui tahap yang ada setelah terjadi perubahan baik dalam standar program atau operasi. Jika nomor 2 tidak dapat dicapai, kemudian mendaur ulang kembali ke tahap 1. Items program mendefinisikan program, kemudian mulai mengevaluasi ketidaksesuaian lagi pada tahap 1. Jika nomor 3 tidak dapat dicapai, maka mendaur ulang program lagi.

3. Persyaratan DEM

DEM yang paling efektif dalam situasi sebagai berikut:

- a) Ketika jenis evaluasi yang diinginkan adalah formal, dan program ini dalam formatif, bukan tahap sumatif.
- b) Ketika evaluasi didefinisikan sebagai manajemen informasi terus menerus menangani program perbaikan dan penilaian, dan evaluasi merupakan komponen dari pengembangan program.
- c) Tujuan dari evaluasi adalah untuk meningkatkan, mempertahankan atau mengakhiri program.
- d) Penekanan utama evaluasi adalah definisi program dan instalasi program.
- e) Peran evaluator adalah sebagai fasilitator, pemeriksa standar, pengamat perilaku yang sebenarnya, dan ahli desain.
- f) Bila pada setiap tahap evaluasi, kinerja program dibandingkan dengan tujuan program (standar) untuk menentukan perbedaan.
- g) Prosedur evaluasi program ini dirancang untuk mengidentifikasi kelemahan dan untuk membuat keputusan tentang koreksi atau penghentian.
- h) Konstruksi teori adalah bahwa semua program setiap tahap terus memberikan umpan balik satu sama lain.
- i) Kriteria untuk menilai program evaluasi meliputi apakah: (1) program telah memenuhi kriteria program yang dibentuk, (2) bahwa tindakan sebenarnya yang dilakukan dapat diidentifikasi, dan (3) tindakan dapat diambil untuk menyelesaikan semua perbedaan.





BAB 10

MESp 4Cs

Pembelajaran abad 21 menuntut terlaksananya 4 karakter keterampilan meliputi: (1) *Critical Thinking and Problem Solving*, (2) *Creativity and Innovation*, (3) *Collaboration*, dan (4) *Communication*, (4Cs) (Partnership, 2015; Griffin, McGaw, & Care, 2012), mempertegas bahwa 4 C's merupakan topik mendesak dalam agenda sistem pendidikan. *Guru dituntut* mentransfer 4C's kepada siswanya, agar berhasil di sekolah, di tempat kerja, dan dalam kehidupan yang lebih luas dan kompleks, (Wolters, 2010; Kareluik, Fahnoe, & Karr, 2013).

Pembelajaran berkualitas sangat dipengaruhi oleh peran guru; sebab guru setiap hari memiliki kesempatan mentransfer pengetahuan dan keterampilan kepada peserta didik. Kompetensi mengajar yang efektif jauh lebih penting daripada mengukur kualifikasi guru (Martínez, Schweig, & Goldschmidt, 2016). Hasil temuan penelitian menyimpulkan bahwa kompetensi dan persepsi tentang supervisi pembelajaran memegang peranan penting dalam meningkatkan produktivitas kerja guru, (Hamdani, & Krismadinata, 2017). Hal ini berarti, kualitas pembelajaran guru akan semakin meningkat, manakala dikuatkan dengan kualitas supervise yang dilakukan oleh kepala sekolah, (OECD, 2005; Sabandi, 2013; Daud, Dali, Khalid, & Fauzee, 2018).

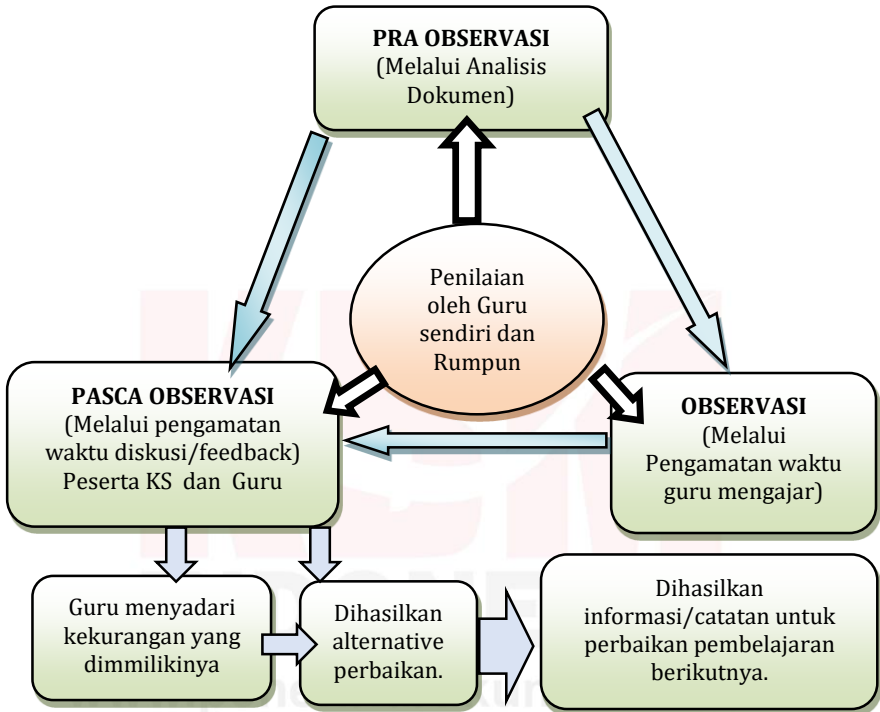
Keterlaksanaan 4C's oleh guru perlu disupervisi, untuk menilai sejauh mana keterampilan tersebut ditransfer oleh guru kepada siswanya. Kepala sekolah sebagai supervisor perlu pula memiliki keterampilan abad 21, melalui ketersediaan instrumen supervisi. Keberhasilan supervise ditentukan oleh: umpan balik dengan penilaian yang konstruktif, pemberian dukungan social dan emosional, hubungan interpersonal yang baik, aliansi pengawasan positif, kejelasan aspek penilaian yang terukur, pelibatan guru dalam penilaian, dan ketersediaan waktu yang cukup dengan pemberian ruang reflektif yang memadai, (Kilminster, 2007; Bambling, & King, 2014; Youngstrom, & Gentile, 2018; OECD, 2009; Bahri, 2014). Ada beberapa temuan terkait pelaksanaan supervise di lapangan, yaitu: kepala sekolah tidak memberi tahu keseluruhan tentang strategi mengajar, kurang cepat memberi umpan balik/saran, pelaksanaan supervisi akademik masih kurang efektif, tidak terjadwal dan tidak ada tindak lanjut, (Winaryati, & Mufnaety, 2012; Yunus, Lestari, & Raharjo, 2016). Rangkaian tahapan pelaksanaan supervisi ditunjukkan dengan table sebagai berikut:

Tabel 5. Tahapan Supervisi Pembelajaran

1.	PRA OBSERVASI.
	Guru melakukan pembelajaran di kelas. Guru sendiri dan guru serumpun melakukan observasi melalui analisis dokumen meliputi persiapan dan perencanaan yang dimiliki oleh guru.
2.	OBSERVASI
	Guru model dan guru serumpun melakukan observasi melalui pengamatan terhadap pelaksanaan pembelajaran yang dilakukan oleh guru model, berkenaan dengan keterampilan instruksionalnya, kemampuannya mengimplementasikan 4 Cs, keterampilan guru dalam mengelola kelas, dan pengetahuan guru tentang perkembangan siswanya.
3.	PASCA OBSERVASI. a) Guru menyerahkan data isian dari kegiatan pra observasi dan observasi kepada kepala sekolah. b) Kepala sekolah melakukan analisis data isian dari guru. c) Guru dan KS melakukan diskusi sebagai kegiatan <i>feedback</i> terkait dengan bagaimana kesesuaian rencana dengan pelaksanaan, respon pembelajaran, dan dampak <i>feedback</i> bagi perbaikan pembelajaran. d) Peran KS dalam forum diskusi sebagai fasilitator.

- e) Dari kegiatan diskusi, dihasilkan catatan-catatan.
- f) Informasi digunakan oleh guru model untuk melakukan perbaikan pembelajaran pada kegiatan berikutnya.

Gambaran proses supervise diilustrasikan sebagai berikut:

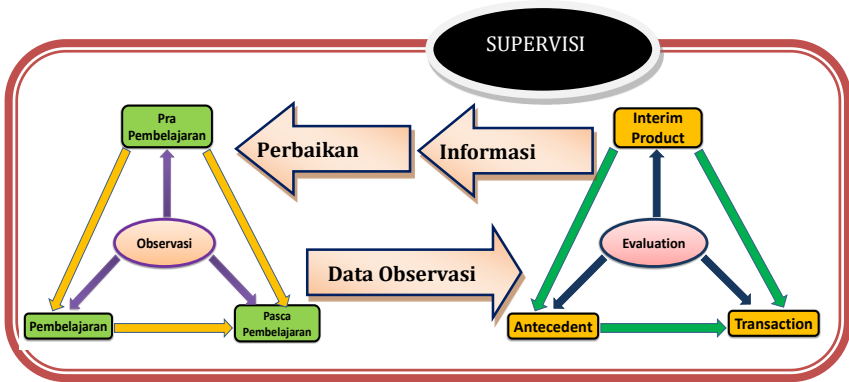


Gambar 12. Alur Pelaksanaan Supervise Pembelajaran

Keterlaksanaan supervisi pembelajaran berbasis 4Cs ini, maka perlu dilakukan evaluasi. Terkait dengan hal ini, maka dibutuhkannya suatu alat evaluasi yang yang mengedepankan aktivitas timbal balik, efektif, efisien, berdampak dan berkelanjutan. Hubungan antara supervise dengan evaluasi adalah siklus. Hasil evaluasi adalah rekomendasi, yang ditindaklanjuti oleh aktivitas supervise untuk melakukan perbaikan pembelajaran, (Winaryati, & Mufnaety, 2012).

Diperlukan model evaluasi yang mendorong kerja evaluasi berkelanjutan untuk menilai aktivitas supervise pembelajaran

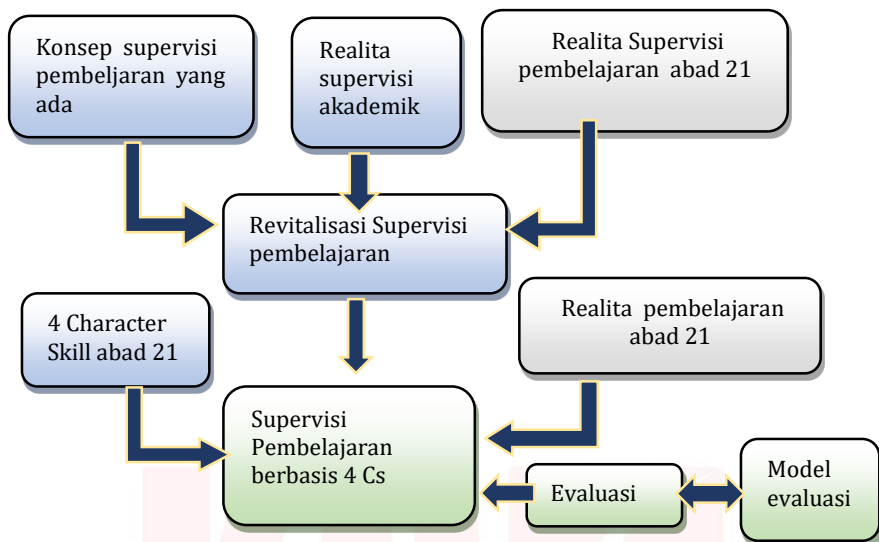
yang terjadi. Model evaluasi yang dikembangkan adalah Model Evaluasi Supervisi Pembelajaran berbasis 4C's (MESp 4C's), (Winaryati, Mardiana & Hidayat, 2020). Hubungan evaluasi pada supervise pembelajaran diilustrasikan sebagai berikut:



Gambar 13. Rangkaian tahap observasi- evaluasi-supervisi

Inti pengertian evaluasi adalah suatu proses **memberikan informasi** yang berarti dan berguna sebagai alternatif keputusan. Sedang pengertian supervisi akademik memberikan penegasan, sebagai suatu kegiatan bantuan pembinaan kearah **perbaikan**, sehingga akan **meningkatkan kualitas proses pembelajaran**. Hubungan antara evaluasi dengan supervisi adalah, bahwa hasil dari kegiatan evaluasi melalui model MESp 4Cs, dihasilkan informasi. Informasi ini dijadikan pengambilan kebijakan oleh pimpinan, dan pijakan bagi guru sendiri untuk melakukan perbaikan pembelajaran.

