

SISTEM PENILAIAN HASIL BELAJAR

A. Evaluasi Acuan kriteria (*Criterion Referenced Evaluation*)

Pengukuran keberhasilan belajar didasarkan atas penafsiran tingkah laku (performance) yang didasarkan atas kriteria atau standar khusus, artinya derajat penguasaan yang ada didasarkan pada tingkat tertentu yang harus dicapai.

B. Evaluasi Acuan Norma (*Norm Referenced Evaluation*).

Pengukuran ini mendudukan individu pada kelompoknya, membandingkan penguasaan individu terhadap rata-rata penguasaan kelompok

Terdapat dua acuan penilaian yaitu penilaian acuan norma dan penilaian acuan patokan

Tujuan penilaian acuan norma adalah untuk mengetahui kedudukan peserta didik dalam kelompoknya (dalam kelas). Oleh karena itu butir-butir soal yang dipakai dalam ujian tidak boleh terlalu sukar atau terlalu mudah, sehingga kisaran indeks kesukarannya 0,3 sampai 0,7. Disamping itu harus dapat membedakan mana peserta didik yang pandai dan yang tidak pandai. (Djemari, 2008)

Karakteristik peniaian acuan norma:

1. Terdapat unsur kompetitif
2. Sangat baik untuk penilaian afektif dan kognitif
3. Tidak dapat untuk menilai kemampuan skill atau materi tertentu
4. Tidak dapat memberi interpretasi secara langsung pada suatu skala
5. Nilai tidak mencerminkan kemampuan yang rinci

Acuan kriteria berasumsi bahwa hampir semua orang bisa belajar apa saja namun waktunya yang berbeda. Konsekuensi dari acuan kriteria adalah adanya program remidi, program pengayaan, dan program percepatan. Penafsiran hasil tes selalu dibandingkan dengan standar atau kriteria yang telah ditetapkan dulu (Djemari, 2004).

Karakteristik penilaian acuan kriteria:

1. Terdapat kemampuan kognitif minimal yang harus dimiliki oleh peserta didik
2. Adanya kemampuan psikomotorik dan sikap mental minimal sebagai prasyarat
3. Meletakkan perbedaan latar belakang peserta didik sebagai unsur individual
4. Sebagai alat diagnosis kesulitan siswa
5. Dapat dufungsikan sebagai embrio tes baku
6. Tidak komparatif terhadap kelompok sehingga dapat melemahkan semangat kompetisi

Terdapat empat jenis penilaian acuan kriteria:

1. *Entry-behaviors test*, yaitu suatu tes yang diadakan sebelum suatu program pengajaran dilaksanakan dan bertujuan untuk mengetahui sampai batas mana penguasaan pengetahuan dan keterampilan yang telah dimiliki peserta didik yang dapat dijadikan dasar untuk menerima program pengajaran yang akan diberikan
2. *Pre-test*, yaitu tes yang diberikan sebelum pengajaran dimulai dan bertujuan untuk mengetahui sampai dimana penguasaan peserta didik terhadap bahan pengajaran (pengetahuan dan keterampilan) yang akan diajarkan
3. *Post-test*, yaitu tes yang diberikan pada setiap akhir program satuan pengajaran dan bertujuan untuk mengetahui sampai dimana pencapaian siswa terhadap bahan pengajaran setelah mengalami suatu kegiatan belajar
4. *Embedded –test*, yaitu tes yang dilaksanakan disela-sela atau pada waktu-waktu tertentu selama proses pengajaran berlangsung dan bertujuan untuk mengetes peserta didik secara langsung sesudah suatu unit pengajaran sebelum post-test dan untuk mencek kemajuan siswa untuk remedial sebelum *post-test*

PENGADMINISTRASIAN TES HASIL BELAJAR

A. Penyusunan perangkat tes

1. Butir tes disusun dari PB awal – akhir
2. Tingkat kesukaran disusun dari mudah – sulit
3. Butir tes yang tipenya sama dikelompokkan dalam satu kelompok
4. Petunjuk mengerjakan tes secara jelas
5. Butir tes disusun dengan baik agar mudah dibaca
6. Stem dan option disusun dalam satu halaman

B. Pelaksanaan Tes

1. Ruang tempat tes harus tenang dari keramaian
2. Tempat duduk diatur dengan jarak tertentu
3. Agar dapat memulai mengejakan bersamaan, lembar soal diletakkan secara terbalik
4. Sebelum tes dimulai perlu ditentukan sangsi
5. Pengawas berlaku wajar
6. Disiapkan daftar hadir yang harus ditandatangani seluruh peserta tes
7. Jika waktu habis peserta diminta menghentikan pekerjaan

8. Membuat berita acara pelaksanaan tes secara lengkap

Petunjuk praktis pelaksanaan tes lisan:

1. Membuat inventarisasi berbagai jenis soal
2. Butir tes yang dipilih, harus disiapkan pedoman jawabannya
3. Skor atau nilai hasil tes lisan harus sudah ditentukan disaat peserta tes selesai dites
4. Jangan sampai menimbulkan rasa takut, panik
5. Jangan berubah menjadi diskusi, simpati, kasihan
6. Ada ancar-ancar waktu yang disediakan
7. Pertanyaan yang diajukan bervariasi

