

EVALUASI PROSES PEMBELAJARAN

Sunardi
Jurusan PLB FIP UPI

A. Pendahuluan

Evaluasi pembelajaran adalah sistem. Artinya suatu rangkaian kegiatan yang melibatkan berbagai unsur sebagai satu kesatuan. Masing-masing unsur mempunyai fungsi dan peran tersendiri dan perubahan dalam salah satu unsur akan berpengaruh pada unsur yang lainnya.

Dalam dunia pendidikan, evaluasi merupakan suatu kegiatan yang tak terpisahkan dan sama pentingnya dengan proses pembelajaran. Pembelajaran tanpa kegiatan evaluasi akan kehilangan makna. Sebab guru tidak akan memperoleh informasi penting tentang tingkat pencapaian tujuan, tingkat penguasaan materi belajar, kekuatan-kelemahan siswa dalam belajar, serta kekuatan-kelemahan guru dalam proses pembelajaran yang dikembangkan.

Walaupun evaluasi dianggap penting dan sudah merupakan pekerjaan rutin guru, namun dalam kenyataan sehari-hari di lapangan sistem evaluasi dalam pembelajaran bukan berarti tanpa persoalan. Berdasar pengamatan sepintas di lapangan, beberapa persoalan tersebut paling tidak berkaitan dengan pemahaman konsep dasar evaluasi, pelaksanaan dan pemanfaatannya, serta evaluasi program pengajaran. Hal-hal ini yang selanjutnya akan dijadikan kajian pokok dalam tulisan ini. Tulisan ini juga perlu perdebatan dan kajian lebih lanjut khususnya dikaitkan dengan persamaan dan perbedaan dengan pembelajaran dan atau siswa yang lainnya.

B. Konsep Dasar Evaluasi

Evaluasi bukan monopoli dunia persekolahan atau milik guru saja, sebab faktanya setiap saat setiap manusia tidak lepas dari pekerjaan evaluasi melalui kegiatan **pengukuran** (*measurement*) dan **penilaian** (*evaluation*).

Misal :

Ketika seseorang ditawarkan untuk memilih antara dua cincin emas yang tidak sama beratnya, tentu kita tidak memilih yang lebih berat atau gram-nya lebih banyak kecuali ada alasan tertentu, atau ketika kita berbelanja di pasar membeli jeruk, maka kita akan memilih jeruk yang lebih baik menurut ukuran kita. Misalnya yang besar, kuning, dan kulitnya halus karena berdasar pengalaman jeruk demikian yang rasanya manis.

Dari contoh di atas, sebelum kita menentukan pilihan kita mengadakan penilaian terlebih dahulu, dan sebelum melakukan penilaian terlebih dahulu melakukan pengukuran. Kita tidak dapat melakukan penilaian sebelum melakukan pengukuran. Dalam contoh pertama sebelum kita memilih, kita menilai dulu mana cincin yang lebih berat, dan untuk mengetahui mana yang lebih berat perlu melakukan pengukuran dahulu, dengan menimbanginya sehingga diketahui gram-nya. Dalam contoh kedua sebelum kita menentukan pilihan, kita menilai dahulu mana jeruk yang kira-kira rasanya

manis dengan “mengukur” berdasar ukuran besar, kuning, dan halus kulitnya. Jeruk mana yang memenuhi ukuran itu yang akan kita pilih.

Dengan demikian ukuran ada yang terstandar (meter, kilogram, dsb) dan yang tidak terstandar (depa, jengkal, dsb. serta ukuran perkiraan berdasar hasil pengalaman) dan keduanya dapat dan biasa digunakan untuk dasar penilaian.

- Mengukur adalah membandingkan sesuatu dengan satu ukuran tertentu. Pengukuran bersifat kuantitatif.
- Penilaian adalah pengambilan suatu keputusan terhadap sesuatu dengan ukuran baik buruk. Penilaian bersifat kualitatif.
- Evaluasi meliputi keduanya, yaitu mengukur dan menilai.

S. Hamid Hasan (1996) menegaskan bahwa dalam evaluasi pendidikan masih dihadapkan pada berbagai persoalan, terutama berkaitan dengan: (1) Pemahaman mengenai istilah evaluasi, (2) peran yang dapat dilakukan evaluasi, (3) pemanfaatan hasil evaluasi, dan (4) kebijakan tentang evaluasi.

Dalam kaitannya dengan pemahaman mengenai istilah evaluasi, dijelaskan bahwa sampai saat ini di samping masih muncul keragaman tentang definisi juga masih terjadi penyalahgunaan istilah evaluasi. Evaluasi cenderung diartikan orang berdasarkan ruang lingkup tugas yang akan dilakukan dalam kegiatan evaluasi tersebut. Evaluasi untuk hasil belajar diartikan agak berbeda dengan evaluasi kurikulum, evaluasi program studi, dsb. Akibatnya pengertian evaluasi sangat tergantung pada konteksnya. Disamping itu, dalam dunia pendidikan kita sehari-hari masih sering terjadi penyalahgunaan atau kesalahan pemakaian istilah evaluasi, dengan istilah lain seperti **ulangan dan tes**. Ulangan, tes, dan evaluasi memang memiliki kesamaan, tetapi juga perbedaan-perbedaan, bahkan perbedaannya cukup prinsipil sehingga ada yang mengatakan bahwa evaluasi dapat dilakukan tanpa adanya tes ataupun ulangan. Tes dan ulangan adalah **alat** pengumpul informasi tentang hasil belajar, sedang informasi hasil belajar tersebut dapat pula dilakukan dengan cara-cara yang lain. Tes dan ulangan juga belum merupakan evaluasi, sebab evaluasi **harus sampai pada mengolah** informasi yang dikumpulkan dan mengambil kesimpulan-kesimpulan mengenai berbagai hal yang berkaitan dengan pendidikan.