Menentukan alur Usulan Proyek/Produk Model



Sumber: Eny Winaryati, 2012b)

Gambar 14. Rancangan Konsep MESp 4Cs



BAB 11

METODE

1. Pendekatan yang digunakan

Pemilihan desain penelitian ini melibatkan beberapa langkah (Crotty, 1998; Sarantakos, 1998). Denzin dan Lincoln (1994) menyarankan pemilihan desain penelitian dimulai dari pendekatan kualitatif. Langkah ini diikuti dengan pemilihan paradigma teoretis penelitian yang dapat memberitahukan dan memandu proses penelitian (Searcy& Mentzer., 2003). Proses ini menghubungkan paradigma penelitian yang dipilih dengan dunia empiris. Dari perspektif ini desain penelitian dibuat dalam skema.

Metode yang digunakan dalam penelitian ini melalui kajian mendalam literatur yang terkait dengan evaluasi supervisi pembelajaran. Literatur diekstraksi dari jurnal dan sumber terbuka. Hasil kajian divalidasi dan didiskusikan dengan ahli supervisi dan teman sejawat (Muller, & Urbach, 2017).

Data dari interview mendalam ditulis lengkap dan dikelompokkan menurut format tertentu (Mansfield, 1986). Dengan cara ini, peneliti dapat mengidentifikasi informasi sesuai pemberi informasi dengan yang berkaitan dengan teori MESp 4Cs dan kontruksi empiris. Transkrip hasil interview kemudian dapat dianalisis dan key points untuk memudahkan coding dan pengklasifikasian (Searcy& Mentzer, 2003). Data dari observasi

dan arsip berupa catatan. selama observasi data diorganisir ke dalam form dengan kolom: tanggal, jam, peristiwa, partisipan, deskripsi peristiwa, dimana terjadinya, bagaimana terjadi, apa yang dikatakan, serta opini dan perasaan peneliti.

Data dari analisis catatan organisasi (arsip) dapat diorganisir ke dalam format tertentu. Format untuk mendukung data dari observasi dan interview (Finlay, 2006). Narasi yang telah diorganisir dapat dikelompokkan kedalam tema tertentu, dengan menggunakan code. Pengelompokan tema tersebut harus koheren dengan tujuan penelitian dan keyakinan yang dibuat oleh peneliti sesuai dengan fenomena penelitian.

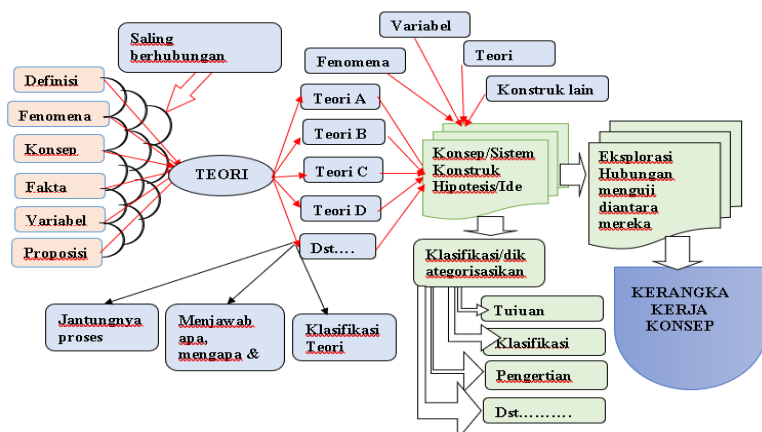
Hasil kajian yang telah dievaluasi peneliti, selanjutnya dikuatkan melalui data penelitian. Kerangka konsep (*conceptual frameworks*) diperoleh melalui langkah berikut;

1. Teori MESp 4Cs diperoleh dari konsep, fakta, fenomena, definisi, proposisi dan variable, (Muller, & Urbach, 2017; Jaccard & Jacoby, 2009; Gioia & Pitre, 1990; Nunnally & Bernstein, 1994; Suddaby, 2010). Teori MESp 4Cs menjelaskan konstruk, keterkaitan, dan kepada siapa konstruk itu berlaku. Teori ini juga menjelaskan kapan dan di mana berlaku. Teori ini juga menjelaskan proses kerja model (Whetten, 1989; Bacharach, 1989; Wheeler & Kip, 2019).
2. Konstruk MESp 4Cs menyajikan suatu kategorisasi atau klasifikasi dari benda-benda atau kejadian-kejadian, menjadi satu symbol, (Marx, 1976). Konsep MESp 4Cs ditambahkan konsep relevan dan kemudian dieksplorasi atau diuji hubungan di antara mereka (Khosro, & Pasand 2019; Wheeler & Kip, 2019).
3. Antar teori MESp 4Cs dicirikan oleh jaringan konstruksi dan hubungan yang unik. Sistem MESp 4Cs dikonstruksi, diproposisi, dan diasumsikan sebagai jaringan nomologis – semacam teori sidik jari (Gregor, 2006).
4. Konsep MESp 4Cs dikonstruksi hipotetis, yang terdiri dari beberapa variable, berjaringan dengan konstruk lainnya, (Snelbecker, 1974; Cronbach & Meehl 1955).
5. Metode berteori MESp 4Cs menggunakan pendekatan yang induktif dan deduktif. metode penalaran literatur MESp 4Cs

menggunakan cara retroduktif dan abduktif, (Ochara 2013; Alvesson & Sköldbberg 1994; Peirce, 1998).

Berdasarkan pemaparan diatas, maka dapat ditarik kesimpulan bahwa penelitian ini menggunakan metode Grounded Theory (GT). Metode ini merupakan penelitian kualitatif yang fokus untuk membuat kerangka kerja konseptual dengan cara membangun analisis induktif dari berbagai data, fenomena, informasi, teori menggunakan sejumlah prosedur sistematis, tujuannya untuk mengembangkan teori.

Alasan pemilihan metode diatas didasarkan dari berbagai rujukan hasil study peneliti terdahulu. Chun, Birks, & Francis, (2019) menyampaikan bahwa metode penelitian GT adalah metode penelitian kualitatif yang menggunakan sejumlah prosedur sistematis guna mengembangkan teori. Glaser dan Holton, (2004: hal 43), menyampaikan bahwa GT sebagai seperangkat hipotesis konseptual terintegrasi yang dihasilkan secara sistematis untuk menghasilkan teori induktif tentang area substantif. Charmaz, K, (2009), mendefinisikan GT sebagai metode melakukan penelitian kualitatif yang berfokus pada pembuatan kerangka kerja konseptual atau teori melalui membangun analisis induktif dari data, (hal. 187). Berikut gambaran alur penerapan metode GT dan bagaimana kerangka konsep itu dibangun. Kelima langkah diatas disusun secara sistematis, berdasarkan Gambar 2.



Gambar 15. Alur Kerangka Konseptual Model Evaluasi Supervise Pembelajaran

Sumber: Winaryati, Mardiana & Hidayat (2020).

Proses keabsahan penelitian ini menggunakan model Triangulasi. Triangulasi dalam penelitian ini peneliti menggunakan berbagai sumber data, teori, metode dan investigator agar informasi yang disajikan konsisten. Peneliti dapat menggunakan lebih dari satu teori, lebih dari satu metode (interview, observasi dan analisis dokumen). Di samping itu, peneliti dalam melakukan interview dari expert dan menginterpretasikan temuan dengan pihak lain (Ajnen & Fishbein 1977). Prosedur ini dilakukan dengan cara mencari tema dan kategori yang konsisten dan menerapkan proses tertentu untuk membuktikan ketidakbenaran temuan tersebut.

Cara ini dilakukan untuk menunjukkan mengapa model evaluasi MESp 4Cs diadopsi. Aspek ini perlu diungkapkan, untuk memverifikasi kredibilitas informasi.

2. Cara Membangun Kerangka Konsep

Alur kerangka konseptual di atas (gambar 3) digunakan untuk membangun MEPs 4Cs. Desain penelitian diawali dengan pemilihan paradigma teoretis penelitian yang dapat memberitahukan dan memandu proses penelitian (Searcy & Mentzer, 2003). Proses ini menghubungkan paradigma penelitian yang dipilih dengan dunia empiris, domain actual yang akan dikembangkan dan fenomena realita yang ada. Dari perspektif ini, desain penelitian dibuat dalam skema. Kerangka kerja konseptual berfungsi sebagai "peta" atau "kemudi" yang akan memandu mewujudkan tujuan pembahasan (Masías, 2005).

Kerangka konseptual adalah tampilan gambaran visual dari teori dengan fenomena yang dipelajari. Peta konsep terdiri dari dua hal: konsep dan hubungan di antaranya (Maxwell, 1996). Kerangka kerja konseptual adalah pemahaman peneliti tentang bagaimana variabel-variabel tertentu dalam studinya terhubung satu sama lain.

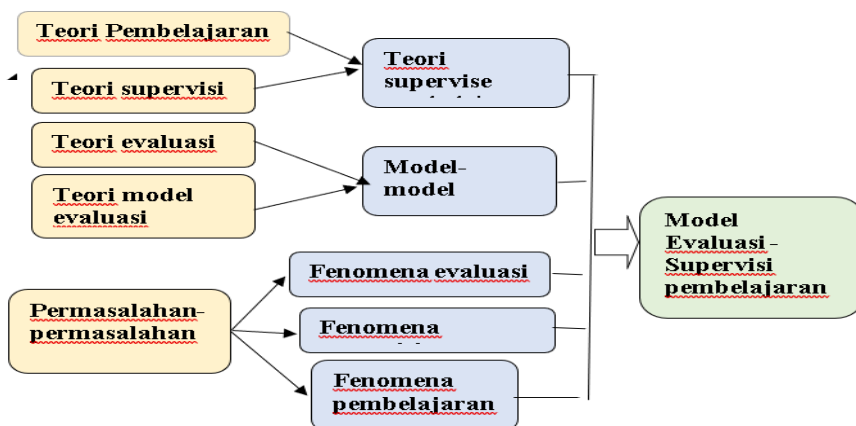
Makalah diekstraksi dari jurnal dan sumber terbuka yang ditinjau oleh sejawat dan ahli, diperkuat dengan data penelitian yang telah dilaksanakan oleh peneliti. Pengertian dari berbagai sumber saling dihubungkan agar kerangka konseptual yang

dirancang sesuai kaidah. Kerangka konseptual dibangun berdasarkan kata kunci berikut:

1. Teori adalah kumpulan variable, masalah untuk diatasi, definisi-definisi, proposisi-proposisi yang saling berhubungan, memberikan pandangan yang sistematis tentang fenomena dengan menspesifikasikan relasi-relasi yang ada diantara beragam variabel, dengan tujuan untuk menjelaskan fenomena yang ada. Teori memiliki komponen-komponen yang terdiri atas: konsep, fakta, fenomena, definisi, proposisi dan variable (Mullerl, & Urbach, 2017; Jaccard & Jacoby, 2009; Gioia & Pitre, 1990; Nunnally & Bernstein, 1994; Suddaby, 2010).
2. Sebuah teori harus menjelaskan "apa konstruknya, bagaimana dan mengapa saling terkait, dan kepada siapa konstruk itu berlaku, kapan dan di mana itu berlaku, dan bagaimana sesuatu itu bekerja (Whetten, 1989; Bacharach, 1989, p. 496; Wheeler, 2019).
3. Teori adalah penjelasan tentang fenomena yang telah berulang kali diuji dan ditemukan konsistennya dalam rentang waktu yang panjang (Wheeler, 2019).
4. Konstruk merupakan semacam konsep, yaitu menyajikan suatu kategorisasi atau klasifikasi dari benda-benda atau kejadian-kejadian, menjadi satu symbol (Marx. 1976). Antar konstruk dapat berjaringan (Snelbecker, 1974; Cronbach & Meehl, 1955). Dalam kerangka kerja konseptual dapat ditambahkan konsep/konstruksi/variabel sendiri, yang menurut kita relevan dan kemudian dieksplorasi atau diuji hubungan di antara mereka (Khosu, 2019; Wheeler, 2019). Konstruksi, adalah konsep dasar yang mencakup teori, dimana antar konstruk dihubungkan oleh proposisi, dan teori dibuat jika semua elemen ini bersatu (Gregor, 2006). Konsep juga bisa berupa ide (Weeler, 2019).
5. Antar teori dicirikan oleh jaringan konstruksi dan hubungan yang unik, dan antar teori-teori serupa dihubungkan berdasarkan batas-batasnya. Sistem konstruksi, proposisi, dan asumsi yang dihasilkan juga disebut sebagai jaringan nomologis–semacam teori sidik jari (Gregor, 2006).