C. Pemeriksaan Tes Hasil Belajar

1. Bentuk Uraian

a. Membuat kunci jawaban sebagai patokan

b. Jika mengacu standar mutlak

1) membandingkan jawaban dengan pedoman

2) memberi skor setiap butir pada bagian kiri

3) menjumlahkan skor yang telah diberikan

c. Jika mengacu standar relatif

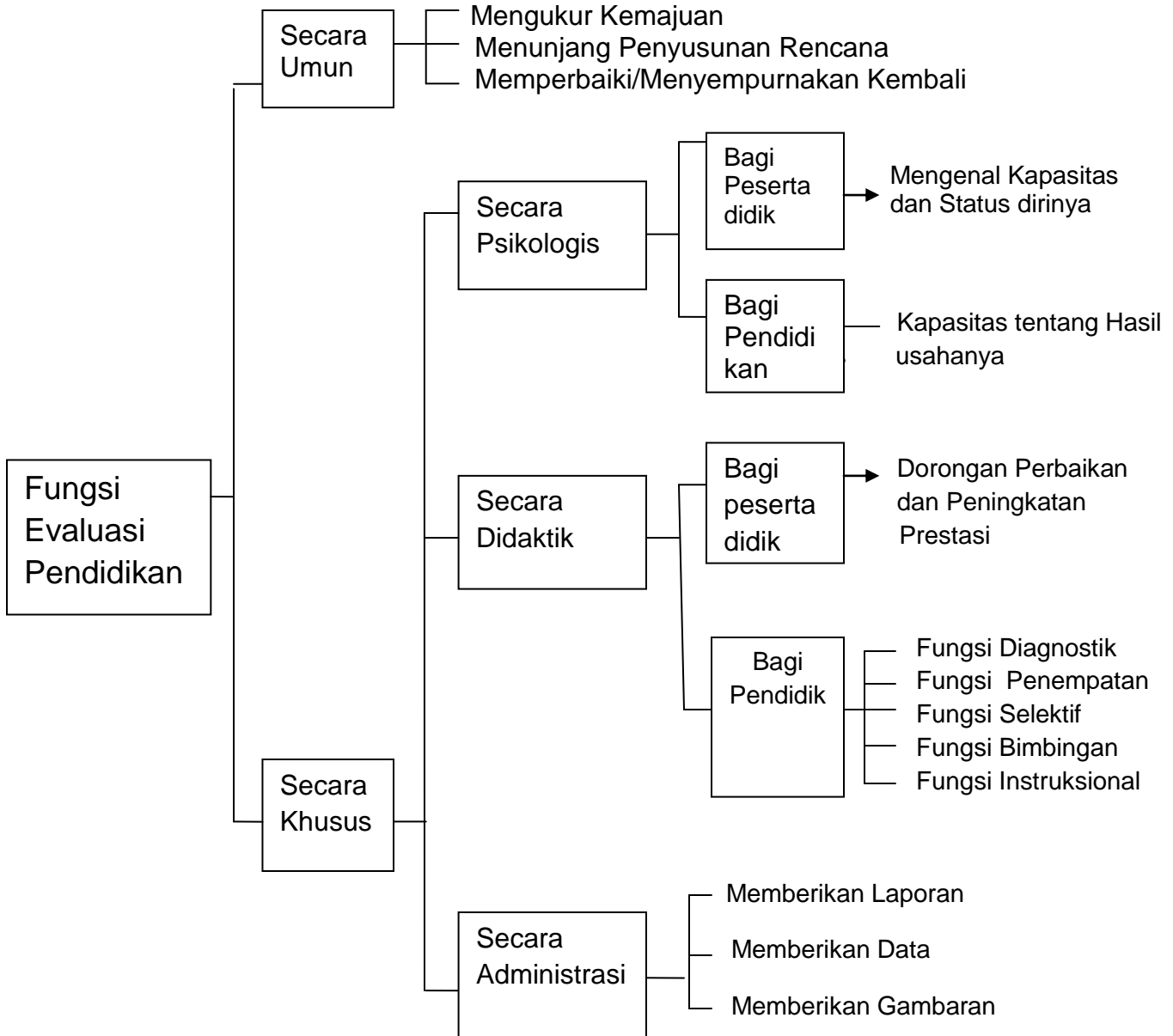
1) memeriksa butir soal nomor 1 untuk memperoleh jawaban seluruh peserta tes

2) memberi skor nomor 1 kepada seluruh peserta tes sesuai tingkat kelengkapan jawaban

3) Dilanjutkan nomor tes berikutnya dengan cara yang sama dan diakhiri dengan penjumlahan skor

D. Pemberian Skor Tes

Pemberian skor merupakan langkah pertama dalam proses pengolahan hasil tes, yaitu proses pengubahan jawaban-jawaban soal tes menjadi angka



JENIS-JENIS EVALUASI

1. EVALUASI FORMATIF

LANGKAH-LANGKAH PENYUSUNAN TES

Untuk dapat memperoleh alat penilaian (tes) yang memenuhi persyaratan, setiap penyusun tes hendaknya dapat mengikuti langkah-langkah penyusunan tes. Sax (1980), mengidentifikasi langkah-langkah pengembangan tes ke dalam sembilan langkah sebagai berikut:

1. Menyusun kisi-kisi (tabel spesifikasi) tes, yang memuat: materi pokok yang akan diteskan, aspek perilaku atau tingkatan kognitif yang akan diukur, dan penentuan jumlah butir tes untuk setiap aspeknya.
2. Menulis butir-butir soal dengan mendasarkan pada aspek-aspek yang telah tercantum pada tabel spesifikasi (kisi-kisi) tersebut.
3. Melakukan telaah soal tes (analisis tes secara logis);

4. Melakukan uji coba soal;
5. Analisis soal secara empiris;
6. Memperbaiki atau merevisi tes;
7. Merakit tes, dengan menyiapkan komponen-komponen pendukung untuk penyelenggaraan tes, yang meliputi: (a) buku tes; (b) lembar jawaban tes; (c) kunci jawaban tes; dan (d) pedoman penilaian atau pedoman pemberian skor.
8. Melaksanakan tes; dan
9. Menafsirkan hasil tes.

TABEL SPESIFIKASI (KISI-KISI)

Materi Pokok	Aspek yang akan diukur						Jumlah butir
	Ingatan 30% (6)	Pemahaman 30% (6)	Aplikasi 20% (4)	Analisis 20% (4)	Sintesis	Evaluasi	
Pengertian busana (13%)	1	1	-	-			2
Fungsi busana (17%)	1	1	-	1			3
Macam-macam busana (40%)	2	2	2	2			8
Persyaratan busana (30%)	2	2	2	1			7

PENILAIAN KOGNITIF

1. Pengetahuan/hafalan/ingatan (*knowledge*)

Adalah kemampuan seseorang untuk mengingat ingat kembali (*recall*) atau mengenali kembali tentang nama, istilah, ide, rumus-rumus, dan sebagainya, tanpa mengharapkan kemampuan untuk menggungkannya. Pengetahuan atau ingatan adalah merupakan proses berfikir yang paling rendah.