Sekalipun ada keragaman definisi, namun secara mendasar terdapat tiga karakteristik penting dalam suatu kegiatan evaluasi, yaitu:

1. Kriteria atau standar

Kriteria adalah **tolok ukur** yang digunakan untuk mengetahui/mengukur aspek-aspek atau apa yang telah terjadi dalam kegiatan pendidikan. Kriteria menentukan jenis data yang akan dikumpulkan.

Misal:

- Untuk melakukan evaluasi mengenai CBSA, maka terlebih dahulu harus dikembangkan aspek atau ciri yang dapat menggambarkan CBSA. Aspek-aspek ini adalah kriteria yang akan menentukan informasi yang diperlukan untuk evaluasi.

- Untuk mengevaluasi hasil belajar siswa melalui keterampilan proses, maka kriterianya adalah bagaimana kemampuan siswa yang sudah dikuasai, mencakup: pengetahuannya, kemampuan merumuskan masalah, menentukan data, sumber data, prosedur pengumpulan, pengolahan, data, penafsiran, dan generalisasi.
- Untuk evaluasi hasil belajar, kriteria dapat dikembangkan dari tujuan yang ingin dicapai kurikulum ke dalam aspek-aspek yang dianggap penting secara rinci. Dalam konteks tujuan instruksional, kriteria dapat digali dari aspek-aspek yang terumuskan dalam tujuan instruksional khusus (TIK).

Standar adalah **kualitas** yang seharusnya dicapai oleh kegiatan pembelajaran pada suatu atau sejumlah kriteria yang digunakan atau merujuk pada sesuatu yang harus (should). Standar menuntut identifikasi yang jelas tentang apa yang seharusnya diketahui atau dapat dilakukan oleh siswa. Standar juga memiliki arti yang lebih luas dari pada kriteria, karena ia mencakup berbagai kriteria yang digunakan. Standar untuk evaluasi terbatas (satu pertemuan, misalnya) berbeda dengan standar yang digunakan untuk evaluasi hasil belajar setelah satu cawu, satu tahun, atau akhir masa pendidikan, karena perbedaan kriteria yang digunakan.

Perbedaan dalam standar sangat mungkin terjadi antara guru yang satu dengan guru yang lain, atau antara sekolah yang satu dengan sekolah yang lain, karena adanya perbedaan pandangan mengenai kualitas belajar. Karena itu angka 6 dari guru atau sekolah A dapat berbeda dengan angka 6 dari guru atau sekolah B. EBTANAS dapat dikatakan menggunakan standar yang sama secara nasional, tetapi tidak ada standar kelulusan yang sifatnya nasional. Keputusan evaluasi hasil belajar ditentukan oleh standar yang digunakan.

2. Tujuan dan fungsi evaluasi

Kegiatan evaluasi selalu memiliki tujuan dan tujuan apa yang ingin dicapai membawa pengaruh terhadap informasi yang akan dikumpulkan. Dalam dunia persekolahan, paling tidak ada dua tujuan besar, yaitu untuk mengetahui:

a. *Tingkat keberhasilan siswa dalam mempelajari sesuatu, sehingga diketahui kedudukannya.*

Evaluasi ini memiliki **fungsi sumatif**. Pengertian sumatif berlaku baik untuk evaluasi proses maupun hasil belajar. Misal nilai 6 dalam rapor mencerminkan status pencapaian tujuan yang minimal dapat diterima atau standar minimal yang disebut lulus.

b. *Kekuatan dan kelemahan yang dimiliki siswa atau suatu proses pembelajaran.*

Aspek mana yang sudah dikuasai dan yang belum agar guru dapat memberikan bantuan pada siswa secara tepat, mungkin secara individual atau kelompok. Tujuan ini berkaitan dengan **fungsi formatif**. Secara teknis guru mungkin tidak harus membuat instrumen dan melakukan pengolahan data secara khusus, tetapi dilakukan dalam konteks ulangan dan dalam pengolahan tidak sekedar menskor tetapi yang lebih penting adalah melakukan identifikasi pada siswa yang belum menguasai serta soal-soal yang dianggap sukar, baik secara individual atau kelompok, sehingga diketahui kelemahannya dan ini dijadikan dasar untuk memberikan bantuan atau remedial. Hal ini penting, agar lebih berkembang sehingga dapat menunjang penguasaan materi berikutnya.

3. Keputusan yang akan diambil.

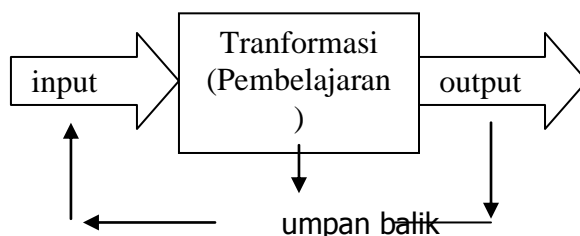
Karakteristik ketiga dari evaluasi adalah jenis keputusan yang akan diambil dan ini erat kaitannya dengan tujuan dan fungsi evaluasi. Evaluasi dengan tujuan dan fungsi formatif akan memberikan warna keputusan yang berbeda dari tujuan dan fungsi sumatif. Misal untuk tujuan dan **fungsi formatif** diharapkan guru dapat mengambil keputusan mengenai **kelemahan dan kekuatan dari siswa** dalam belajar, proses belajar yang dialami, proses pembelajaran yang dilakukan guru, sumber belajar yang tersedia, maupun alat evaluasi yang digunakan. Sedangkan untuk tujuan dan **fungsi sumatif**, keputusan yang diambil lebih banyak berkaitan dengan **kenaikan kelas atau kelulusan**. Keputusan apa yang diambil hendaknya tidak lepas dari pertimbangan aspek psikologis, sosial, maupun pedagogis.