6. Konsep adalah sebagai suatu konstruk hipotetis, yang terdiri dari beberapa variable, dapat pula berjaringan dengan konstruk lainnya (Snelbecker, 1974; Cronbach & Meehl, 1955).
7. Penelitian ini memuat domain real, actual dan empirik, sehingga termasuk realisme kritis, dengan menggunakan pendekatan abduktif. Dasar pendekatannya adalah menginterpretasikan dan merekontekstualisasi sebuah fenomena dengan sebuah kerangka konseptual atau sejumlah konsep pemikiran, dan menjadi kerangka konseptual yang baru. Pendekatan abduktif diartikan sebagai pengembangan hipotesis kerja pendahuluan sebelum analisis data induktif (Peirce, 1998).

Berdasarkan alur kerangka konseptual yang digambarkan pada gambar 16. Berikutnya dapatlah dibuat tahapan-tahapan dalam menyusun kerangka konseptual MESp 4Cs.



Gambar 16. Kerangka Konsep MESp 4 Cs

BAB 12

PEMBAHASAN

1. Kerangka Konseptual Awal Yang Dibangun

Kerangka konseptual model, dibangun berdasarkan analisis beberapa teori, fakta-fakta yang dipeoleh, fenomena yang terjadi, kemudian dikonstruksi dengan pendekatan induktif. Ada beberapa variabel yang harus diperoleh datanya. Analisis mendalam telah dilakukan pada saat penelitian terkait dengan pelaksanaan supervise. Diperoleh fenomena: kepala sekolah tidak memberi tahu keseluruhan tentang strategi mengajar, kurang cepat memberi umpan balik/saran, pelaksanaan supervisi akademik masih kurang efektif, tidak terjadwal dan tidak ada tindak lanjut. Selain itu diperoleh beberapa data lapangan yang memperkuat fenomena di atas. Fakta lapangan diperoleh suatu gambaran bahwa, perlu adanya evaluasi terhadap pelaksanaan supervise.

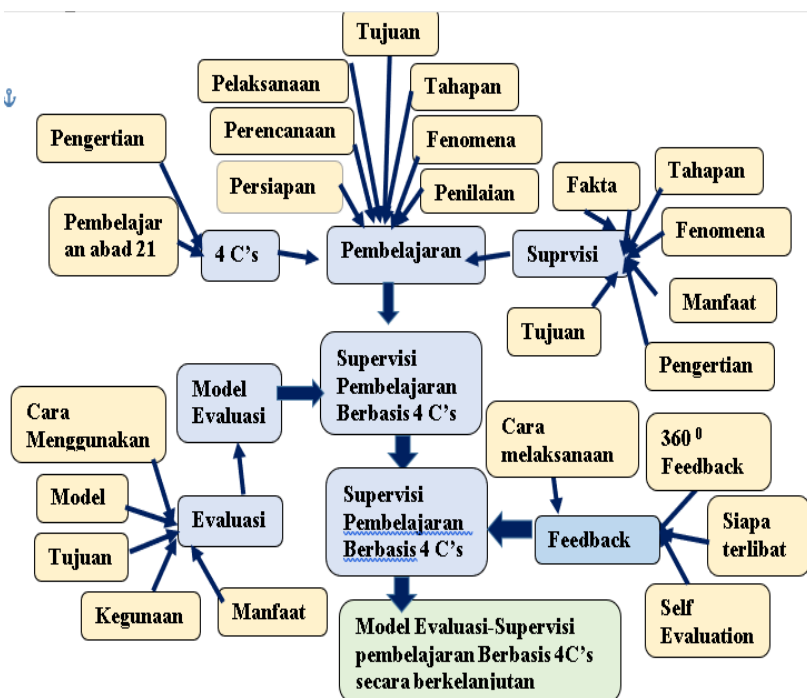
Perlunya pemahaman yang lengkap terkait dengan beberapa teori tentang pembelajaran, supervise, dan evaluasi. Selain itu perlu pemahaman dan keterlaksanaan terkait kesiapan perencanaan, proses pembelajaran dan penilaian serta umpan balik. Definisi mengajar, ketersediaan silabus, RPP, tujuan pembelajaran, indikator, metode dan media yang digunakan, dan penilaian harus telah disiapkan.

Karakter keterampilan (4C's) pembelajaran abad 21, merupakan tuntutan yang harus dipahami dan dilaksanakan. Guru mentransfer keterampilan diatas pada siswa, sehingga siswa memiliki kompetensi 4Cs. Harapannya siswa telah memiliki kesiapan dalam menghadapi kebutuhan, persoalan dan tantangan di masa depan. Supervisor perlu melakukan kepengawasan sejauh mana 4Cs telah dilaksanakan oleh guru dalam pembelajarannya. Evaluasi dilakukan untuk mendapatkan data sejauh mana pelaksanaan supervise berbasis 4C's dilaksanakan.

Adanya pemahaman yang lengkap terkait dengan pengertian, definisi, tujuan, kebutuhan dan harapan dari supervise pembelajaran. Apa yang harus dilakukan supervisor ketika guru melakukan persiapan, bagaimana pelaksanaan proses pembelajaran, dan penilaian dilakukan, serta umpan balik secara reflektif dilaksanakan. Sekaligus menjadi dasar bagaimana umpan balik 360 derajat dilaksanakan, strateginya seperti apa, dan siapa saja yang akan terlibat dalam umpan balik. Pertanyaan apa, mengapa dan bagaimana ini menjadi dasar peneliti untuk menyusun instrument supervise dan evaluasi.

Adanya pengertian dan relevansi antara supervise dengan evaluasi. Bagaimana cara konstruksi model evaluasi dibangun, untuk menjawab apa, mengapa dan bagaimana aktivitas evaluasi dijalankan. Beberapa data, teori, konsep awal, dan bukti empiris menjadi variable-variabel yang menjadi saling berjarangan untuk membentuk suatu konstruksi dan menjadi kerangka konsep model yang diharapkan.

Pemilihan model evaluasi yang digunakan didasarkan pada definisi, tujuan, fungsi, manfaat, dan cara penerapannya. Ada beberapa model evaluasi, yang masing-masing memiliki kelebihan dan keterbatasan. Hal inilah yang menjadi dasar perlunya kombinasi model evaluasi. Model evaluasi yang ditetapkan telah mengakomodir relevansinya dengan pembelajaran, 4 C's dan supervise pembelajaran. Kerangka konseptual model evaluasi ini disusun berdasarkan langkah-langkah yang alurnya didasarkan pada urutan yang sistematis.



Gambar 17. Kerangka konsep membangun model evaluasi
 Sumber: Winaryati, Mardiana & Hidayat (2020).

2. Landasan Model Evaluasi

Evaluasi adalah suatu proses memberikan informasi yang berarti dan berguna sebagai alternatif keputusan. Evaluasi merupakan proses atau kegiatan pemilihan, pengumpulan, analisis dan penyajian informasi yang sesuai untuk mengetahui sejauh mana suatu tujuan program, prosedur produk atau strategi yang dijalankan telah tercapai. Harapannya dapat bermanfaat bagi pengambilan keputusan serta dapat menentukan beberapa alternatif keputusan untuk program selanjutnya (Stark & Thomas, 1994: 12).

Berkenaan dengan hasil informasi dari proses evaluasi, ada dua pengembangan pendekatan: 1) berfokus pada bagaimana hasil evaluasi digunakan; 2) usaha untuk melakukan pengembangan metode yang digunakan. Pendekatan pengembangan metode memberikan kebebasan memilih metode yang digunakan untuk

mengumpulkan, menganalisis, dan menafsirkan data. Tumpuan dari kedua pendekatan di atas, adalah bahwa proses evaluasi menumbuhkan komitmen untuk perbaikan program dengan pendekatan yang lebih analitis menghasilkan perubahan yang diinginkan

3. Membangun Konten Evaluasi Supervisi Pembelajaran Berbasis 4 C's

Tahapan langkah penyusunan model evaluasi MESp 4Cs untuk supervise pembelajaran berbasis 4Cs adalah sebagai berikut:

1. Kompetensi mengajar yang dibangun berisi domain: guru, ahli pengetahuan konten, fasilitator proses pembelajaran, Penyelenggara pembelajaran (Tigelaar, *et.al* 2004). Pendekatan pembelajaran lebih menekankan akan peran guu sebagai fasilitator, dengan pendekatan pengajaran menjadi lebih berpusat pada siswa. Siswa dipandang sebagai pelajar yang aktif dan mampu mengatur diri sendiri menciptakan makna dari pengalamannya sendiri dengan cara yang bermakna (Alexander, Harris, & Alexander, 1998). Oleh karenanya penguatan kompetensi guru untuk meningkatkan proses belajar-mengajar di sekolah sangat penting.
2. Penguatan kompetensi guru untuk meningkatkan proses belajar-mengajar di sekolah, sangat penting dilakukan. Selvi (2010) menyampaikan bahwa kontribusi semua kompetensi mengajar secara simultan atau bersama-sama dinyatakan signifikan memiliki pengaruh dalam meningkatkan kualitas kinerja dalam proses pembelajaran. Hal ini berarti proses pembelajaran menjadi sasaran yang harus dikedepankan. Oleh karenanya supervisi program pembelajaran harus dilakukan.
3. *Daresh, 1989 & Glickman, et al., 2007*), menyampaikan bahwa supervisi akademik merupakan serangkaian kegiatan membantu guru mengembangkan kemampuannya mengelola proses pembelajaran untuk mencapai tujuan pembelajaran. Hal ini mengindikasikan bahwa supervisi akademik tidak terlepas dari penilaian kinerja guru dalam mengelola

pembelajaran (Sergiovanni, 1987) dan meningkatkan kualitas pembelajaran (Kemendikbud, 2017; Prabowo, & Yoga, 2016). Supervisi pembelajaran dilakukan pada tahap perencanaan, pelaksanaan, dan penilaian pembelajaran, serta umpan balik pasca pembelajaran. Supervisi merupakan kegiatan pembimbingan yang bersifat manusiawi melalui hubungan yang demokratis, keterbukaan dan persahabatan (Bafadal, 1992; Maralih, 2014).

4. Supervisi yang sangat erat dengan aspek pelaksanaan proses pembelajaran adalah supervisi akademik. Supervisi akademik merupakan kegiatan pembinaan dengan memberi bantuan teknis kepada guru dalam melaksanakan proses pembelajaran, yang bertujuan untuk meningkatkan kemampuan profesional guru dan meningkatkan kualitas pembelajaran (APSI, 2006; Prabowo, & Yoga, 2016). Supervisi merupakan kegiatan pembimbingan yang bersifat manusiawi melalui hubungan yang demokratis, keterbukaan dan persahabatan (Bafadal, 1992; Maralih, 2014). Supervisi akademik merupakan upaya membantu guru-guru mengembangkan kemampuannya mencapai tujuan pembelajaran, dan meningkatkan kinerja guru, (Bahri, 2014; Suryani, 2015).
5. Tahap supervisi akademik terdiri dari tiga tahap yaitu: pra-observasi (pengamatan/pertemuan sebelum pembelajaran), observasi (pengamatan pembelajaran) dan pasca observasi (pertemuan balikan/*feedback*) (Depdikbud, 2017). Pra-observasi berisi persiapan dan perencanaan pembelajaran, sedang tahap observasi adalah pelaksanaan dari perencanaan yang telah disiapkan. Umpan balik dapat digunakan untuk meningkatkan efektivitasnya di ruang kelas, sekaligus sebagai penilaian formatif (Hattie, & Timperley, 2007; Shute, 2008; Darto, 2014). Pelaksanaan supervisi akademik oleh kepala sekolah dilakukan dalam tiga tahap yaitu: survei awal sebelum melakukan supervisi akademik, kunjungan kelas untuk mengetahui jalannya pembelajaran, dan mereview hasil

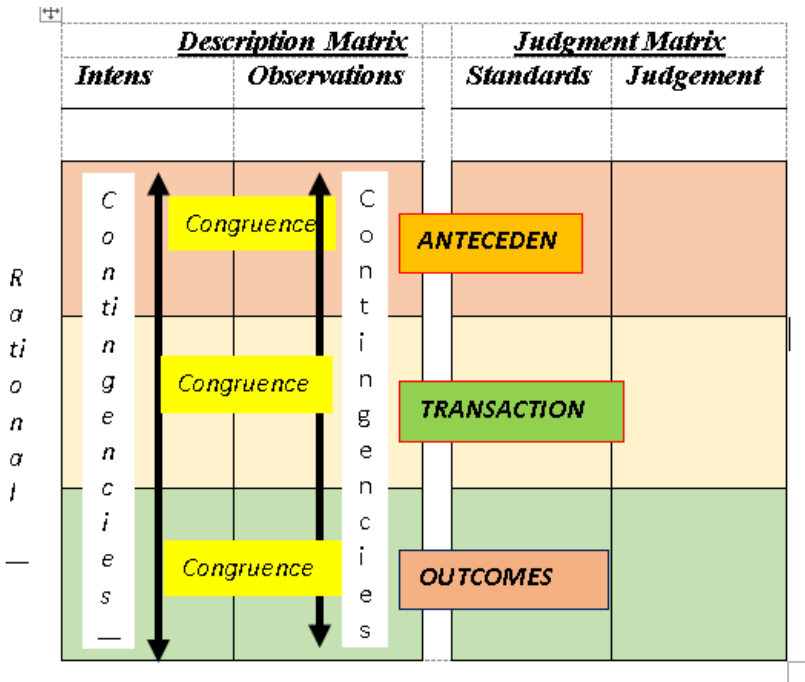
kunjungan kelas serta memberikan solusi yang tepat dalam mengatasi masalah yang dihadapi guru (Ajasan, 2016).