2. Pemahaman (*comprehension*)

Adalah kemampuan seseorang untuk mengerti atau memahami sesuatu setelah sesuatu itu diketahui dan diingat. Dengan kata lain, memahami adalah mengetahui tentang sesuatu dan dapat melihatnya dari berbagai segi. Seseorang peserta didik dikatakan memahami sesuatu apabila ia dapat memberikan penjelasan atau memberi uraian yang lebih rinci tentang hal itu dengan menggunakan kata-katanya sendiri.

3. Penerapan (*application*)

- Adalah kesanggupan seseorang untuk menerapkan atau menggunakan ide-ide umum, tata cara ataupun metode-metode, prinsip-prinsip, rumus-rumus, teori-teori dan sebagainya, dalam situasi yang baru dan kongkret. Penerapan ini adalah merupakan proses berfikir setingkat lebih tinggi ketimbang pemahaman.

4. Analisis (*analysis*)

- Adalah kemampuan seseorang untuk merinci atau menguraikan suatu bahan atau keadaan menurut bagian-bagian yang lebih kecil dan mampu memahami hubungan di antara bagian-bagian atau faktor-faktor yang satu dengan faktor-faktor lainnya. Jenjang analisis adalah setingkat lebih tinggi ketimbang jenjang aplikasi.

5. Sintesis (*syntesis*)

- Adalah kemampuan berfikir yang merupakan kebalikan dari proses berfikir analisis. Sintesis merupakan suatu proses yang memadukan bagian-bagian atau unsur-unsur secara logis, sehingga menjelma menjadi suatu pola yang berstruktur atau berbentuk pola baru

6. Penilaian/penghargaan/evaluasi (*evaluation*)

- Adalah merupakan jenjang berpikir paling tinggi dalam ranah kognitif dalam taksonomi Bloom. Penilaian/evaluasi disini merupakan kemampuan seseorang untuk membuat pertimbangan terhadap suatu kondisi, nilai atau ide, misalkan jika seseorang dihadapkan pada beberapa pilihan maka ia akan mampu memilih satu pilihan yang terbaik sesuai dengan patokan-patokan atau kriteria yang ada.

Tes Bentuk Pilihan Ganda

- Butir soal tes bentuk pilihan ganda ini merupakan salah satu bentuk tes obyektif yang paling luwes dan banyak dikembangkan akhir-akhir ini, karena dapat digunakan untuk mengukur berbagai tataran pengetahuan dari yang sederhana sampai yang kompleks.
- Tes pilihan ganda terdiri atas bagian pokok soal/pertanyaan yang disebut STEM, dan bagian alternatif jawaban disebut OPTIONS. Opsi jawaban, terdiri atas: satu jawaban BENAR, yaitu kunci jawaban, dan beberapa alternatif jawaban yang disebut pengecoh (distraktor).

Kaidah Penyusunan Butir Tes Pilihan Ganda

1. Soal harus sesuai dengan indikator.
2. Pokok soal harus dirumuskan secara jelas dan tegas.
3. Rumusan pokok soal dan pilihan jawaban hendaknya merupakan pertanyaan yang diperlukan saja. Rumusan persoalan hendaknya jangan bertele-tele yang tidak relevan dengan persoalan.
4. Pokok soal hendaknya jangan memberi petunjuk ke arah jawaban yang benar. Atau hindarkan sifat asosiatif antara pokok soal dengan alternatif jawabannya.
5. Pokok soal hendaknya jangan menggunakan pernyataan yang bersifat negatif ganda.
6. Pilihan jawaban harus homogen dan atau logis ditinjau dari segi materi.

7. Semua alternatif jawaban benar (kunci jawaban) hendaknya harus sulit dibedakan dengan pengecoh-pengecohnya, khususnya bagi mereka yang belum mencapai tujuan belajarnya.
8. Panjang rumusan pilihan jawaban hendaknya relatif sama. Atau, panjang alternatif jawaban hendaknya tidak memberikan isyarat akan jawaban yang benar.
9. Pilihan jawaban hendaknya jangan menggunakan pernyataan yang berbunyi "semua pilihan jawaban di atas salah" atau "semua jawaban di atas benar".
10. Pilihan jawaban yang berbentuk angka harus disusun berdasarkan ukuran besar kecilnya, pengurutan tersebut dimaksudkan untuk memudahkan peserta tes melihat pilihan jawabannya.

11. Gambar/grafik/tabel/diagram dan sejenisnya yang terdapat pada soal harus berfungsi.
12. Setiap soal harus mempunyai satu jawaban yang benar atau paling benar.
13. Butir soal hendaknya jangan bergantung pada jawaban soal sebelumnya.
14. Penempatan alternatif jawaban yang benar (kunci jawaban) hendaknya tidak mengikuti pola sistematis, sehingga tidak memberikan isyarat secara jelas kepada peserta tes tentang jawaban yang benar.

RANAH AFFEKTIF

Domain affektif dipergunakan dalam pengukuran penampilan khusus. Domain affektif berkaitan dengan sikap, perasaan, apresiasi, dan minat.

Terdapat 5 kategori dalam domain affektif, yaitu:

1. Menerima (*receiving*)

Kesediaan dan kesiapan siswa untuk menerima suatu fenomena khusus atau suatu rangsangan tertentu

TU: mendengarkan dengan tekun, menunjukkan kesadaran pentingnya belajar

TK: identifikasikan, pilihlah, tunjukkan, dll

2. Merespon (*responding*)

Responding menunjuk kepada partisipasi aktif dari siswa. Terhadap fenomena yang ada tidak cukup hanya memperhatikan tetapi melibatkan diri secara aktif dalam situasi tersebut

TU: berpartisipasi aktif dalam diskusi

TK: berinisiatif, menceriterakan dll

3. Penghargaan (*valuing*)

Valuing menunjuk kepada penghargaan atau penilaian oleh siswa terhadap obyek khusus, fenomena atau perilaku

TU: menghargai karya yang baik (seni, musik), memiliki suka memecahkan masalah

TK: lengkapi, deskripsikan, ikuti, buat keputusan

4. Mengorganisasikan (*organization*)

Menunjuk kepada mengorganisasikan nilai-nilai dari berbagai nilai yang berbeda-beda, bahkan bertolak belakang satu dengan yang lain

TU: bertanggung jawab atas perbuatan sendiri, mengerti dan dapat menerima kelebihan dan keterbatasannya