C. Evaluasi Pembelajaran

1. Konsepsi Pembelajaran

Dalam pembelajaran di sekolah atau khususnya di kelas, guru adalah pihak yang paling bertanggung jawab atas hasilnya, sehingga guru perlu dibekali dengan evaluasi sebagai ilmu yang mendukung tugasnya, yakni mengevaluasi hasil belajar siswa. Dalam hal ini guru bertugas mengukur apakah atas bimbingan guru siswa sudah menguasai ilmu yang dipelajari sesuai dengan tujuan yang dirumuskan.

Dalam konsep lama, prestasi belajar adalah hasil dari kegiatan belajar mengajar semata. Artinya proses belajar mengajar adalah satu-satunya factor determinan bagi hasil belajar. Pendapat ini sudah tidak berlaku lagi. Proses belajar mengajar bukanlah satu-satunya factor yang menentukan prestasi belajar siswa, karena prestasi merupakan merupakan hasil kerja (ibarat mesin) yang keadaanya sangat kompleks. Apabila sekolah diumpamakan sebagai tempat mengolah sesuatu dan calon siswa sebagai bahan mentahnya (input), maka lulusan dari sekolah itu dapat disamakan dengan hasil olahannya (output). Dalam istilah inovasi yang menggunakan teknologi, tempat pengolah itu disebut transformasi. Dalam konteks pembelajaran di sekolah, transformasi adalah seluruh factor yang menentukan kegiatan pembelajaran, yang meliputi: (a) siswa itu sendiri, (b) guru dan personal lainnya, (c) bahan pelajaran, (d) metode mengajar dan sistem evaluasi, (e) sarana penunjang, dan (f) sistem administrasi. Dalam prosesnya terdapat umpan balik yang diperlukan untuk memperbaiki input maupun transformasi. Jika digambarkan tampak sebagai berikut:



2. Makna Penilaian

Dalam pembelajaran, penilaian mempunyai beberapa **makna**, baik bagi siswa, guru, maupun sekolah.

Bagi siswa, dengan penilaian siswa dapat mengetahui sejauh mana telah berhasil dalam mengikuti pelajaran yang diberikan guru. Hasilnya mungkin memuaskan atau tidak memuaskan. Kalau memuaskan mungkin dapat meningkatkan motivasi belajarnya sehingga lain kali hasilnya lebih memuaskan lagi atau sebaliknya, merasa puas dan kurang giat untuk belajar. Kalau hasilnya tidak memuaskan, mungkin ia akan berusaha lebih giat lagi atau sebaliknya menjadi putus asa.

Bagi guru, dengan penilaian guru dapat mengetahui tentang: (1) tingkat keberhasilan siswa, apakah siswa sudah menguasai pelajaran atau belum, dengan demikian guru dapat lebih memusatkan perhatian untuk memberikan bantuan pada siswa yang belum berhasil. (2) ketepatan materi yang diberikan untuk digunakan pertimbangan dalam penetapan materi berikutnya, dan (3) ketepatan metode pembelajaran yang dilakukan, sehingga guru dapat mawas diri terhadap upaya yang ditempuhnya.

Bagi sekolah, dengan penilaian dapat diketahui bagaimana hasil belajar siswa-siswanya, kesesuaian kondisi belajar, kesesuaian kurikulum, kualitas pembelajaran yang dilangsungkan, dan sebagainya.

Sesuai dengan urutan kejadiannya, dalam proses transformasi penilaian dapat dibedakan atas tiga jenis, yaitu penilaian sebelum, selama, dan sesudah terjadi proses pembelajaran.

3. fungsi Penilaian

Dalam pembelajaran evaluasi memiliki berbagai fungsi, yaitu:

- a. fungsi selektif, yaitu untuk menjaring siswa, mana yang sudah berhasil dan yang belum.
- b. Fungsi diagnostik, yaitu untuk mengetahui kelemahan-kelemahan siswa dalam mempelajari bahan yang diberikan dan sebab-sebabnya.
- c. Fungsi penempatan, yaitu untuk menentukan pada kelompok mana anak itu harus berada.
- d. Fungsi pengukur keberhasilan, yaitu untuk mengetahui sejauh mana program pembelajaran berhasil diterapkan.

4. Ciri-ciri

Dalam pembelajaran, evaluasi memiliki cirri-ciri sebagai berikut:

- D. Dilakukan tidak langsung, yaitu untuk mengukur kepandaian siswa dalam mempelajari bidang studi melalui ukuran kemampuannya dalam menyelesaikan soal-soal yang diberikan.
- E. Menggunakan ukuran kuantitatif, yaitu dengan menggunakan symbol bilangan sebagai hasil pertama pengukurannya, baru diinterpretasikan secara kualitatif.