6. Harapannya melalui pelaksanaan supervise pembelajaran ini, dapat terlaksana secara berkelanjutan. Proses umpan balik secara reflektif ini dilakukan melalui supervisi pembelajaran. Supervisi yang lemah akan berdampak kualitas guru yang kurang dan kualitas pembelajaran yang tidak maksimal (Zhou, 2018). Hal ini mendasarkan bahwa berdasarkan kajian terkait umpan balik 360 derajat dapat terjadi suatu penilaian dari orang-orang yang bekerja di sekitar guru, baik dari atasan, teman sebaya (guru lainnya), dan bawahan (siswa), rekan kerja (staf lainnya).
7. Pembelajaran di abad 21 menuntut dimilikinya 4 (empat) keterampilan yang terdiri dari: (1) *Communication*, (2) *Collaboration*, (3) *Critical Thinking and Problem Solving*, dan (4) *Creativity and Innovation*. Empat keterampilan abad 21 diatas seringkali diistilahkan dengan keterampilan 4C's. 4 Cs ini harus dimiliki baik oleh siswa, guru dan juga kepala sekolah. Maka supervise yang dilakukan oleh kepala sekolah dapat melakukan penilaian ketercapaian 4 Cs ini dalam pembelajaran, dan evaluasi dilakukan untuk menilai sejauh mana pelaksanaan supervise berbasis 4C's.
8. Berdasarkan substansi dari tahap praobservasi dan observasi di atas, mengindikasikan:
 - a) Kesesuaian antara perencanaan dengan pelaksanaan.
 - b) Supervisi menitikberatkan kesenjangan antara formulasi standar yang telah ditetapkan dengan realita yang terjadi (hasil pengamatan).
 - c) Pada tahap praobservasi dan observasi, masing-masing memiliki proses pentahapan (meliputi input, proses dan output), sampai dihasilkannya suatu produk. Produk dari tahap praobservasi akan memberi sumbangan/ mempengaruhi pada tahap observasi, demikian juga terhadap pasca observasi.
9. Berdasarkan substansi dari tahap pra-observasi dan observasi di atas, mengindikasikan: (a) Kesesuaian antara perencanaan

dengan pelaksanaan. (b) Supervisi menitikberatkan kesenjangan antara formulasi standar yang telah ditetapkan dengan realita yang terjadi (hasil pengamatan). (c) Pada tahap pra-observasi dan observasi, masing-masing memiliki proses pentahapan (meliputi input, proses dan output), sampai dihasilkannya suatu produk. Produk dari tahap pra-observasi akan memberi sumbangan/mempengaruhi pada tahap observasi, demikian juga terhadap pasca observasi.

4. Membangun kerangka MESp 4C's

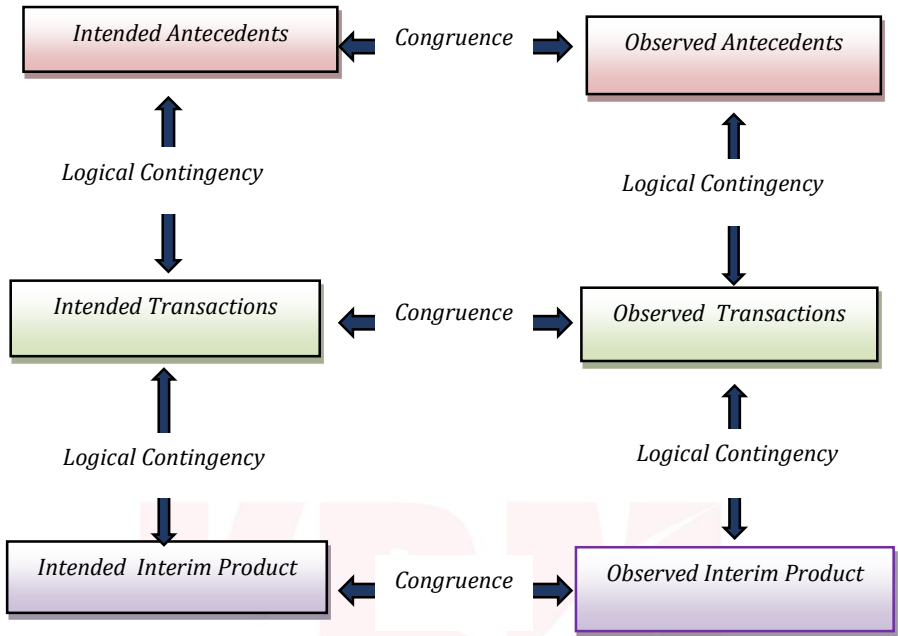
Berdasarkan tahapan supervise pembelajaran diatas, maka *Stake Model (Countenance Evaluation Model)* merupakan pilihan yang tepat. Stake membagi objek evaluasi mejadi 3 (tiga) kategori, yaitu: *Antecedents*, *Transaction*, *Outcomes*. *Antecedents* yaitu sumber/model/input yang ada pada sistem itu untuk dikembangkan, seperti tenaga, keuangan, karakteristik siswa, dan tujuan yang ingin dicapai, kondisi yang ada sebelum instruksi yang mungkin berhubungan dengan hasil. *Transaction*, mencakup rencana kegiatan maupun proses pelaksanaannya pada saat di lapangan, termasuk didalamnya urutan kegiatannya, penjadwalan waktu, bentuk interaksi guru-murid, cara menilai hasil belajar, dsb. Keterlibatan secara berurutan atau pertemuan yang dinamis yang merupakan proses instruksi. *Outcomes* merupakan hasil yang dicapai oleh para siswa, reaksi guru terhadap sistem tersebut, dan efek samping dari sistem yang bersangkutan, dampak dari pengalaman instruksional, (Popham, 1993: 5-15; Stake, 1977: 372-390; Fernandes, 1984:8-10; Wood, 2001: 18-27; Owston, 2008).



Sumber Stake's, 2000:351.

Gambar 18. Implementasi Stake Model

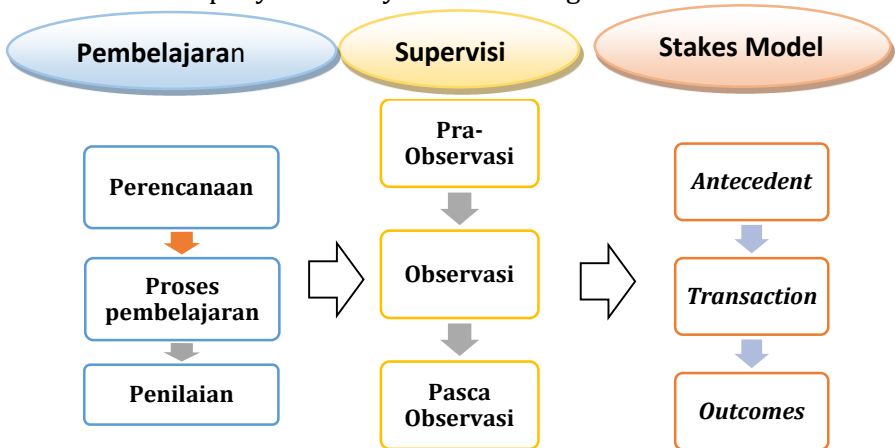
Pendekatan model Stake's ini ada dua yaitu: *contingencies* dan *congruence*. *Contingencies* adalah hubungan logis antara 3 dimensi di atas (*antecedent, transaction, outcomes*). *Congruence* adalah kesesuaian antara yang diharapkan (kriteria) dan yang terjadi/dihasilkan baik rencana, proses dan hasil, (Popham, 1993: 5-15; Stake, 2000: 350-351). *Outcome* tergantung pada *transaction* dan pada kondisi *antecedent*. Melalui analisis secara simultan dari beberapa aspek, evaluasi akan menemukan interaksi yang signifikan untuk perbaikan. Gambar alur evaluasi tertera pada gambar di bawah ini.



Sumber: Stake, 1977: 372-390

Gambar 19. Proses diskripsi data contingencies dan congruence

Berdasarkan relevansi tahapan pembelajaran yang harus dilakukan oleh guru, ada kesesuaian dengan tahapan supervise pembelajaran, dan adanya kesesuaian dengan model evaluasi Stake's. Alur penyesuaiannya adalah sebagai berikut:



Gambar 20. Alur penyesuaian Tahap Pembelajaran-Supervisi-Stakes Model Evaluasi

Hasil penelitian diatas menyampaikan bahwa keberhasilan supervise pembelajaran diantaranya adalah adanya kegiatan umpan balik yang konstruktif dan adanya ruang reflektif yang memadai. Hasil feedback ini memberi ruang terlaksananya evaluasi yang berkelanjutan. Terkait hal ini maka MESp 4C's perlu adanya umpan balik sehingga akan terlaksanan kegiatan bersiklus. Gagasan model di atas selaras dengan yang disampaikan oleh Tang, & Chow, (2007), hasil penelitiannya menyampaikan bahwa komunikasi umpan balik dari hasil observasi pembelajaran sebagai diskusi hasil mentoring dan peer pembinaan. Konten feedback dihasilkan suatu penilaian yang berorientasi pada peningkatan pembelajaran dalam konteks evaluasi sumatif untuk menghasilkan wawasan yang lebih dalam dari supervisi praktik mengajar.

Penjelasan diatas dikuatkan dengan penjelasan bahwa pada tahap pra-observasi dan observasi dihasilkannya produk sementara untuk menuju ke tahap berikutnya. Pasca observasi adalah produk sementara sebagai hasil umpan balik untuk dilakukan perbaikan pada tahap pra-observasi berikutnya. Berkenaan dengan pengertian ini berarti dihasilkannya produk sementara. Disamping itu outcome merupakan hasil jangka panjang. Maka perlu adanya modifikasi pada tahap *outcome* dengan model evaluasi yang lainnya. Supervisi adalah menilai antara standar yang ada dengan kinerja yang dilakukan, maka penilaian terhadap kesenjangan yang terjadi menjadi suatu pilihan. Model evaluasi yang mendasarkan teori kesenjangan adalah *The Descrpancy Evaluation Model (DEM)*.

Organizational Elements Model (OEM) dikembangkan oleh Roger Kaufman sebagai alat yang dapat digunakan untuk mengidentifikasi elemen yang berbeda dalam suatu sistem. Sebuah sistem adalah "satu set komponen yang saling terkait yang bekerja sama untuk mencapai tujuan bersama", (Porter, 2005a). OEM menyediakan kerangka sistemik, merancang dan menerapkan cara yang efektif untuk mencapai hasil akhir yang diinginkan. Kaufman membagi OEM menjadi lima unsur yang berinteraksi satu sama lain: Input, proses, dan produk, output, *outcome*. Produk dan output adalah hasil akhir yang terjadi dalam organisasi; *outcome* adalah hasil akhir luar organisasi, (Chyung, 2000). Input: adalah bahan

baku; proses adalah bagaimana untuk melakukannya; produk adalah hasil sementara dalam proses; output adalah prestasi organisasi, sebagai produk yang dikirim kepada masyarakat; outcome adalah efek bagi masyarakat, (Porter, 2005). Kaufman (2006: 6-16), memberikan penegasan bahwa OEM adalah suatu tahapan bagi evaluator menentukan model evaluasi yang tepat untuk digunakan.