TK: buat pilihan, modifikasikan, organisasikan, hubungkan

5. Mempribadi (*mewatak*)

Individu telah memiliki suatu sistem nilai yang mantap yang dapat mengontrol perilakunya dalam kehidupan sehari-hari

TU: dapat melaksanakan kerja kelompok dengan baik

TK: praktekan, paparkan, pecahkan

DOMAIN PSIKOMOTOR

Domain psikomotor menyangkut keterampilan gerak otot (phisik), dapat dibagi atas 7 kategori, yaitu:

1. Menirukan (*perception*),
Berkenaan dengan perasaan untuk menirukan gerakan yang ditunjukkan
 - a. kesadaran /adanya rangsangan
 - b. seleksi terhadap gerakan-gerakan yang relevan
 - c. melaksanakan dalam suatu penampilanTU: dapat membedakan rasa macam-macam makanan
2. Set (*setting*): memanipulasikan
Kesiapan untuk melaksanakan suatu tindakan yang khusus (dapat menampilkan suatu tindakan (*action*) seperti yang diajarkan
3. Imitasi (*Guided Response*)
Kecepatan gerakan sesuai dengan apa yang telah ditentukan oleh instruktur
4. Mekanisme (*mechanism*)
Berkaitan dengan penampilan tindakan dimana hal-hal yang diajarkan telah menjadi satu kebiasaan gerakan-gerakan yang ditampilkan menjadi lebih meyakinkan dan lebih ahli

5. Responsi yang bebas dan kompleks (*complex overt response*)
Menunjuk pada penampilan kecakapan gerak pada pola-pola gerakan yang kompleks
6. Adaptasi (*adaptation*)
Berkaitan dengan keterampilan yang sudah dikembangkan begitu baiknya sehingga individu sudah dapat memodifikasi pola-pola gerakan yang ada untuk keperluan-keperluan khusus guna menghadapi kesulitan-kesulitan yang timbul
7. Mengorganisasi (*Organization*)
Menunjuk kepada penciptaan pola-pola gerakan baru pada situasi yang khusus atau problema yang khusus.
Penekanannya kepada kreativitas yang didasarkan atas keterampilan yang tinggi perkembangannya

SISTEM PENILAIAN HASIL BELAJAR

Pada dasarnya sistem evaluasi dapat dibagi dua, yaitu:

1. Evaluasi acuan kriteria (*Criterion Referenced Evaluation*)
2. Evaluasi Acuan Norma (*Norm Referenced Evaluation*).

Pada evaluasi acuan norma menempatkan kedudukan kemampuan siswa dibandingkan dengan kelompoknya yang mengikuti tes pada saat yang sama.

Sedangkan evaluasi acuan kriteria interpretasi skor tes didasarkan atas seberapa jauh kemampuan yang dicapai siswa dibandingkan dengan standar yang telah ditetapkan sebelumnya (Joesmani, 1988)

Terdapat dua acuan penilaian yaitu penilaian acuan norma dan penilaian acuan patokan

Tujuan penilaian acuan norma adalah untuk mengetahui kedudukan peserta didik dalam kelompoknya (dalam kelas). Oleh karena itu butir-butir soal yang dipakai dalam ujian tidak boleh terlalu sukar atau terlalu mudah, sehingga kisaran indeks kesukarannya 0,3 sampai 0,7. Disamping itu harus dapat membedakan mana peserta didik yang pandai dan yang tidak pandai. (Djemari, 2008)

Karakteristik penilaian acuan norma:

1. Terdapat unsur kompetitif
2. Sangat baik untuk penilaian afektif dan kognitif
3. Tidak dapat untuk menilai kemampuan skill atau materi tertentu
4. Tidak dapat memberi interpretasi secara langsung pada suatu skala
5. Nilai tidak mencerminkan kemampuan yang rinci

Acuan kriteria berasumsi bahwa hampir semua orang bisa belajar apa saja namun waktunya yang berbeda. Konsekuensi dari acuan kriteria adalah adanya program remidi, program pengayaan, dan program percepatan. Penafsiran hasil tes selalu dibandingkan dengan standar atau kriteria yang telah ditetapkan dulu (Djemari, 2004).

Karakteristik penilaian acuan kriteria:

1. Terdapat kemampuan kognitif minimal yang harus dimiliki oleh peserta didik
2. Adanya kemampuan psikomotorik dan sikap mental minimal sebagai prasyarat
3. Meletakkan perbedaan latar belakang peserta didik sebagai unsur individual
4. Sebagai alat diagnosis kesulitan siswa
5. Dapat dufungsikan sebagai embrio tes baku
6. Tidak komparatif terhadap kelompok sehingga dapat melemahkan semangat kompetisi

Terdapat empat jenis penilaian acuan kriteria:

1. *Entry-behaviors test*, yaitu suatu tes yang diadakan sebelum suatu program pengajaran dilaksanakan dan bertujuan untuk mengetahui sampai batas mana penguasaan pengetahuan dan keterampilan yang telah dimiliki peserta didik yang dapat dijadikan dasar untuk menerima program pengajaran yang akan diberikan
2. *Pre-test*, yaitu tes yang diberikan sebelum pengajaran dimulai dan bertujuan untuk mengetahui sampai dimana penguasaan peserta didik terhadap bahan pengajaran (pengetahuan dan keterampilan) yang akan diajarkan
3. *Post-test*, yaitu tes yang diberikan pada setiap akhir program satuan pengajaran dan bertujuan untuk mengetahui sampai dimana pencapaian siswa terhadap bahan pengajaran setelah mengalami suatu kegiatan belajar
4. *Embedded –test*, yaitu tes yang dilaksanakan disela-sela atau pada waktu-waktu tertentu selama proses pengajaran berlangsung dan bertujuan untuk mengetes peserta didik secara langsung sesudah suatu unit pengajaran sebelum post-test dan untuk mengecek kemajuan siswa untuk remedial sebelum *post-test*

LATIHAN:

Seorang guru matematika beberapa kali mengadakan ulangan sampai diperoleh nilai akhir pada buku rapornya dengan ranking 3 di kelasnya

- a. Jelaskan antara pekerjaan pengukuran, penilaian, dan evaluasi
- b. Sistem penilaian apa yang tepat dipergunakan untuk mata pelajaran di atas