- F. Menggunakan unit atau ukuran yang tetap, artinya perolehan nilai 6 dianggap memenuhi standar minimal untuk pemberian angka dengan skala 1-10.
- G. Bersifat relatif, artinya penilaian yang dihasilkan bukan sesuatu yang menetap, tetapi ada perubahan dari waktu ke waktu lain. Prestasi belajar saat ini kemungkinan berbeda untuk waktu berikutnya.
- H. Bahwa dalam penilaian pembelajaran tidak menutup kemungkinan terjadinya kesalahan-kesalahan. Sumber kesalahan tersebut dapat terjadi pada factor alat ukurnya (tidak valid), guru (subyektif, terlalu murah/mahal, halo effect-kesan, pengaruh hasil tes sebelumnya, kekliruan menjumlahkan, dsb), siswa (karena factor fisik, emosi, atau sosial), dan situasi (bising, ramai, gaduh, kurang pengawasan).
- I. Menganut prinsip triangulasi, yaitu mempertimbangkan keeratan hubungan antara tiga komponen, yakni: (1) tujuan pembelajaran, (2) KBM, dan (3) evaluasi. Artinya, evaluasi harus mengacu pada tujuan, KBM harus mengacu pada tujuan, dan evaluasi juga harus mengacu pada KBM yang dilaksanakan.
- J. Alat atau instrumen yang digunakan harus baik dan tepat, artinya mampu mengukur apa yang akan dievaluasi.
- K. Pengambilan keputusan hasil evaluasi sepatutnya tidak hanya berdasar atas evaluasi hasil, tetapi juga mempertimbangkan hasil evaluasi terhadap proses yang dilakukan.

D. Pelaksanaan Evaluasi Pembelajaran

1. Alat evaluasi

Alat biasa juga disebut instrumen, yaitu sesuatu yang digunakan untuk mempermudah seseorang dalam melaksanakan tugas atau mencapai tujuan secara lebih efektif dan efisien. Dalam evaluasi fungsi alat juga digunakan untuk memperoleh hasil yang lebih baik sesuai dengan kenyataan yang akan dievaluasi. Alat evaluasi dikatakan baik apabila mampu mengevaluasi sesuatu yang dievaluasi dengan hasil seperti yang dievaluasi.

Dalam menggunakan alat, evaluator menggunakan cara atau teknik, sehingga dikenal adanya teknik evaluasi dan secara garis besar dapat dikelompokkan menjadi dua bagian, yaitu teknik tes dan non tes.

a. Teknik tes

Tes dapat diartikan sebagai suatu alat atau prosedur yang sistematis dan obyektif untuk memperoleh data atau keterangan yang diinginkan seseorang melalui serentetan pertanyaan, tugas, atau latihan. Tes juga dapat diartikan sebagai alat atau prosedur yang digunakan untuk mengetahui atau mengukur sesuatu, dalam suasana, dengan cara, dan dengan aturan-aturan yang sudah ditentukan. Karena itu tes sifatnya lebih resmi karena penuh dengan batasan-batasan.

Ditinjau dari segi kegunaannya dalam pembelajaran , tes dapat dibedakan menjadi 3, yaitu:

1) *Tes diagnostik*

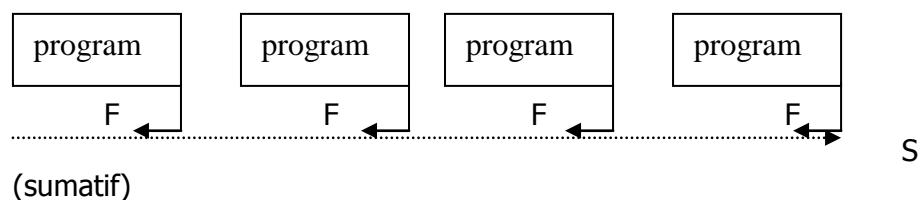
Tes diagnostik adalah tes yang digunakan untuk mengetahui kelemahan-kelemahan siswa sehingga berdasar kelemahan-kelemahan tersebut dapat dilakukan pemberian bantuan atau perlakuan yang tepat. Secara teknis tes diagnostik dapat diberikan kepada siswa sebelum mengikuti program pembelajaran untuk mengetahui kondisi awal (*entering behavior*) atau dalam rangka pengelompokan., pada saat berlangsungnya pembelajaran untuk mengetahui bagian mana dari materi pembelajaran yang belum dikuasai serta deteksi sebabnya guna perencanaan bantuannya, atau pada menjelang akhir pelajaran, sehingga dapat diketahui tingkat penguasaan siswa terhadap materi yang diberikan.

2) *Tes formatif*

Tes formatif adalah tes yang diberikan pada setiap akhir program untuk mengetahui sejauh mana siswa telah "terbentuk" atau menguasai bahan pelajaran setelah mengikuti program tertentu. Tes ini juga dapat dipandang sebagai tes diagnostik, bila didalamnya dimaksudkan untuk mengetahui kelemahan-kelemahannya. Bagi siswa tes ini memberikan manfaat, karena dapat mengetahui bahan yang sudah dan belum dikuasai, dapat sebagai bahan penguatan (*reinforcemant*), untuk bahan perbaikan, sebagai bahan diagnosis (mengetahui kesulitan-kesulitannya). Sedang manfaat bagi guru, dapat dijadikan informasi tentang bahan mana yang sudah dan belum dikuasai siswanya, bahan perbaikan selanjutnya, dan untuk meramalkan sukses tidaknya program yang diberikan.