<i>Model</i>	<i>Input</i>	<i>Proses</i>	<i>Produk</i>	<i>Output</i>	<i>Outcome</i>
<i>Formatif Sumatif</i>		<i>Formatif</i>	<i>Sumatif</i>	----- -----	
<i>CIPP</i>	<i>Context Input</i>	<i>Proses</i>	<i>Produk</i>	----- -----	
<i>CSE</i>	<i>Stage 1</i>	<i>Stage 2</i>	<i>Stage 3</i>	<i>Stage 4</i>	<i>Stage 5</i>
<i>Stake</i>	<i>Antecedent</i>	<i>Transaction</i>	<i>Outcome</i>	----- -----	
<i>Tyler</i>	<i>Goal 1 Identification</i>	<i>Attainment</i>	----- -----		
<i>Provus (DEM)</i>	<i>Installation</i>	<i>Implementa Tion</i>	<i>Fulfil Objectives?</i>	----- <i>Cost</i> ----- ----- <i>Benefit</i>	
			<i>(Depending upon objectives)</i>		
<i>Sciven's Goal Free</i>	-----				
<i>Stake Resp</i>	-----				

Sumber: Kaufman & Thomas (1980: 137).

Gambar 21. Hubungan OEM dengan Beberapa Model Evaluasi.

OEM adalah suatu tahapan bagi evaluator menentukan model evaluasi yang tepat untuk digunakan (Kaufman, 2006: 10-16). Berdasarkan pengertian ini, maka perlu dilakukan suatu pemetaan, untuk menempatkan tahap evaluasi yang tepat. Berdasarkan gambaran tahap OEM, maka tahap DEM yang tepat adalah *interim product* (produk sementara). *Interim product* memberi gambaran

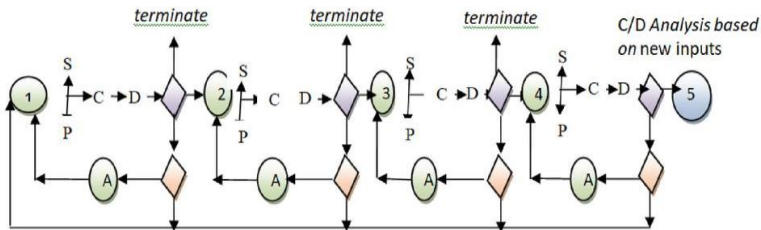
hubungan antara proses program dan produk sementara. Pemetaan melalui OEM selaras dengan kesamaan evaluasi yang mendasarkan teori kesenjangan. Model evaluasi yang memiliki substansi berdasarkan teori kesenjangan (antara standar dengan pengamatan, dan antara intensi dengan pelaksanaan) dari Stake's Model adalah *The Discrepancy Evaluation Model (DEM)*. Kedua model ini menitikberatkan teori kesenjangan, yakni sebagai alat untuk membuat pertimbangan (*judgement*) berdasarkan standar dan kinerja (Provus, 1969: 9,18). Perbedaannya pada Stake ada tambahan kegiatan *congruence* (kesesuaian antara *instens* dengan observasi), dan *contingencies* (hubungan logis *Antecedents, Transaction, Outcomes*). DEM ada proses memperbaiki, mendaur ulang atau menghentikan program.

The Discrepancy Evaluation Models (DEM), yang dirancang oleh Malcolm Provus pada tahun 1969 adalah cara yang efektif untuk mengevaluasi program-program akademik. DEM disebut dengan evaluasi kesenjangan program. Kesenjangan program adalah sebagai suatu keadaan antara yang diharapkan dalam rencana dengan yang dihasilkan dalam pelaksanaan program. Evaluasi kesenjangan dimaksudkan untuk mengetahui tingkat kesesuaian antara standar yang sudah ditentukan dalam program dengan penampilan aktual dari program tersebut. Standar adalah: kriteria yang telah dikembangkan dan ditetapkan dengan hasil yang efektif. Provus mendefinisikan evaluasi sebagai alat untuk membuat pertimbangan (*judgement*) atas kekurangan dan kelebihan suatu objek berdasarkan antara standar dan kinerja. Model ini juga dianggap menggunakan pendekatan formatif dan berorientasi pada analisis system, (Provus, 1969: 10-14; Steinmetz, 2000: 135).

DEM menawarkan pendekatan pragmatis sistematis untuk berbagai kebutuhan evaluasi. DEM dapat dimanfaatkan untuk struktur pengumpulan informasi penting baik untuk informasi pengambilan keputusan. Penekanan utama dari DEM adalah evaluasi diri dan perbaikan program secara sistematis.

DEM membagi tahap evaluasi menjadi 5 tahap, yaitu: *Program Design, Program Operation, Program Interim Products, Program Terminal Products, Program Cost*. Argumentasi Provus, bahwa

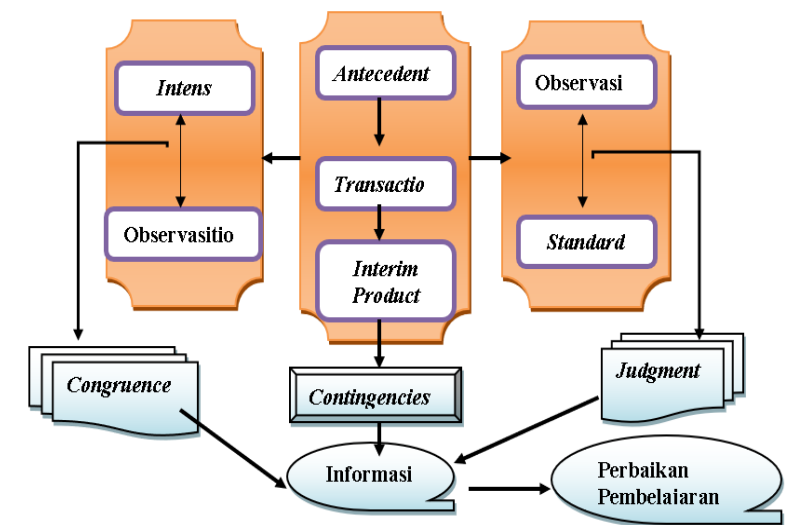
semua program memiliki daur hidup (*life cycle*). Karena program terdiri atas langkah-langkah pengembangan, aktivitas evaluasi yang banyak diartikan adanya integrasi pada masing-masing komponennya.



(Sumber: Provus 1969: 13)

Gambar 22. Diagram Alir Tahap Proses Evaluasi DEM

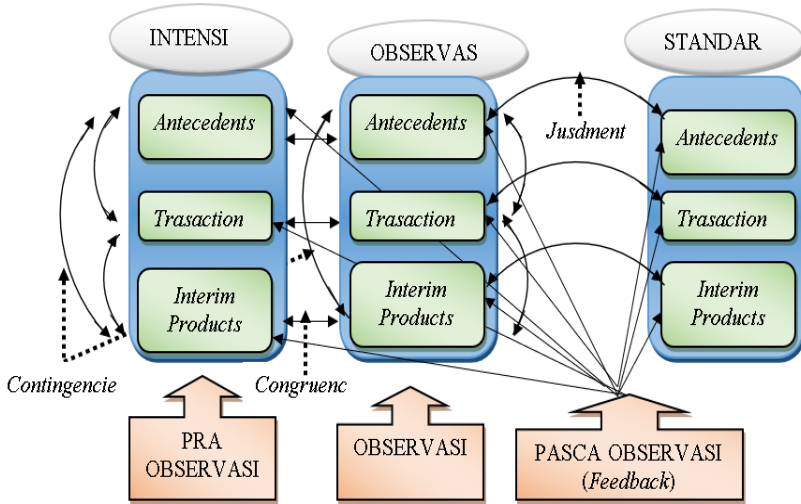
Berdasarkan tahapan dari kombinasi Stake's model dengan DEM menghasilkan tahapan model *Antecedents, Transaction, Interim Product, disingkat ATIp*. Dalam model evaluasi *ATIp* ada beberapa pengertian terkait standar, observasi dan intens. *Standard* adalah patokan/ukuran yang harus terpenuhi, dan yang diharapkan oleh stakeholder/pemerintah dan telah ditetapkan. Pada artikel ini rumusan standar didasarkan dari buku panduan supervise akademik (2017), 4 character keterampilan pembelajaran abad 21; Permendiknas RI, (No 21,22,23,24, tahun 2016). *Intens* adalah apa yang dimaksudkan oleh guru. *Observation* adalah apa yang dirasakan pengamat. Tahapan proses evaluasi *ATIp* digambarkan sebagai berikut:



Gambar 23. Tahapan Proses Evaluasi

Dari gambar ada beberapa keterangan tambahan yaitu: 1) *judgment* adalah kesenjangan/kesesuaian antara standar dengan observasi; 2) *congruencies* adalah kesenjangan/kesesuaian antara *intens* dengan *observation*; 3) *contingencies* adalah kesenjangan/hubungan antara *anteseden*, *transaction*, dan *interim product*. Hasil penilaian dari kegiatan *judgment*, *congruencies*, dan *contingencies*, dihasilkan informasi yang digunakan untuk perbaikan pembelajaran.

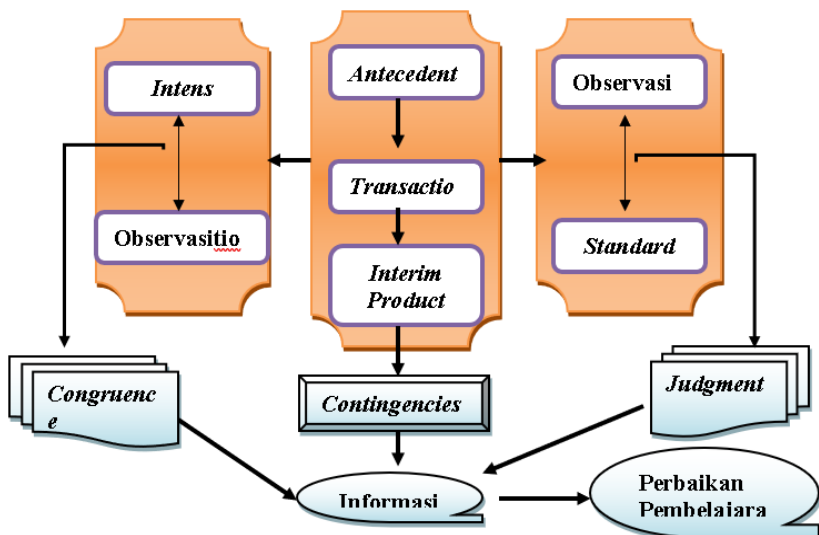
Berdasarkan penjelasan diatas dan diperjelas dengan gambar maka dihasilkan kerangka konsep model evaluasi pada supervise pembelajaran berbasis 4 C's. Persoalan feedback pasca supervise pembelajaran diadopsi melalui aktivitas evaluasi-supervisi yang berkelanjutan. Kerangka konsep MESP 4C's menghasilkan hubungan 6 cell Stake models yang dimodifikasi dan telah ada kesesuaiannya dengan tahapan supervisi pembelajaran. Enam cell MESP 4C's digambarkan sebagai berikut:



Gambar 24. Hubungan 6 Cell MESP 4C's dan Tahapan Supervisi Pembelajaran

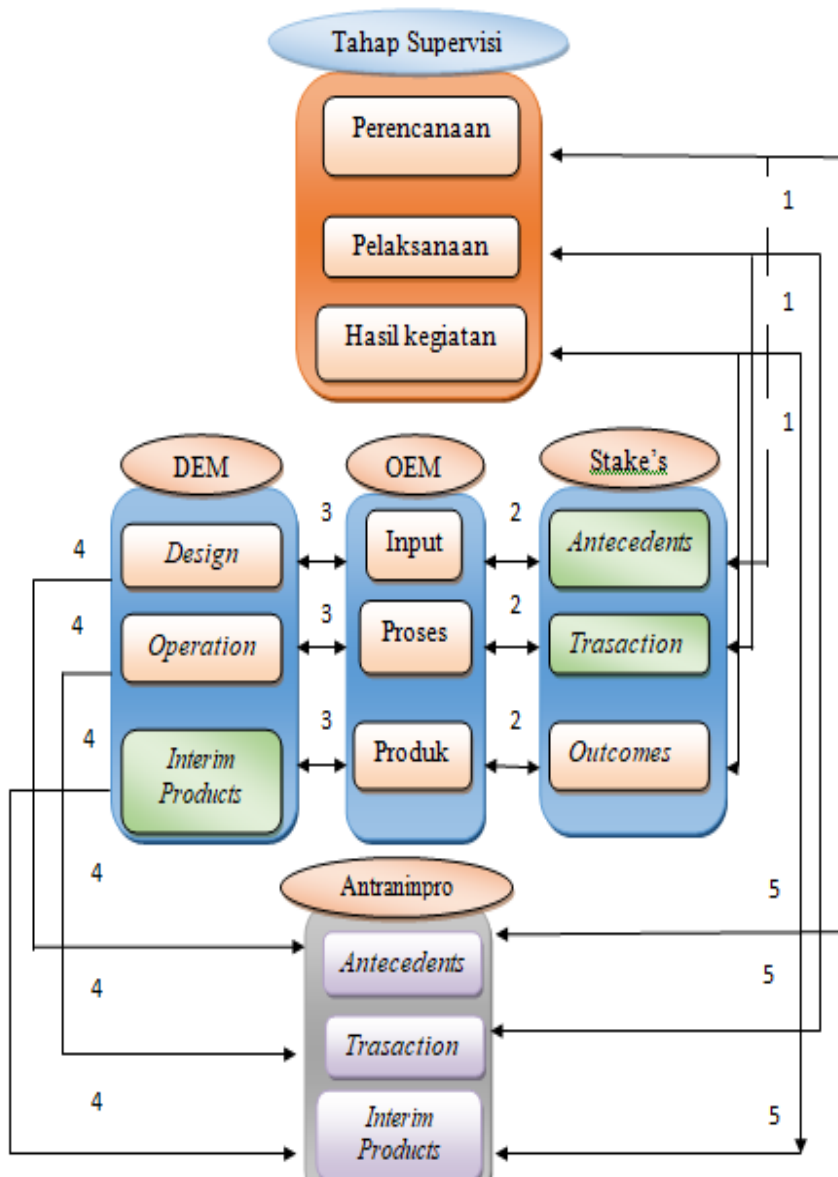
Model evaluasi yang memiliki substansi berdasarkan teori kesenjangan (antara standar dengan pengamatan, dan antara intensi dengan pelaksanaan) dari Stake's Model adalah *The Descrepancy Evaluation Model (DEM)*. Kedua model ini menitikberatkan teori kesenjangan, yakni sebagai alat untuk membuat pertimbangan (*judgement*) berdasarkan standar dan kinerja (Provus, 1969: 9,18). Perbedaannya pada Stake ada tambahan kegiatan *congruence* (kesesuaian antara *instens* dengan observasi), dan *contingencies* (hubungan logis *Antecedents, Transaction, Outcomes*). DEM ada proses memperbaiki, mendaur ulang atau menghentikan program.

Kombinasi Stake's model dengan DEM menghasilkan tahapan model *Antecedents, Transaction, Interim Product*, (Winaryati, 2013). Nama model evaluasi yang dikembangkan adalah Model Evaluasi Supervisi Pembelajaran berbasis 4 (empat) karakter skllil abad 21 (MESP 4Cs). Tahapan proses MESP 4Cs adalah sbb:



Gambar 25. Tahapan Proses Evaluasi Melalui MESp 4Cs

Penjelasan gambar di atas adalah: 1) judgment adalah kesenjangan/kesesuaian antara standar dengan observasi; 2) congruencies adalah kesenjangan atau /kesesuaian antara *intens* dengan *observation*; 3) contingencies adalah kesenjangan/hubungan antara *anteseden*, *transaction*, dan *interim product*. Hasil penilaian dari kegiatan *judgment*, *congruencies*, dan *contingencies*, dihasilkan informasi yang digunakan untuk perbaikan pembelajaran.



Gambar 26. Tahapan Pengembangan MESP 4Cs



- (1) Kualitas pembelajaran guru akan berdampak pada kualitas lulusan; kualitas supervise akan berdampak pada kualitas pembelajaran, dan evaluasi dilakukan untuk menilai sejauh mana proses supervise dilakukan. Pembelajaran di era abad 21 ini dituntut mengkreasi pembelajarannya dengan 4 (empat) karakter skill meliputi: *Critical Thinking and Problem Solving, Creativity and Innovation, Collaboration, dan Communication (4 C's)*. Evaluasi yang dilakukan diharapkan dapat memberikan informasi pelaksanaan supervise berbasis 4 C's ini.
- (2) Penelitian tentang supervisi banyak ditemukan, namun terkait supervise terhadap pelaksanaan 4 Cs belum ada. Rekomendasi hasil evaluasi akan berdampak pada perbaikan pembelajaran dikelas oleh guru, dan evaluasi terhadap penguasaan kepala sekolah sebagai supervisor terkait pelaksanaan 4 C's. Tulisan ini akan mengarahkan pada kerangka konsep penyusunan model evaluasi yang tepat, berdasarkan tahapan pada supervise pembelajaran berbasis 4 C's.
- (3) MESp 4C's dibangun berdasarkan dengan 5 (lima) tahap.
 - a) Pertama mendasarkan pada relevansi tahapan pembelajaran (perencanaan, proses pembelajaran dan penilaian) dengan fase supervise pembelajaran (pra-

- observasi, observasi, dan pasca observasi), serta tahapan evaluasi model Stake (*antecedent, transaction, outcomes*).
- b) Ke-dua, perlunya memodifikasi model evaluasi Stake dengan model evaluasi lainnya melalui Organizational Elements Model (**OEM**). OEM adalah suatu tahapan bagi evaluator menentukan model evaluasi yang tepat untuk digunakan. Diperoleh modifikasi model evaluasi yang sesuai yaitu *The Discrepancy Evaluation Models* (DEM). Fase Interim Product lebih tepat digunakan dibanding outcome, sehingga tahapan model evaluasinya menjadi *Antecedent, Transaction, dan Interim Product (ATIp)*.
 - c) Ke-tiga, setiap tahap supervisi keterlaksanaan 4C's, akan selalu dievaluasi.
 - d) Ke-empat adanya kegiatan feedback (self, peer, atasan) untuk menghasilkan kegiatan evaluasi-supervisi pembelajaran yang berkelanjutan, berdasarkan teori 360° feedback.
 - e) Ke-lima dihasilkan MESp 4C's sebagai model evaluasi dengan 6 cell Stake model yang dimodifikasi. MESp 4C's dilengkapi dengan data kesesuaian deskripsi antara instens dengan observasi, data penilaian antara observasi dengan standar, data *contingencies* dan *congruence*, serta data aktivitas feedback..

SARAN dan Batasan Penelitian

Rekomendasi penelitian adalah perlunya disusun desain MESp 4C's dan melakukan ujicoba pada beberapa sekolah. Keterbatasan penelitian ini adalah data hasil uji efektivitas model tidak dapat diperoleh dengan cepat; sebab membutuhkan waktu ujicoba yang lama dan melibatkan banyak sekolah.



DAFTAR PUSTAKA

- Ajasan, N. (2016). *Efektivitas pelaksanaan supervisi akademik oleh kepala sekolah dalam meningkatkan kinerja guru di SMK negeri 1 Meulaboh [The effectiveness of the implementation of academic supervision by the principal in improving the performance of teachers at Meulaboh State Vocational School]*. *Jurnal Administrasi Pendidikan Program Pascasarjana Unsyiah*, 4(3). <http://jurnal.unsyiah.ac.id/JAP/article/view/4795>
- Alkin, M. C. (1985). *A guide for evaluation decision makers*. Beverly Hills, CA: Sage.
- Alkin, M.C, & Christie, C.A. (2004). *An evaluation theory tree*.
- Allen, M.J & Yan, W.M. (1979). *Introduction to measurement theory*. California: Brooks/Cole Publishing Company.
- Aman. (2009). *Kajian model-model evaluasi program pendidikan*. Yogyakarta: Laporan penelitian pendidikan.
- Azwar, S. (2009). *Penyusunan Skala Psikologi*. Yogyakarta: Pustaka Pelajar
- Bacharach, S.B. (1989). *Organizational theories: some criteria for evaluation*. *Academy of Management Review* (14:4), pp. 496-515. Retrieved from https://gent.uab.cat/diego_prior/sites/gent.uab.cat.diego_prior/files/Bacharach_AMR_1989_14_4.pdf
- Bafadal, I. (1992). *Supervisi pengajaran: teori dan aplikasinya dalam membina profesional guru [Teaching supervision: theory*

and its application in fostering teacher professionalism].
Jakarta: Bumi Aksara

- Bahri, S. (2014). *Supervisi akademik dalam peningkatan profesionalisme guru*. *Visipena Jurnal*, 1(1), 100–112. DOI: 10.46244/visipena.v5i1.236
- Bahri, S. (2014). *Supervisi akademik dalam peningkatan profesionalisme guru [Academic supervision in improving teacher professionalism]*. *Visipena Jurnal*, 1(1), 100–112. <https://visipena.stkipgetsempena.ac.id/?journal=home&page=article&op=view&path%5B%5D=74>
- Bambling, M., & King, R. (2014). *Supervisor social skill and supervision outcome*. *Counselling and Psychotherapy Research*, 14(4), 256–262. <https://doi.org/10.1080/14733145.2013.835849>
- Bennet, J. (2006). *Evaluation methods in research*. New York: Continuum
- BNSP. (2006). *Standar Penilaian Badan Standar Nasional Pendidikan*. Jakarta.
- BNSP. (2007). *Standar Proses untuk Satuan Pendidikan Dasar dan Menengah*. Jakarta.
- BSNP. (2006). *Standar Isi untuk satuan pendidikan dasar dan menengah, standar kompetensi dan kompetensi dasar SMA/MA*. Jakarta.
- Charmaz, K. (2009). *Shifting the grounds: constructivist grounded theory methods for the*
- Chyung, y. (2005). *Improve the motivational appeal of online instruction for adult learners: what's in it for me?*. *Instructional & Performance Technology*. Boise State University.
- Chyung, Y. (2006). Donald Winiiecki., & Jo Ann Fenner. (2006). *Evaluation of effective interventions to solve the dropout problem in adult distance education*. *Instructional & performance technology*. *College of Engineering Boise State University*.

- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302. doi: 10.1037/h0040957
- Cronbach, L.J. (1982). Designing evaluation of educational and social program. San Fransisco: Jossey-Bass.
- Cyung, Y. (2005). Improve the motivational appeal of online instruction for adult learners: what's in it for me?. *Instructional & Performance Technology*. Boise State University.
- Daresh, J.C. (1989). *Supervision as a proactive process*. White Plains, NY: Longman. o from <https://www.amazon.com/Supervision-Proactive-Process-John-Daresh/dp/B003MAMGHA>
- Darto. (2014). Supervisi akademik terprogram dalam peningkatan kompetensi guru pada pelaksanaan pembelajaran [Programmed academic supervision in improving teacher competence in the implementation of learning]. *Jurnal Penelitian Tindakan Sekolah Dan Kepengawasan*, 1(1). <http://www.i-rpp.com/index.php/jptsk/article/view/75/74>
- Daud, Y., Dali, P. D., Khalid, R., & Fauzee, M. S. O. (2018). Teaching and learning supervision, teachers' attitude towards classroom supervision and students' participation. *International Journal of Instruction*, 11(4), 513-526. <https://doi.org/10.12973/iji.2018.11432a>
- Eisner, E. (1985). The art of educational evaluation: A personal view. Philadelphia, PA: The Falmer Press.
- Fernandes, H.J.X. (1984). *Evaluation educational of program*. Jakarta: National Education planning evaluation and curriculum development. Retrieved from https://www.researchgate.net/publication/338445328_Evaluation_of_Curriculum_Implementation_2013_Batubara_District
- Fetterman, D. (1984). *Ethnography in educational research: The dynamics of diffusion*. In D. M. Fetterman (Ed.),

Ethnography in educational evaluation (pp. 21-35).
Beverly Hills, CA: Sage.