- **LANGKAH-LANGKAH PENYUSUNAN TES**
- Untuk dapat memperoleh alat penilaian (tes) yang memenuhi persyaratan, setiap penyusun tes hendaknya dapat mengikuti langkah-langkah penyusunan tes. Sax (1980), mengidentifikasi langkah-langkah pengembangan tes ke dalam sembilan langkah sebagai berikut:
 - 1. Menyusun kisi-kisi (tabel spesifikasi) tes, yang memuat: materi pokok yang akan diteskan, aspek perilaku atau tingkatan kognitif yang akan diukur, dan penentuan jumlah butir tes untuk setiap aspeknya.
 - 2. Menulis butir-butir soal dengan mendasarkan pada aspek-aspek yang telah tercantum pada tabel spesifikasi (kisi-kisi) tersebut.
 - 3. Melakukan telaah soal tes (analisis tes secara logis);

4. Melakukan uji coba soal;
5. Analisis soal secara empiris;
6. Memperbaiki atau merevisi tes;
7. Merakit tes, dengan menyiapkan komponen-komponen pendukung untuk penyelenggaraan tes, yang meliputi: (a) buku tes; (b) lembar jawaban tes; (c) kunci jawaban tes; dan (d) pedoman penilaian atau pedoman pemberian skor.
8. Melaksanakan tes; dan
9. Menafsirkan hasil tes.

KONSTRUKSI INSTRUMEN NON TES

1. Informasi tentang hasil belajar peserta didik tidak hanya dapat diperoleh melalui tes, tetapi dapat juga diperoleh melalui pengukuran bukan tes.
2. Alat ukur yang digunakan untuk mengukur perubahan tingkah laku
3. Aspek penting pengukuran tingkah laku adalah aspek pengamatan dan aspek perbuatan
4. Aspek pengamatan adalah kegiatan penilaian dan perbuatan adalah kegiatan yang dinilai
5. Penilaian tes perbuatan berorientasi proses dan produk

Karakteristik Penilaian Proses dan Produk

A. Karakteristik penilaian proses:

1. lebih sulit diukur daripada variabel produk, yaitu dengan cara observasi
2. Pengukuran dapat dilakukan secara obyektif jika hanya mengukur penampilan tunggal (*single performance*)
3. Penilaian proses umumnya cenderung lebih bersifat subyektif

B. Karakteristik Penilaian produk

1. Mudah dilakukan
2. Sering menggunakan standar kemanfaatan, kesesuaian dengan tujuan, nampak luar, tingkat penyimpangan, dan kekuatan

TIPE-TIPE SISTEM PENILAIAN

1. Penilaian oleh instruktur atau guru. Penilaian tipe ini paling banyak digunakan di dalam penilaian perbuatan atau penampilan
2. Penilaian oleh teman. Penilaian mempunyai beberapa keuntungan, a) memperkenalkan peserta didik mengenal kompleksitas, b) mendorong peserta didik dalam melakukan penilaian keterampilan dan usahanya, c) mendorong keterlibatan peserta didik dalam proses belajar mengajar
3. *Self assessment*, penilaian melibatkan peserta didik untuk menilai pekerjaannya, baik dalam proses maupun produk. Penilaian ini akan memberi informasi yang bermanfaat untuk keperluan diagnostik atas kemampuannya

TIPE-TIPE SKALA PENILAIAN

1. **Skala Rangking**, suatu model penilaian dengan cara merangking, baik baik proses maupun produk. Skala rangking merupakan metode penilaian paling dasar dan banyak digunakan dalam sistem penilaian *performance*
2. **Skala Produk**, dikembangkan dengan tujuan untuk melakukan penilaian seperti skala rangking tetapi mendasarkan pada acuan patokan. Untuk mengevaluasi produk dilakukan dengan cara mengkomparasikan antara produk dengan skala yang telah ditetapkan. Model penilaian produk umumnya lebih baik dari pada skala rangking

3. **Check List**, merupakan daftar isian yang berisi sekumpulan perilaku atau aktivitas yang dapat dijadikan pedoman guru selama melakukan observasi dalam suatu kegiatan penilaian. Salah satu bentuk check list yang sangat sederhana adalah dengan cara memberi tanda chek (v) pada kotak-kotak yang tersedia pada setiap skill yang akan diukur
4. **Skala numerik**, skala penilaian numerik ditandai dengan cara memberikan skor angka untuk setiap kategori kualitas tertentu. Skala tersebut berkisar antara 1 sampai 10, dimana 1 berarti sangat jelek dan 10 berabti sempurna. Keuntungan penilaian ini adalah sederhana dan mudah digunakan

KONSTRUKSI BUTIR TES HASIL BELAJAR

Pengertian tes hasil belajar

- TEST** : alat atau prosedur yang dipergunakan dalam rangka pengukuran dan penilaian
- TESTING** : saat dilaksanakannya atau peristiwa berlangsungnya pengukuran dan penilaian
- TESTER** : orang yang melaksanakan tes atau pembuat tes atau eksperimenter
- TESTEE** : orang yang sedang dikenai tes atau pihak yang sedang dikenai percobaan

Dalam pendidikan, apa yang dimaksud dengan TES?

CIRI-CIRI TES HASIL BELAJAR

1. VALIDITAS

Tes hasil belajar dinyatakan valid apabila tes hasil belajar (sebagai alat pengukur keberhasilan peserta didik) dengan secara tepat, benar, sah telah dapat mengukur atau mengungkap hasil-hasil belajar yang telah dicapai oleh peserta didik setelah mereka mempunyai proses belajar mengajar dalam jangka waktu tertentu

2. RELIABILITAS

Apabila skor-skor atau nilai-nilai yang diperoleh para peserta ujian untuk pekerjaan ujiannya adalah stabil, kapan sana, dimana saja, dan oleh siapa saja itu dilaksanakan, diperiksa, dan dinilai

3. OBYEKTIVITAS

Tes hasil belajar yang obyektif apabila tes tersebut disusun dan dilaksanakan menurut apa adanya

Dari segi materi berdasarkan bahan pelajaran yang telah diberikan sesuai atau sejalan dengan tujuan instruksional khusus yang telah diberikan

Dari segi pemberian skor dan penentuan nilai hasil tesnya maupun dalam melakukan koreksi terhindar dari unsur-unsur subyektivitas yang melekat pada diri penyusun tes