3) *Tes sumatif*

Tes sumatif adalah suatu kegiatan evaluasi yang diberikan setelah berakhirnya pemberian sekelompok atau seluruh program yang lebih besar, misalnya program untuk satu semester. Manfaatnya untuk menentukan nilai dan kedudukan anak dalam kelompoknya, menentukan dapat tidaknya siswa mengikuti program berikutnya, dan untuk mengisi catatan atau laporan kemajuan belajar siswa (*rapor*). Bila digambarkan, hubungan antara tes formatif dan tes sumatif adalah sebagai berikut:



Dalam kehidupan sehari-hari di sekolah tes formatif dikenal dengan ulangan harian, sedangkan tes sumatif disebut dengan TPB atau THB. Tetapi THB juga dapat disebut sebagai tes formatif atau sub sumatif bila yang dianggap tes sumatif itu sendiri adalah EBTA/UAN.

Bila dibandingkan antara tes diagnostik, formatif, dan sumatif dapat dijelaskan sebagai berikut:

	Tes Diagnostik	Tes Formatif	Tes Sumatif
Funsgi	Bahan prasyarat Tingkat penguasaan Pengelompokkan kesulitan	Umpan balik pelaksanaan program	Penguasaan materi program, posisi siswa dalam kelompoknya
Waktu	Penyaringan Pembagian kelas Selama PBM berlangsung	Selama pelajaran berlangsung	Akhir cawu, semester, tahun, atau akhir pendidikan
Titik berat penilaian	Kognitif, afektif, psikomotor atau fisik, psikologis, dan lingkungan	kognitif	Kognitif, tetapi ada kalanya juga afektif dan psikomotor.
Alat	Tes diagnostik	Tes prestasi belajar	Tes Ujian akhir
Tujuan yang dievaluasi	Keterampilan prasyarat, aspek fisik, mental	Semua tujuan instruksional khusus	Semua tujuan instruksional umum
Tingkat kesulitan	mudah	Tak dapat ditentukan	Merata mulai mudah – sukar
Tingkat pencapaian	Tidak sama, kecuali untuk prasyarat seharusnya 100%	75%	Tidak ada
Cara pencatatan hasil	Dalam bentuk profil	Gagal atau berhasil menguasai tugas	Berdasar skor dari tujuan-tujuan yang dicapai
Informasi belajar	Proses dan hasil	Proses dan hasil	Hasil

b. Teknik non tes

Termasuk teknik non tes adalah: (1) skala bertingkat – rating scale, (2) kuesioner – angket - questionnaire, (3) daftar cocok – check list, (4) wawancara – interview, (5) pengamatan – observation, (6) laporan tugas – task report, (7) catatan siswa, dan (8) catatan sosial,

Documentary Based Questions (DBQ).

DBQ adalah suatu jenis tes yang dibuat dengan cara mengembangkan pertanyaan-pertanyaan yang didasarkan atas benda-benda, dokumen, catatan, gambar, chart, grafik, atau catatan/tulisan yang yang dapat dijadikan sebagai sumber belajar . Tujuannya untuk mengetahui apakah siswa memahami makna yang terkandung di dalamnya. Keuntungan bentuk tes ini adalah mampu mengumpulkan informasi tertulis mengenai kemampuan intelektual tingkat tinggi serta kecenderungan-kecenderungan emosi dan sosial siswa, dan bentuk soal yang digunakan dapat berbentuk soal obyektif ataupun uraian. Apapun bentuk

soalnya, apabila pertanyaan yang diajukan tidak berkaitan dengan pernyataan dalam dokumen (misalnya), maka hal tersebut tidak termasuk DBQ.

2. Masalah tes

Tes yang baik harus memenuhi minimal dua hal, yaitu: (1) bermutu, dan (2) diadministrasikan dengan baik. Berkaitan dengan mutu, tes sebaiknya memenuhi persyaratan validitas (ketepatan), reliabelitas (dapat dipercaya-ajeg), obyektivitas, praktikabilitas (mudah dilaksanakan, mudah pemeriksaannya, dan dilengkapi petunjuk-petunjuk yang jelas), dan ekonomis. Berkaitan dengan pengadministrasian, tes harus memuat dengan jelas petunjuk cara pelaksanaan, cara menskor, dan cara menginterpretasikan.

3. Tas standar dan tes buatan guru

Dalam evaluasi pembelajaran jarang digunakan tes standar atau baku, karena harus melalui prosedur uji coba, analisis, dan revisi sehingga penyusunannya sangat rumit dan perlu waktu lama. Umumnya digunakan tes buatan guru. Tes buatan guru, umumnya disusun atas bahan atau tujuan khusus yang dirumuskan oleh guru kelas, mencakup pengetahuan dan keterampilan yang sempit, disusun sendiri oleh guru tanpa bantuan orang lain/tenaga ahli, jarang diujicobakan terlebih dahulu, reliabilitasnya sedang atau mungkin malah rendah, menggunakan kelompok sebagai acuan normanya, dan tidak ada buku manualnya.

4. Penyusunan tes

Dalam menyusun tes evaluasi pembelajaran, hal-hal yang perlu diperhatikan adalah:

a. Tujuan dan fungsi tes

Tes untuk tujuan dan fungsi yang bersifat diagnostik, formatif, atau sumatif akan menentukan bentuk atau isi tes yang akan dirumuskan, karena masing-masing berbeda.

b. kekomprehensifan

Tes sebaiknya menyeluruh, mampu menggambarkan seluruh kemampuan siswa (berdasar taksonomi), menyebar dari soal yang mudah sampai yang sukar, mewakili seluruh materi yang sudah diberikan (representatif).

c. Kesenambungan (kontinuitas)

Disusun sedemikian rupa yang menggambarkan seluruh materi sejak awal anak mengikuti kegiatan pembelajaran sampai akhir program secara sinambung.

d. Kebervariasian

Tes akan lebih baik dan menarik apabila soal-soal disusun dalam berbagai bentuk secara bervariasi. Misal ada pilihan ganda, benar salah, asosiasi, menjodohkan, pengelompokan, sebab-akibat, isian, uraian singkat, dsb.