- Fetterman, D. (1988). *A national ethnographic evaluation: An executive summary of the ethnographic component of the Career Intern Program Study*. In D. M. Fetterman (Ed.), *Qualitative approaches to evaluation in education: The silent scientific revolution* (pp. 262-273). New York: Praeger.
- Framework, A C. (2009). *A Conceptual framework and examples of country practices. Evaluation*, (December).
- Gioia, D.A., & Pitre, E. (1990). Multiparadigm perspectives on theory building. *Academy of Management Review* (15:4), pp. 584-602 <https://journals.aom.org/doi/abs/10.5465/amr.1990.4310758?journalCode=amr>
- Gioia, D.A., & Pitre, E. (1990). Multiparadigm perspectives on theory building. *Academy of Management Review* (15:4), pp. 584-602. <https://doi.org/10.2307/258683>
- Gregor, S. (2006). The Nature of theory in information systems, *MIS Quarterly* (30:3), pp. 611-642 <https://misq.org/skin/frontend/default/misq/pdf/TheoryReview/Gregor.pdf>
- Griffin, P., McGaw, B., & Care, E. (2012). Assessment and teaching of 21st century skills. (Vol. 9789400723245, pp. 1-345). *Springer Netherlands*. <https://doi.org/10.1007/978-94-007-2324-5>
- Hamdani, H., & Krismadinata, K. (2017). Kontribusi kompetensi dan persepsi tentang supervisi pembelajaran terhadap produktivitas kerja guru. *Invotek: Jurnal Inovasi Vokasional Dan Teknologi*, 17(2), 61-66. <https://doi.org/10.24036/invotek.v17i2.37>
- Hamm, R.W. (1985). A systematic evaluation of an environmental investigation course (*Doctoral dissertation. Georgia State University*) (ERIC Document.Reproduction Service No. ED-256-622).

- Hannum, K.M., & Kaufman, R. (2005). *Applying the organizational elements model to leadership development evaluation*.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*.
- Jaccard, J. & Jacoby, J. (2009). *Theory construction and model building skills*, (1. ed.). New York, NY, USA: The Guilford Press. Retrieved from <http://eprints.qums.ac.ir/788/1/Theory%20Construction%20and%20Model%20%20Building%20Skills.pdf>
- Journal of the Learning Sciences*, 11(1), 105-121.
- Kaufman, R. (1988). Preparing useful performance indicators. *Training & Development Journal*, 42(9), 80-84. Retrieved from <http://libproxy.boisestate.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=tfh&AN=9126902&site=ehost-live>
- Kaufman, R. (2006). Defining and delivering measurable value: a mega thinking and planning primer. *Performance Improvement Quarterly*, 18(3) pp. 6-16. Florida State University. DOI: 10.1111/j.1937-8327.2005.tb00337.
- Kaufman, R., & Thomas, S. (1980). *Evaluation without fear*. New Viewpoints. New York. London.
- Kemendikbud. (2017). *Panduan supervisi akademik*. Jakarta: Kemendikbud. Retrieved from http://repositori.kemdikbud.go.id/9042/1/KSP_M10.pdf
- Kereluik, K., Fahnoe, C., & Karr, J.A. (2013). What knowledge is of most worth: Teacher knowledge for 21 st century learning. *Journal of Digital Learning in Teacher Education*, 29 (4), 127-140
- Khoso, P.A. (2019). What are the differences between conceptual framework and theoretical framework?. *Universiti Kebangsaan Malaysia. School of Social, Development & Environmental*. SOSDES. Retrieved from https://www.researchgate.net/post/What_are_the_differences_between_conceptual_framework_and_theoretical_framework/2

- Kilminster, S., et al. (2007). AMEE Guide No. 27: Effective educational and clinical supervision. *Medical Teacher* 29 (1):2-19. doi: 10.1080/01421590701210907.
- Kirkpatrick, D. L. (2009). *Kirkpatrick's training evaluation model*. San Francisco: Berrett-Koehler Publisher, Inc.
- Kirkpatrick, D. L. (1998). *Evaluating training programs: The Four Levels*. San Francisco: Berrett-Koehler Publisher, Inc.
- Kirkpatrick, D. L. 1998: 21. *Evaluating training programs: The Four Levels*.
- Kosecoff, J. & Fink, A. (1985). *Evaluation basics, a1 practitioner's manual*. California: Sage Publication, Inc1
- Levinger, G., & Rubin, J. (1994). Bridges and barriers to a more general theory of conflict. *Negotiation Journal*, 10(3), 20-35. Levinger, G., & Rubin, J. (1994). Bridges and barriers to a more general theory of conflict. *Negotiation Journal*, 10(3), 20-35.
- Madaus, G. F., Scriven, M. S. & Stufflebeam. (1983). The Cipp model for program evaluation. *Evaluation Models* (pp. 117-141). Springer Netherlands. https://doi.org/10.1007/978-94-009-6675-8_7
- Maralih. (2014). Peranan supervisi dalam peningkatan kualitas pendidikan [The role of supervision in improving the quality of education]. *Jurnal Qathruna*, 1(1), 179-192. Retrieved from <http://jurnal.uinbanten.ac.id/index.php/qathruna/article/view/251>
- Mardapi, D. (2008). *Teknik penyusunan instrumen tes dan non tes*. Jogjakarta: Mitra Cendekia.
- Martínez, J. F., Schweig, J., & Goldschmidt, P. (2016). Approaches for combining multiple measures of teacher performance. *Educational Evaluation and Policy Analysis*, 38(4), 738-756. Retrieved from <https://doi.org/10.3102/0162373716666166>

- Maxwell, J A. (1996). *Qualitative research design, an interactive a halfroach*. California: Sage Publications, Inc
- Maxwell, J A. (1996). *Qualitative research design, an interactive a halroach*. California: Sage Publications, Inc. Retrieved from [https://www.scirp.org/\(S\(351jmbntvnsjt1aadkposzje\)\)/reference/ReferencesPapers.aspx?ReferenceID=1424076](https://www.scirp.org/(S(351jmbntvnsjt1aadkposzje))/reference/ReferencesPapers.aspx?ReferenceID=1424076)
- McKenney, S., Nieveen, N., & van den Akker, J. (2002). Computer Support for curriculum developers CASCADE. *ETR&D*, Vol. 50, No. 4, pp. 25–35 ISSN 1042–1629
- Mullerl, B & Urbach, N. (2017). *Understanding the why, what, and how of theories in is research*. Research Center. Finance & Information management: University of Groningen. Retrieved from <https://pdfs.semanticscholar.org/780c/0d7e229e0ae37f213185c639d621ac4d9620.pdf>
- Nickols, F. (2010). The organizational elements model. Diunduh pada 2 Oktober 2011, dari <http://www.nickols.us/organi-zationalelements.pdf>
- Nunnally & Bernstein, I.H. (1994). *Psychometric theory* (3nd ed). New York: McGraw Hill. Retrieved from <https://journals.sagepub.com/doi/abs/10.1177/014662169501900308>
- Nunnally & Bernstein, I.H. (1994). *Psychometric theory*. third edition. New York: McGraw Hill.
- OECD (2005). *Teacher matter: Attracting, developing and reraining efective teacher*. Paris: OECD. Retrieved from <https://www.oecd.org/education/school/34990905.pdf>
- OECD (2009). A Conceptual framework and examples of country practices. *Evaluation*, (December). Retrieved from <http://www.oecd.org/education/school/44568106.pdf>
- Owen, J.M. (2004). Evaluation Form. Toward an inclusive framework for evaluation practice. In M.C, Alkin (Ed). *Evaluation rootd: Traicing theorits' views and influences* (pp. 356-369). Thousand Oaks, CA: Sage

- Owen, J.M., & Rogers, P.J. (1999). *Program evaluation, forms and approaches*-International Edition, SAGE Publications, London.
- Owston, R. (2008). *Models and methods for evaluation*. Toronto Canada: York University.
- Owston, R. (2008). *Models and methods for evaluation*. Toronto Canada: York University. [http://daneshnamehicsa.ir/userfiles/file/Resources/7\)%20Evaluation/Models%20and%20Methods%20for%20Evaluation.pdf](http://daneshnamehicsa.ir/userfiles/file/Resources/7)%20Evaluation/Models%20and%20Methods%20for%20Evaluation.pdf)
- Partnership*. (2009). P21 framework definitions. Retrieved from <https://files.eric.ed.gov/fulltext/ED519462.pdf>
- Partnership. (2015). P21 Partnership for 21st century learning. *Partnership for 21st Century Learning*, 9. Retrieved from http://www.p21.org/documents/P21_Framework_Definitions.pdf
- Partnership. (2015). P21 Partnership for 21st century learning. *Partnership for 21st Century Learning*, 9. Retrieved from http://www.p21.org/documents/P21_Framework_Definitions.pdf
- Peirce, C.S. (1998). On the Logic of drawing history from ancient documents. *The Essential Peirce* (Vol. 2), C.S. Peirce (ed.). Bloomington, IL, USA: Indiana University Press, pp. 75-11
- Popham, W J. (1993). *Educational evaluation*. Boston: Allyn and Bacon.
- Popham, W J. (1993). *Educational evaluation*. Boston: Allyn and Bacon. Retrieved from <https://www.worldcat.org/title/educational-evaluation/oclc/25675424>
- Porter, C. (2005a). Week 1: Job technology: hard & soft. Retrieved August 19, from IPT 536 4158/4159 course database.
- Prabowo, S., & Yoga, D. S. (2016). Supervisi kunjungan kelas sebagai upaya membina profesional guru SLTP/SLTA. *Jurnal Sosial Humaniora*, 9(1), 96. <https://doi.org/10.12962/j24433527.v9i1.1281>

- Provus, M. (1969). *The discrepancy evaluation models. An approach to local program improvement and development*. Pittaburgh Public School.
- Provus, M. (1971). *Discrepancy evaluation for educational program improvement and assessment*. Berkeley, CA: McCutchan Publishing Corporation.
- Reinhorn, S. K., Johnson, S. M., & Simon, N. S. (2017). Investing in development: Six high-performing, high-poverty schools implement the massachusetts teacher evaluation policy. *Educational Evaluation and Policy Analysis*, 39(3), 383–406. <https://doi.org/10.3102/0162373717690605>
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: a systematic approach*. New Delhi: Saga Publications.
- Sabandi, A. (2013). Supervisi pendidikan untuk pengembangan profesionalitas guru berkelanjutan. *Jurnal Ilmiah Ilmu Pendidikan*, XIII(2), 1–9.
- Sabandi, A. (2013). Supervisi pendidikan untuk pengembangan profesionalitas guru berkelanjutan [Educational supervision for continuous teacher professional development]. *Jurnal Ilmiah Ilmu Pendidikan [Scientific Journal of Educational Sciences]*. XIII(2), 1–9. Retrieved from <http://ejournal.unp.ac.id/index.php/pedagogi/article/view/4275>
- Schuh, G., at.al. (2015). Promoting work-based learning through industry 4.0. In *Procedia CIRP* (Vol. 32, pp. 82–87). Elsevier B.V. <https://doi.org/10.1016/j.procir.2015.02.21>
- Sciven, M. (1967). *The methodology of evaluation*. In R. W. Tyler, R.M.
- Scott, C. L. (2015). The futures of learning 3: what kind of pedagogies for the 21st century?. *Education Research and Foresight* (pp. 1–21). Retrieved from <https://unesdoc.unesco.org/ark:/48223/pf0000243126>

- Searcy, D & J.T. Mentzer. (2003). A Framework for Conducting and Evaluating Research. *Journal of Accounting Literature*. 22 (130-167) <https://search.proquest.com/openview/57795fd99db05acb093900bce161c56c/1?pq-origsite=gscholar&cbl=31366>
- Searcy, D & J.T. Mentzer. (2003). A Framework for conducting and Evaluating Research. *Journal of Accounting Literature*. 22 (07374607) (130-167). Retrieved from <https://www.mendeley.com/catalogue/5198beaf-abec-366f-994f-01c03c9f02c5/>
- Selvi, K. (2010). Teachers' competencies. *Cultura. International Journal of Philosophy of Culture and Axiology. Philosophy Documentation Center*. <https://doi.org/10.5840/cultura20107133>
- Sergiovanni, T. J., & Starratt, R. J. (1998). *Supervision: A re-definition* (6th ed.). Boston: McGraw-Hill. <https://www.amazon.com/Supervision-Redefinition-Thomas-Sergiovanni/dp/0073131261>
- Sergiovanni, T.J. (1982). *Supervision of teaching*. Alexandria: ASCD.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. <https://doi.org/10.3102/0034654307313795>
- Stake, R. E. Program evaluation, particularly responsive evaluation. Sumber: www.fivehokies.com/evaluation/evaluation%... Diakses 5 September 2012.
- Stake, RE. (1967). The countenance of educational evaluation. *Teacher's Coolege Record*. Vol. 68, no:7.
- Stake, RE. (1977). The countenance of educational evaluation. In A.A.Ballack & H/M. Kliebard Eds). *Curriculum and evaluation* (pp 372-390). Berkeley CA: McCutchan. Retrieved from <https://dx.doi.org/10.4135/9781412984157.n13>
- Stake, RE. (1977). *The countenance of educational evaluation*. In A.A.Ballack & H/M. Kliebard Eds). *Curriculum and evaluation* (pp 372-390). Berkeley.CA McCutchan.