4. PRAKTIKABILITAS

Sebuah tes bersifat praktis apabila:

- a. Sederhana, tidak memerlukan peralatan yang banyak atau atau peralatan yang sulit pengadaannya
- b. Tes dilengkapi dengan petunjuk cara mengerjakannya, dilengkapi kunci jawaban dan pedoman skoring dan penentuan nilainya
- c. Tes hasil belajar tidak memakan waktu yang lama, tenaga yang banyak serta biaya yang banyak

PRNSIP-PRINSIP PENYUSUNAN TES HASIL BELAJAR

1. Tes hasil belajar harus mengukur secara jelas hasil belajar yang telah ditetapkan sesuai tujuan instruksional
2. Butir-butir soal tes hasil belajar harus merupakan sampel representatif dari populasi
3. Bentuk tes hasil belajar harus dibuat bervariasi, yang benar-benar cocok untuk mengukur hasil belajar yang diinginkan sesuai dengan tujuan tes
4. Tes hasil belajar harus didisain sesuai dengan kegunaannya untuk memperoleh hasil yang diinginkan
5. Tes hasil belajar harus memiliki reliabilitas yang dapat diandalkan
6. Tes hasil belajar harus dapat dijadikan alat pengukur keberhasilan belajar peserta didik, harus dapat dijadikan alat mencari informasi untuk memperbaiki cara belajar peserta didik dan cara mengajar guuru

BENTUK TES HASIL BELAJAR

Secara garis besar, tes hasil belajar dibagi menjadi dua:

1. Tes hasil belajar bentuk uraian
2. Tes hasil belajar bentuk obyektif

TES BENTUK URAIAN

Adalah tes yang berbentuk pertanyaan tulisan, jawabannya merupakan karangan (*essay*) atau kalimat yang panjang

1. Karakteristik

- a. Memberikan kebebasan kepada *testee*
- b. Hilangnya unsur menebak
- c. Cocok untuk grup kecil
- d. Mudah dalam membuat dan sulit dalam menentukan *scoring*

2. Kebaikan tes bentuk uraian

- a. Pembuatan tes bentuk uraian dapat dilakukan dengan mudah dan cepat
- b. Memberi kebebasan kepada testee dalam menjawab dan mengeluarkan buah pikirannya
- c. Penyusun tes dapat mengetahui seberapa jauh tingkat kedalaman dan tingkat penguasaan testee dalam memahami materi yang ditanyakan
- d. Melatih testee untuk berani mengemukakan pendapat dengan menggunakan kalimat dan gaya bahasa yang merupakan hasil pemikiran sendiri

3. Kelemahan tes bentuk uraian

- a. Kurang dapat menampung tu mencakup luasnya materi yang seharusnya diuji sehingga kurang dapat menilai isi pengetahuan testee yang sebenarnya
- b. Jawaban yang sifatnya heterogen akan menyulitkan dalam mengoreksi jawaban
- c. kecenderungan memberikan skor hasil tes yang lebih subyektif
- d. Koreksi lembar jawaban sulit untuk diserahkan kepada orang lain
- e. Daya ketepatan mengukur dan daya keajegan mengukur kurang dapat diandalkan sebagai alat pengukur hasil belajar yang baik

PENYUSUNAN TES BENTUK URAIAN

Pedoman menyusun butir-butir soal tes bentuk uraian:

1. Soal tes mencakup ide-ide pokok dari materi pelajaran-pelajaran yang sifatnya komprehensif
2. Susunan kalimat soal tidak disalin langsung dari buku pelajaran atau bahan-bahan lain
3. Penyusunan soal dilengkapi dengan kunci jawaban serta pedoman penilaian
4. Penyusunan soal diusahakan agar pertanyaannya bervariasi
5. Soal disusun secara ringkas, padat dan jelas, sehingga dapat dipahami oleh testee
6. Hendaknya dikemukakan pedoman tentang cara menjawab butir soal

TES HASIL BELAJAR BENTUK OBYEKTIF

Tes obyektif atau tes jawaban pendek adalah jenis tes hasil belajar yang terdiri dari butir-butir soal yang dijawab *testee* dengan memilih salah satu jawaban yang dipasangkan atau mengisi jawabannya berupa kata-kata atau simbol-simbol tertentu pada tempat yang telah disediakan

Macam Tes Obyektif

1. Tes obyektif bentuk Benar-Salah (True-False Test)

Adalah butir soal yang terdiri dari pernyataan, yang disertai dengan alternatif jawaban yaitu menyatakan pernyataan tersebut benar atau salah, atau keharusan memilih satu dari dua alternatif jawaban lainnya. Alternatif jawaban itu dapat saja berbentuk benar-salah atau setuju-tidak setuju, baik- tidak tidak

Kebaikan Tes Obyektif Bentuk B-S

1. Perangkat soal dapat mewakili seluruh pokok bahasan
2. Mudah menyusunnya
3. Mudah diskor
4. Alat baik untuk mengukur fakta dan hasil belajar langsung, terutama dengan ingatan
5. Dapat dipergunakan berkali-kali

Kelemahan tes obyektif B-S

1. Mendorong testee untuk menebak jawaban
2. Terlalu menekankan kepada ingatan
3. Meminta respon testee yang berbentuk penilaian absolut
4. Banyak masalah yang tidak dapat dinyatakan hanya dengan dua kemungkinan benar dan salah

Penyusunan tes obyektif B-S

1. Butir soal harus menguji atau mengukur hasilajar testee yang penting dan bermakna
2. Butir soal harus menguji pemahaman, tidak hanya pengukuran terhadap daya ingat
3. Kunci jawaban yang ditentukan haruslah benar
4. Butir soal yang baik haruslah yang secara jelas membedakan testee yang belajar dengan testee yang tidak belajar
5. Butir soal harus dinyatakan secara jelas dan menggunakan bahasa yang baik dan benar

Contoh:

Cara mengolah skor:

$S = R - W$, dimana: S = skor terakhir

R = jumlah soal yang dijawab benar (right)

w = jumlah soal yang dijawab salah (wrong)

TES OBYEKTIF PILIHAN GANDA

Adalah suatu butir soal yang alternatif jawabannya lebih dari dua

Sebutir soal pilihan ganda terdiri dari dua bagian:

1. Pernyataan atau disebut *stem*
2. Alternatif jawaban atau disebut *option*

Kebaikan:

1. Dapat digunakan untuk mengukur segala level tujuan instruksional, mulai dari yang sederhana sampai dengan yang paling kompleks
2. Dapt menggunakan jumlah butir soal yang relatif banyak
3. Penskoran secara obyektif
4. Menuntut kemampuan testee untuk membedakan berbagai tingkat kebenaran

5. Mengurangi keinginan peserta tes untuk menebak karena jumlah option yang disediakan lebih dari dua
6. Informasi yang diberikan lebih kaya

KELEMAHAN

1. Kesukaran dalam mengkonstruksi butir soal, terutama untk menemukan alternatif jawaban yang homogen
2. Kecenderungan dalam mengkonstruksi butir soal dengan hanya menguji aspek ingatan, atau aspek paling rendah dalam ranah kognitif
3. Mempunyai pengaruh yang berarti terhadap hasil tes siswa

PENYUSUNAN TES PILIHAN GANDA

1. Saripati permasalahan harus ditempatkan pada pokok soal (*stem*)
2. Hindari pengulangan kata-kata yang sama dalam pilihan
3. Hindari rumusan kata yang berlebihan
4. Susunan alternatif jawaban dibuat teratur dan sederhana
5. Hindari kata-kata atau istilah aneh
6. Semua pilihan jawaban harus homogen dan diinginkan sebagai jawaban yang benar
7. Hindari jawaban yang benar selalu ditulis lebih panjang dari jawaban yang salah
8. Minimal alternatif jawaban tiga
9. Hindari penggunaan pilihan “semua yang di atas salah”
10. Pokok soal sedapat mungkin dalam pernyataan atau pertanyaan positif

Contoh:

Pulau yang terpadat penduduknya di Indonesia adalah pulau:

- a. Sumatra
- b. Jawa
- c. Kalimantan
- d. Sulawesi

Cara mengolah skor:

$$S = R - W/O - 1$$

Dimana:

S = skor terakhir

R = jawaban yang benar

W = jawaban yang salah

O = banyaknya option

TES OBYEKTIF BENTUK MENJODOHKAN

(matching-test)

Istilah tes menjoodohkan sering dikenal dengan tes mencari pasangan, tes menyesuaikan, atau tes menyesuaikan

Butir soal ditulis dalam dua kolom, kolom pertama adalah pokok soal atau stem sedangkan kolom kedua adalah jawaban

Kebaikan:

1. Baik untuk menguji hasil belajar yang berhubungan dengan pengetahuan tentang istilah, definisi, peristiwa atau penanggalan
2. Dapat menguji kemampuan menghubungkan dua hal baik yang berhubungan langsung maupun tidak langsung
3. Mudah dikonstruksi sehingga diperoleh jumlah butir soal yang cukup untuk menguji satu pokok bahasan tertentu
4. Dapat meliputi seluruh bidang studi yang diuji, perangkat soal yang menggunakan tipe lebih merata dan keseluruhan pokok tipe lebih merata
5. Mudah diskor

• KELEMAHAN

1. Terlalu mengandalkan pada pengujian aspek ingatan
2. Karena mudah disusun, maka sering dijadikan pelarian bagi pengajar apabila tidak sempat untuk membuat tes bentuk lain
3. Karena jawaban yang pendek, maka kurang baik untuk mengevaluasi pengertian dan interpretasi

Penyusunan Tes Obyektif bentuk menjodohkan:

1. Pernyataan pada kolom pertama dan kolom kedua masing-masing haruslah terdiri dari kelompok yang homogen
2. Pernyataan pada kolom kedua harus lebih banyak dari kolom yang pertama
3. Hendaknya diukur sedemikian rupa sehingga kelompok soal maupun kelompok jawabannya berada pada satu halaman kertas
4. Petunjuk cara mengerjakan dibuat singkat dan setegas mungkin

Contoh:

Kolom pertama

(....) 1. Suharto

(....)2. Moh. Hatta

(....)3. Soekarno

(....)4. Sultan Hamengku

Buwono

Kolom kedua

A. Bapak Koperasi Indonesia

B. Bapak Palang Revolusi Indonesi

C. Bapak Pramuka Indonesia

D. Bapak Revolusi Indonesia

E. Bapak Pembangunan

Cara mengolah skor:

$$S = R$$

TES OBYEKTIF BENTUK ISIAN

Tes bentuk isian biasanya berbentuk cerita atau karangan. Kata-kata penting dikosongkan, tugas testee mengisi bagian yang dikosongkan

KEBAIKAN

1. Masalah yang diuji tertuang secara keseluruhan dalam konteksnya
2. Mengungkap pengetahuan testee secara bulat
3. Cara penyusunan item mudah

KELEMAHAN

1. Cenderung lebih banyak aspek pengetahuan
2. Banyak memakan tempat
3. Kurang komprehensif
4. Memberi peluang testee untuk bermain tebak teka

PENYUSUNAN TES BENTUK ISIAN

1. Sebaiknya jawaban yang harus diisi testee dituliskan pada lembar jawaban secara terpisah
2. Ungkapan cerita bahan tes disusun secara ringkas, padat
3. Diusahakan mengungkap selain pengetahuan atau pengenalan
4. Soal dapat dituangkan dalam bentuk gambar, peta dsb

Contoh:

Pada pembuatan busana, terlebih dahulu harus mengambil.....
(1), karena berguna untuk membuat.....(2), kemudian bisa
sebagai pedoman dalam merancang bahan dan harga

Mengolah skor akhir: $S = R$

TES OBYEKTIF BENTUK MELENGKAPI

Tes bentuk melengkapi terdiri dari susunan kalimat yang bagian-bagiannya sudah dihilangkan, bagian yang dihilangkan diganti dengan titik-titik, dan titik dilengkapi oleh testee

KEBAIKAN

1. Sangat mudah penyusunannya
2. Menghemat tempat
3. Persyaratan komprehensif dapat dipenuhi
4. Dapat mengukur berbagai taraf kompetensi

KELEMAHAN

1. Cenderung mengungkap daya ingat atau aspek hafalan saja
2. Butir –butir tes kurang relevan untuk diujikan
3. Tester kurang berhati-hati dalam menyusun kalimat soalnya

Contoh ?