Sedangkan langkah-langkahnya adalah sebagai berikut:

- a. Menentukan tujuan mengadakan tes
- b. Mengadakan pembatasan bahan yang akan diteskan
- c. Mengidentifikasi TIK dari tiap bagian atau pokok bahasan

- d. Menderetkan semua TIK dan mengidentifikasi tingkah laku yang dikehendaki
- e. Menyusun table spesifikasi (kisi-kisi) yang memuat pokok materi, aspek berpikir yang akan diukur (ingatan, pemahaman, aplikasi, analisis, sintesis, dan evaluasi).
- f. Menuliskan butir-butir soal.
- g. Membuat lembar jawaban
- h. Membuat kunci jawaban
- i. Membuat pedoman penilaian

5. Tes tertulis untuk prestasi belajar

a. Bentuk-bentuk tes

1) tes subyektif, esai, atau uraian

Tes ini memerlukan jawaban yang bersifat pembahasan atau uraian, bisa uraian terbatas atau uraian bebas. Soal atau pertanyaannya biasanya dimulai dengan kata-kata: jelaskan, uraikan, mengapa, bagaimana, bandingkan, dan sebagainya. Kebaikannya, mudah dipersiapkan, tidak mengundang spekulasi, mendorong siswa untuk mengemukakan pendapatnya, menuntut elaborasi lanjut, dan dapat mengetahui kedalaman penguasaan materi oleh siswa. Keburukannya, kadar validitas dan reliabilitas rendah, kurang representatif, mengundang masuknya unsur subyektif, sulit pemeriksaannya dan perlu waktu lama. Dalam penyusunan hendaknya usahakan komprehensif, variatif, mudah dipahami, dan tegaskan model jawaban yang dikehendaki.

2) tes obyektif

Kebaikannya lebih representatif, obyektif, dan lebih cepat pemeriksaannya bahkan bisa diserahkan pada orang/guru lain. Kelemahannya, penyusunannya sulit, cenderung mengungkap aspek ingatan, mengundang sikap spekulatif, membuka peluang kerja sama antar siswa.

b. Macam-macam tes obyektif

Termasuk tes obyektif adalah: (1) tes benar-salah: with correction atau without correction, (2) tes pilihan ganda (multiple choice test): biasa, hubungan sebab-akibat, atau asosiasi, (3) Menjodohkan (matching test), (4) tes isian (completion test)

6. Tabel spesifikasi

Dalam evaluasi pembelajaran, perlu penggunaan tabel spesifikasi, grid, atau kisi-kisi agar tes yang disusun tidak menyimpang dari bahan, TIK, KBM, dan aspek tingkah laku yang akan dicakup dalam tes. Aspek yang akan diungkap paling tidak mencakup tiga hal, yaitu ingatan, pemahaman dan aplikasi. Namun dapat juga mencantumkan aspek yang lain.

Tabel spesifikasi mempunyai kolom dan baris sehingga tampak hubungan materi dengan aspek yang tergambar dalam TIK. Namun sebenarnya guru tidak hanya mengingat hubungan antara dua hal tersebut tetapi empat hal, yaitu bahan, TIK, KBM, dan evaluasi karena hal-hal ini berhubungan sangat erat. Misal apabila materinya fakta, maka TIK-nya ingatan, KBM-nya informasi, dan evaluasinya dapat uraian, isian singkat, B-S, atau pilihan ganda. Namun bila bahannya konsep atau hubungan antar konsep maka TIK, KBM, dan evaluasinya juga akan lain.

Hubungan ke empat hal tersebut dapat dicontohkan sebagai berikut:

- TIK : siswa dapat menyebutkan kota-kota besar di Pulau Jawa
 Materi : Kota-kota besar di Pulau Jawa
 KBM : Informasi dan Tanya jawab
 Evaluasi : Termasuk tiga kota besar di pulau Jawa adalah
- Jakarta, Bandung, Bogor
 - Jakarta, Bandung, Surabaya
 - Jakarta, Yogyakarta, Surabaya

Contoh table spesifikasi adalah:

PM \ aspek	Ingatan (%)	Pemahaman (%)	Aplikasi (%)	Jumlah (100%)
Pokok materi 1 (%)	(a)	(b)	©	
Pokok materi 2 (%)	(d)	(e)	(f)	
Pokok materi 3 (%)	(g)	(h)	(i)	
Pokok materi 4 (%)	(j)	(k)	(l)	
Jumlah (100%)				

Langkah-langkah pembuatan:

- Identifikasi pokok-pokok materi yang akan diteskan
- Berikan perimbangan bobot berdasar luasnya materi atau kepentingannya untuk di teskan, dalam bentuk indek dan kemudian ubah dalam bentuk persentase, serta tuliskan dalam kurung di belakang masing-masing pokok materi.

Misal:

- PM 1 (2) menjadi $2/14 \times 100\%$ = 14 %
 PM 2 (4)..... menjadi $4/14 \times 100\%$ = 29 %, dst
 PM 3 (5)
 PM 4 (3)

- Tentukan banyak soal yang akan dibuat berdasar waktu yang tersedia dan bentuk soal yang akan dibuat.
- Memerinci banyaknya soal untuk tiap pokok materi, caranya dengan mengalikan jumlah soal yang akan dibuat dengan perimbangan bobotnya. Misal jumlah soal = 50 buah, maka banyak soal untuk PM 1 = $14\% \times 50 = 7$ soal
- Untuk menentukan jumlah soal pada masing-masing aspek tingkah laku yang akan diukur, tergantung homogenitas atau heterogenitas materi (pokok bahasan) yang akan diteskan.