- Stake, RE. (2000). *Program evaluation, particularly responsive evaluation*. Dalam *Evaluation models viewpoints on educational and human services evaluation* 2th Ed (pp 340-362). EwYork: Kluwer Academic Publisher.
- Stake, RE. (2006). *The countenance of educational evaluation*. Center for Instructional Research and Curriculum Evaluation. University of Illinois.
- Steinmetz, A. (2000). *The discrepancy evaluation model*. *Evaluation Models Viewpoints on Educational and Human Services Evaluation* Second Edition. New York: Kluwer Academic Publishers.
- Stufflebeam & Shinkfield. (1985). *Systematic evaluations*. Boston: Kluwar Academic Publishers.
- Stufflebeam, D. L, Madaus, G. F., &Kellaghan, T. 2000: 280. *Evaluation models: viewpoints of educational and human services evaluation..*
- Stufflebeam, D. (2001). *Evaluation models*. *New Directions for Evaluation*, 85, 8-98
- Stufflebeam, D. L. (1983). The CIPP model for program evaluation. In G. F. Madaus, M. S. Scriven, & D. L. Stufflebeam (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 117-141). Boston: KluwerNijhoff.
- Stufflebeam, D.L. (1966). *Evaluation under Title I of the element Act of 1967*. Addt=ress delivered at the Title I Evaluation Conference sp Departement of Education, Lansing.
- Stufflebeam, D.L. (1969). *Evaluation under title I of the elemen act of 1967*. Address delivered at the Title I Evaluation Conference sp Departement of Education, Lansing. Retrieved from https://books.google.co.id/books/about/The_Discrepancy_Evaluation_Model.html?id=MdzYGwAA CAAJ&redir_esc=y
- Stufflebeam, D.L. (1973). *Educational evaluation: theory and practice*. *Evaluation as enlightenment for decision-making*. In B. R. Worthen & J. R. Sanders (Eds.),

- Suddaby, R. (2010). Construct clarity in theories of management and organization. *Academy of Management Review* (35:3), pp. 346-357.
- Tang, S. Y. F., & Chow, A. W. K. (2007). Communicating feedback in teaching practice supervision in a learning-oriented field experience assessment framework. *Teaching and Teacher Education*, 23(7), 1066–1085. doi:10.1016/j.tate.2006.07.013
- Tyler, R. W. (1942). General Statement on Evaluation. *The Journal of Educational Research*, 35(7), 492-501.
- Weiss, C. (1983). The stakeholder approach to evaluation: Origins and promise. *New Directions in Program Evaluation*, 17, 3-14
- Weiss, C. H. (1972). Utilization of evaluation: toward comparative study." In C. H. Weiss (ed.), *Evaluating Action Programs: Reading in Social Action and Education*. Boston: Allyn & Bacon.
- Wheeler, K. (2019). *What is the difference between a theory and a concept?*. Faculty sponsor for Alpha Chi Honor Society, Tennessee Delta Chapter.
- Wheeler, K. (2019). *What is the difference between a theory and a concept?*. Faculty sponsor for Alpha Chi Honor Society, Tennessee Delta Chapter. Retrieved from <https://www.quora.com/What-is-the-difference-between-a-theory-and-a-concept>
- Whetten, D.A. (1989). What Constitutes a Theoretical Contribution?. *Academy of Management Review* (14:4), pp. 490-495.
<http://cmsdev.aom.org/uploadedFiles/Publications/AMR/WhettenWhatconstitutes.pdf>
- Whetten, D.A. (1989). What Constitutes a Theoretical Contribution?. *Academy of Management Review* (14:4), pp. 490-495. Retrieved from <https://doi.org/10.5465/amr.1989.4308371>

- Winaryati, E. (2012). Kompetensi kepala sekolah sebagai supervisor pada pembelajaran IPA SMP di kota Semarang [Principal's competence as a supervisor in junior high school science learning in the city of Semarang]. *Prosiding Univ. Muhammadiyah Semarang*. Univ Muhammadiyah Semarang. ISSN: 978-602-18809-0-6 <http://jurnal.unimus.ac.id/index.php/psn...>
- Winaryati, E., & Mufnaety. (2012). Kompetensi kepala sekolah sebagai supervisor pada pembelajaran IPA SMP di kota Semarang [The competence of the principal as a supervisor in junior high school science learning in the city of Semarang]. *Prosiding Univ. Muhammadiyah Semarang (UNIMUS)*. ISSN: 978-602-18809-0-6. Retrieved from <http://jurnal.unimus.ac.id/index.php/psn...>
- Winaryati, E., Mardiana & Hidayat, M.T. (2020). Conceptual framework of evaluation model on 4 c's-based learning supervision. *International Journal of Learning, Teaching and Educational Research*. 19(8), pp. 173-193. <https://doi.org/10.26803/ijlter.19.8.10>
- Winaryati, E., Suyata & Sumarno (2013). Model Evaluasi dalam Supervisi Pembelajaran IPA Berbasis Lima Domain Sains. *jurnal Penelitian dan Evaluasi Pendidikan*. Vol 17, No 2 (2013)
- Wolters, C. (2010). Self-regulated learning and the 21 st century competencies. *USA University of Houston*, (May), 27. Retrieved from http://www.hewlett.org/uploads/Self_Regulated_Learning_21st_Century_Compencies.pdf
- Wood, B.B. (2001). Stake's countenance model: evaluating an environmental education professional development course. *The Journal of Environmental Education*, Vol. 32 No. 2 18-27. 10.1080/00958960109599134
- Youngstrom, M. J., & Gentile, P. A. (2018). Supervision. In *Willard and Spackmans Occupational Therapy, 13th Edition* (pp.

1134–1155). Wolters Kluwer Health. <https://doi.org/10.12968/hmed.2009.70.4.41628>

Youngstrom, M. J., & Gentile, P. A. (2018). *Supervision*. In *willard and spackmans occupational therapy, 13th Edition* (pp. 1134–1155). Wolters Kluwer Health. <https://doi.org/10.12968/hmed.2009.70.4.41628>

Yunus, G. A., Lestari, W., & Raharjo, T. J. (2016). Pengembangan Model Supervisi Akademik Berbasis Evaluasi Diri Bagi Guru SMA. *Educational Management*, 5(1), 12–22. Retrieved from <http://journal.unnes.ac.id/sju/index.php/eduman>

Yunus, G. A., Lestari, W., & Raharjo, T. J. (2016). Pengembangan Model Supervisi Akademik Berbasis Evaluasi Diri Bagi Guru SMA [Development of Self-Evaluation Based Academic Supervision Model for High School Teachers]. *Educational Management*, 5(1), 12–22. Retrieved from <http://journal.unnes.ac.id/sju/index.php/eduman>

Zhou, Z. H. (2018). A brief introduction to weakly supervised learning. *National Science Review*. Oxford University Press. <https://doi.org/10.1093/nsr/nwx106>

www.penerbitbukumurah.com

BIODATA PENULIS



Dr. Eny Winaryati, M.Pd. Lahir di Rembang Jawa Tengah, pada 25 Desember 1963. Lulus S1 FKIP Pendidikan Kimia UNS (1989), S2 Pendidikan Sains, konsentrasi Pendidikan Kimia UNNES (2009), S3 Penelitian dan Evaluasi Pendidikan UNY (2013). Saat ini menjadi dosen Pendidikan Kimia, FMIPA, Universitas Muhammadiyah Semarang (UNIMUS). Mempunyai hobi menulis puisi

yaitu: *Perjalanan* (2004), *Manakala Tangan Menengadah (Berlaksa Harap Menjuntai Rahmat)* ditulis tahun 2005, *Dua Perempuan* (2006), *Merajut Mimpi* (2007), *Empat puluh tahun dalam empat puluh lima Puisi untuk 'AISYIYAH* (Kenangan di bulan Desember 2008), *Memori yang Tercecer diantara gedung PPS* (2009), *Tanpa Kata Ada Rasa* (2010), *Percikan, Goresan dan Tulisan* (2011), *Menggapai Hidayah* (2012), *Lambaian Kenang (Benamkan Asa, Terbangkan Cita* (2013), *Babak Baru* (Arus Skenario Illahi, 2014). *Lentera Hati* (2015), *Kado Untuk Anakku* (2016), *POTRET* (2017). *Buku Ilmiah yang ditulis: Pendidikan Karakter Bagi Mahasiswa* (2014), *Evaluasi Supervi Pembelajaran* (2015), *Model Pembelajaran "Wisata Lokal"*, (2017), *Action Research dalam Pendidikan* (Teori dan Aplikasinya, 2018). *Cercular Model of RD&D* (Model Digunakan Pada Pengembangan Platfork Digital MESp 4Cs) (2021).

Alamat email: enywinaryati@unimus.ac.id. Telp: 081325678400



Muhammad Munsarif, S.Kom., M.Kom. Lahir di Demak Jawa Tengah, pada 14 Maret 1972. Lulus D3 Informatika UDINUS (1998). S1 Sistem informasi UDINUS (2002), Lulus S2 Teknik Informatika UDINUS (2004). Saat ini menjadi dosen Informatika, Fakultas Teknik, Universitas Muhammadiyah Semarang (UNIMUS). Mempunyai hobi Olahraga, Bisnis, dan membaca. Buku yang sudah di tulis : *Kewirausahaan dan Bisnis Online* (2020), *Media pembelajaran* (2020), *Tourism Marketing* (2020), *Teori-Teori Manajemen Sumber Daya Manusia* (2020), *Sistem Basis Data Lanjutan* (2021). *Cercular Model of RD&D (Model Digunakan Pada Pengembangan Platfork Digital MESp 4Cs)* (2021). Email: m.munsarif@unimus.ac.id. Telp: 081238931088.



Dr. Mardiana lahir di metro Lampung tanggal 14 Januari 1974. Menempuh pendidikan di SD Negeri 8 Metro (1985), SMP Negeri 4 Metro (1989), MAN 1 Metro (1991), IAIN Raden Intan Lampung S1 (1996). Pendidikan S2 di PPs IAIN Raden Intan Lampung, Prodi Pengembangan Kurikulum (2009), dan pendidikan Doktor ditempuh di PPs Universitas Negeri Yogyakarta, Prodi Penelitian dan Evaluasi Pendidikan (2017). *Cercular Model of RD&D (Model Digunakan Pada Pengembangan Platfork Digital MESp 4Cs)* (2021). Hingga saat ini bertugas di Universitas Muhammadiyah Lampung. Email mardhiyana.rahma@yahoo.com. Telp. 081229949101.



Dr. SUWAHONO, M.Pd. lahir pada 20 Mei 1975 di Kota Semarang, Jawa Tengah. Berasal dari keluarga sederhana. Penulis mengenyam pendidikan formal di SDN Jrahak II, lulus tahun 1985. Selanjutnya meneruskan sekolah di SMP Negeri 16 Semarang, tamat tahun 1989, dan di SMA Negeri 6 Semarang, tamat tahun 1991. 2 tahun 1991 sampai tahun 1993 bekerja sebagai

buruh kasar untuk mengumpulkan uang buat biaya kuliah. Di tahun 1993 penulis Kuliah di Jurusan Pendidikan Kimia FPMIPA IKIP Negeri Semarang. Pada tahun 2009, penulis meraih gelar Magister Pendidikan (M.Pd.). Dan Penulis Menyelesaikan studi Doktorat di Universitas Negeri Yogyakarta pada Tahun 2017. Di luar kesibukannya sebagai Dosen di program studi pendidikan kimia Fakultas Sains dan Teknologi, penulis juga mengajar di Universitas terbuka dan aktif mengedukasi tentang statistik dan metodologi penelitian bidang pendidikan. buku-buku yang sudah diterbitkan antara lain: (1) *Pemodelan Rasch Sebagai Kerangka Acuan Penilaian* (2018). (2) *Pengukuran Keterampilan Generik Kimia*, (2018). (3) *Asesmen Keterampilan Dasar Kimia*, (2020). (4) *Principia Kimia : Sejarah Dan Filosofi Kimia* (2021). *Circular Model of RD&D (Model Digunakan Pada Pengembangan Platform Digital MESp 4Cs)* (2021).

Penulis ini dapat dihubungi pada alamat Rumah jalan Segaran III no 19 Ngaliyan Semarang 50185 dengan Nomor HP 08122925061, dengan alamat Surel Suwahono@walisongo.ac.id