1) Untuk materi yang homogen, misal persentasenya: ingatan 50%, pemahaman 30%, aplikasi 20%. (tapi ini tidak baku tergantung jenjang pendidikannya). Berdasar contoh di atas, untuk PM 1 dapat diketahui bahwa jumlah soal untuk materi ingatan = $50/100 \times 7$ soal = 3,5 (4 soal), pemahaman = $30/100 \times 7$ soal = 2,1 (2 soal), dan aplikasi = $20/100 \times 7$ soal = 1,4 (1 soal), dan seterusnya. Cara lainnya adalah dengan menghitung dulu jumlah soal untuk masing-masing sel, baru diketahui jumlahnya. Misal untuk sel PM 1-Ingatan (sel a) , maka jumlah soalnya = $50/100 \times 14/100 \times 50$ soal = 3,5 (4 soal), dst.

- 2) Untuk materi yang tidak seragam (misal PM 1, ingatan= 50%, pemahaman= 30%, dan aplikasi= 20% sedangkan untuk PM 2 ingatan= 30%, pemahaman= 50% dan aplikasi= 20%) untuk menentukan jumlah soal dapat dilakukan sesuai contoh-contoh di atas.

Setelah selesai membuat table spesifikasi, langkah berikutnya adalah:

- Menentukan bentuk soal. Untuk materi tentang fakta (seperti: istilah, tahun, dan nama) masuk dalam kelompok ingatan dan bentuk soalnya dapat benar-salah. Untuk materi konsep/hubungan dua konsep, masuk dalam kelompok pemahaman dan bentuk soalnya dapat pilihan ganda, sedangkan untuk mempertentangkan konsep, generalisasi, atau menghubungkannya dengan kehidupan sehari-hari masuk dalam kelompok aplikasi dan bentuk soalnya dapat tes uraian.
- Menuliskan soal-soalnya. Usahakan bahasa sederhana, tidak mengandung penafsiran ganda, dan jelaskan petunjuk pengerjaannya untuk masing-masing kelompok soal (bila bentuk soalnya beragam).
- Mencantumkan nomor-nomor soal dalam table sesuai dengan spesifikasinya.

7. Menganalisis Hasil tes

Pada anak normal, hasil tes pembelajaran yang baik akan menggambarkan kurva normal. Sedangkan untuk anak-anak luar biasa kemungkinan itu sulit didapat, disamping populasinya rendah juga kondisinya cenderung homogen. Untuk itu dalam kepentingan PLB analisis dapat dilakukan secara sederhana, yaitu dengan meneliti secara jujur soal-soal yang sudah dibuat, baik berdasar aspek bahasa, taraf kesukaran, kecenderungan jawaban siswa, kejelasan petunjuk, penyebaran, dan sebagainya. Misalnya dengan pertanyaan-pertanyaan: apakah soal yang kita susun cukup mudah dipahami, tidak membingungkan, dapat dikerjakan, tidak terlalu sulit, tingkat kesukarannya cukup menyebar, materinya sudah kita ajarkan, dsb.

8. Menskor dan Menilai

Untuk menentukan skor pada soal benar-salah kita dapat menggunakan dua cara, yaitu:

- tanpa hukuman atau tanpa denda
Skor = jumlah jawaban benar – jumlah jawaban salah ($S=R-W$)
- dengan hukuman atau dengan denda
Rumusnya : $S = T - 2W$, dimana T = jumlah soal

Untuk soal pilihan ganda, bila dengan hukuman dapat digunakan rumus:

$$S = R - \frac{(W)}{(n - 1)}$$

S = score; R = jumlah jawaban benar, W = jumlah jawaban salah, dan n = banyaknya pilihan jawaban.

Untuk menskor soal dalam bentuk jawaban singkat dan menjodohkan, mengingat relatif lebih sulit dari pada soal bentuk B-S atau pilihan ganda, maka setiap soal dapat diberi bobot dua. Bila ini yang ditempuh maka penskorannya harus dikalikan 2.

Untuk soal esai, penskoran relatif sulit. Namun, secara umum dapat digunakan dengan menggunakan: (1) **norm reference test**, dengan menggunakan kelompok sebagai pembanding caranya dengan membaca semua jawaban soal pertama, sehingga diketahui sitausi jawabannya baru kemudian dilakukan penilaian untuk masing-masing berdasar kelengkapan, kerincian, atau kedalaman jawabannya. Baru kemudian dilanjutkan soal kedua, dan seterusnya. Skor maksimum dijadikan patokan untuk menentukan skor yang lain. Misal skor maksimum 40, si B dapat skor 20, maka skor B = $20/40 \times 100\% = 50\%$, jadi skor B=50. Untuk di SLB penerapan model ini sulit dan karena berbagai hal menjadi kurang relevan. Dan (2) **criterion reference test**, caranya tidak dibandingkan dengan kelengkapan jawaban dari siswa lain, tetapi dibandingkan dengan jawaban lengkap yang dikehendaki guru, standar mutlak yaitu standar 100 atau 10.

9. Mengolah nilai

Untuk kepentingan evaluasi dalam pembelajaran di SLB, pengolahan nilai biasanya menggunakan skala penilaian 1-10 atau 1-100, bukan skala bebas dan skala huruf. Sedangkan standar nilai yang digunakan biasanya menggunakan standar enam (berkisar angka 4 – 9), bukan Standar Nines atau Staninines (angka 1-9). Pada standar enam, angka 4 diinterpretasikan kurang sekali, 5 kurang, 6 cukup, 7 lebih dari cukup, 8 baik, dan 9 baik sekali.

Untuk mengubah skor (nilai mentah) menjadi nilai jarang menggunakan standar 10 disamping sulit karena harus mencari terlebih dahulu mean, standar deviasi, dan table konversi, hal ini juga kurang relevan dengan asas-asas dalam pembelajaran untuk anak luar biasa. Untuk jenis anak tunagrahita menilai dengan angka malah menjadi tidak relevan karena lebih mengutamakan deskri tentang kemampuan-kemampuan yang dicapai sebagai gambaran tentang pencapaian prestasi belajarnya.

10. Menetapkan nilai akhir

Nilai (dalam bentuk kuantitatif maupun kualitatif) mempunyai berbagai fungsi, yakni fungsi instruksional (gambaran seberapa jauh siswa mencapai tujuan pembelajaran), informative (sebagai laporan terhadap kemajuan belajar anaknya kepada orang tua/wali), bimbingan (untuk mengetahui bagian-bagian mana yang masih memerlukan bantuan), dan adminsitratif (keperluan kenaikan kelas, mutasi, pemberian beasiswa, studi lanjut, dsb).

Dalam menetapkan nilai akhir, akan sangat bijaksana apabila guru tidak sekedar berpatokan pada evaluasi hasil, berdasar nilai yang diperoleh dalam tes formatif dan sumatif, tetapi juga mempertimbangkan segi proses. Untuk itu, aspek-aspek kerajinan, motivasi, usaha, atau aspek-aspek pribadi dan sosial perlu menjadi pertimbangan tersendiri, sehingga dapat menggambarkan prestasi belajar secara lebih komprehensif.

E. Evaluasi Program Pembelajaran

Evaluasi program pembelajaran, , adalah suatu rangkaian kegiatan yang dilakukan dengan sengaja untuk melihat tingkat keberhasilan program pembelajaran yang sudah dilaksanakan. Tujuannya untuk memperoleh masukan sebagai bahan pertimbangan dalam mengambil kebijakan-kebijakan dalam pembelajaran berikutnya. Masukan-masukan tersebut terutama dalam hubungannya dengan dimana letak kelemahan-

kelemahan atau faktor-faktor yang menjadi penghambat dari keseluruhan program pembelajaran yang telah direncanakan sebelumnya. Berdasar ini mungkin akan diperoleh masukan tentang hal-hal atau aspek-aspek apa yang perlu dipertahankan, disempurnakan, dimodifikasi, ditingkatkan, atau bahkan harus dibuang sehingga proses pembelajaran selanjutnya dapat diperkirakan atau diramalkan menjadi lebih baik terutama dalam kepentingannya dengan kemajuan belajar siswa.

Dalam evaluasi program pembelajaran, obyek atau sasaran evaluasi adalah seluruh komponen yang terlibat dengan proses pembelajaran (proses transformasi), yang meliputi: masukan mentah (raw input: siswa dengan segala karakteristiknya), masukan instrumental (instrumental input: materi/kurikulum, guru, metode mengajar, sarana/media, dan sistem evaluasi), serta factor masukan lingkungan (environmental input: baik lingkungan manusia maupun bukan manusia).

Cara yang dapat ditempuh untuk evaluasi program yang paling efektif adalah melalui penelitian tindakan kelas. Dengan cara ini guru dituntut untuk terus berpikir kritis dan sistematis guna memperoleh masukan-masukan berharga bagi perbaikan dalam pembelajaran yang dilakukannya menuju dicapainya prestasi siswa secara maksimal dan memuaskan semua pihak. Untuk itu kegiatan penelitian kelas ini harus dibudayakan. Cara lain adalah dengan menyusun dan mencari jawaban terhadap seluruh komponen, atau komponen-komponen tertentu yang dirasakan menghambat proses pembelajaran telah yang dilakukan.

F. Evaluasi Proses Pembelajaran Di SLB

1. Permasalahan / kecenderungan

- a. Dilaksanakan sepintas
- b. Tanpa kisi-kisi (terutama untuk formatif)
- c. kurang representatif (komplain eksternal)
- d. kurang komprehensif (pengetahuan – cend ingatan)
- e. kurang variatif

2. Idealnya

- a. Menyesuaikan dgn karakteristik perilaku belajar siswa :
 - 1) *Tunanetra* : lisan, braille : tambah waktu
 - 2) *Tunarungu* : Bahasa jelas, padat, singkat, sederhana.
 - 3) *Tunagrahita* : Bahasa jelas, singkat, sederhana, gambar, perbuatan (pragmatis)
 - 4) *Tunadaksa* : relatif sama anak normal, kecuali perbuatan/fisik
 - 5) *Tunalaras* : Relatif sama, aspek afeksi perlu penekanan, perbuatan (pragmatis)
- b. Gunakan dgn alat yang bervariasi
- c. Kontinuitas
- d. konteks inklusi : pembandingan diri sendiri
- e. komprehensif (hasil & proses)

Sumber utama:

Hamid Hasan, S. (1996), **Pendidikan Ilmu Sosial**, Jakarta: Dirjen Dikdik – PPTA
Suharsimi Arikunto (1999), **Dasar-Dasar Evaluasi Pendidikan**, Jakarta: Bumi Aksara.