LANGKAH-LANGKAH PENYUSUNAN TES HASIL BELAJAR

Tes hasil belajar akan berarti bila terdiri dari butir-butir soal yang menguji tujuan penting dan mewakili ranah kognitif, afektif dan psikomotor

Langkah-langkah yang perlu dilakukan:

1. Menentukan tujuan mengadakan tes, misalnya: formatif, sumatif, selektif, placement, dsb
2. Pembatasan terhadap bahan yang akan diteskan
3. Memastikan tujuan instruksional yang akan diujikan
4. Menderetkan semua TIK dalam tabel persiapan, yang memuat aspek tingkah laku
5. Memilih butir tes disesuaikan dengan tujuan instruksional

6. Membuat tabel spesifikasi (*blue-print*)

Blue-print atau kisi-kisi adalah sebuah tabel yang memuat tentang perincian materi dan tingkahlaku beserta proporsi yang dikehendaki oleh penilai.

7. Menulis butir soal dengan memperhatikan rambu-rambu penulisan

8. Melakukan uji coba soal

9. Menganalisis butir soal menggunakan statistik atau uji konstruksi

10. Merevisi dan merakit soal tes

11. Memberi label pada soal tes setelah tes memenuhi syarat. Dilengkapi dengan petunjuk cara menggunakan tes

LANGKAH MEMBUAT *BLUE-PRINT*

1. Mencantumkan pokok materi
2. Menentukan persentasenya
3. Menentukan aspek yang diukur dan persentasenya
4. Merinci banyaknya butir soal untuk tiap pokok materi
5. Menentukan aspek yang diukur dan persentasenya
6. Menentukan banyaknya butir soal

Contoh: Spesifikasi Penyusunan Tes Hasil Belajar

Aspek yang diukur	Ingatan (50%)	Pemahaman (30%)	Aplikasi (30%)	Jumlah (100%)
Pokok Mater				
Pengertian busana (14%)				
Fungsi busana (21%)				
Macam-macam busana (36%)				
Persyaratan busana (29%)				

LATIHAN

Buatlah suatu tes obyektif untuk mata pelajaran bidang studi busana dengan mengacu pada kompetensi dasar untuk SMK Busana

1. Tentukan Tujuan instruksional khususnya
2. Buatlah tabel spesifikasinya
3. Buatlah tes obyektif bentuk
 - a. True-false (5soal)
 - b. Multiple-choise (2o butir soal)
 - c. Completion test (10 butir soal)
4. Tentukan rumus perhitungan skornya

TELAAH BUTIR SOAL SECARA EMPIRIS

Soal yang sudah baik secara teoritis atau menurut penilaian ahli, kemudian diujicobakan secara empiris kepada calon peserta tes. Hasil uji coba soal kemudian dianalisis dan ditindak lanjuti untuk mengetahui tingkat kesulitan, daya pembeda, validitas, dan reliabilitas

a. Tingkat Kesulitan

Cara menghitung tingkat kesukaran suatu butir dapat digunakan rumus sebagai berikut:

$$\text{Tingkat kesukaran} = \frac{\text{jumlah peserta tes yang menjawab benar}}{\text{jumlah seluruh peserta tes}}$$

Contoh: jika terdapat 35 peserta tes yang menjawab benar dari 70 peserta yang ikut dalam tes tersebut, maka Indeks Kesukaran butirnya adalah:

$$IK (1) = 35/70 \times 100 = 0,5$$

Tingkat kesulitan butir dikategorikan menjadi 3 yaitu kategori sulit, sedang, dan mudah. Kriteria yang digunakan untuk menginterpretasikan hasil analisis mengacu pada Tabel 6

Tabel 6. Kriteria Indeks Tingkat Kesulitan Butir (b)

Proporsi benar	Kategori
$IK > 0,7$	Mudah
$0,3 \leq IK \leq 0,7$	Sedang
$IK < 0,3$	Sulit

Sumber: Bahrul Hayat (1997: 18)

Berdasarkan rumus tersebut maka dapat dikatakan bahwa tingkat kesukaran suatu butir soal dipengaruhi oleh tingkat kemampuan dari anggota kelompok peserta tes. Hal ini berarti bahwa tingkat kesukaran butir soal tidak semata-mata menunjukkan ukuran kesukaran butir soal, tetapi juga menunjukkan kemampuan rata-rata peserta tes.

2). Daya Pembeda Butir

Daya pembeda butir soal berfungsi untuk melihat kemampuan butir soal dalam membedakan peserta didik antara yang berkemampuan tinggi dengan yang berkemampuan rendah. Cara yang digunakan untuk menentukan daya pembeda butir dapat menggunakan korelasi *Point-Biserial*. Korelasi ini digunakan untuk data yang berbentuk dikotomi yaitu benar = 1 dan salah = 0. Koefisien korelasi (r_{pbis}) merupakan indeks korelasi/hubungan antara skor satu butir tes terhadap skor tes secara keseluruhan. Rumus korelasi Point-Biserial (Hinkle, 1979: 97) yang digunakan adalah:

$$r_{pbis} = \frac{N \sum XY - \sum X \sum Y}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}}$$

Indeks pembeda (D/r_{pbis})	Kategori
$D \geq 0,4$	Sangat baik
$0,3 \leq D \leq 0,39$	Baik, tanpa revisi

PENGEMBANGAN INSTRUMEN PENILAIAN PRAKTEK BUSANA

Emy Budiastuti
Jurusan PTBB FT UNY

• **PENGUKURAN, PENILAIAN & EVALUASI**

- **Mengukur adalah menentukan dimensi kuantitatif berdasarkan suatu standar/ukuran/instrumen yang telah ditetapkan sebelumnya.**
- **Menilai adalah menentukan dimensi kualitatif terhadap suatu hasil pengukuran, berdasarkan kriteria tertentu yang telah ditetapkan.**
- **Pengertian evaluasi dalam konteks evaluasi hasil belajar adalah sama dengan penilaian.**

Ciri-ciri penilaian (pengukuran) dlm bidang pendidikan

- **Pengukuran (penilaian) tidak dapat dilakukan secara langsung, tetapi hanya didasarkan pada indikator-indikator atau gejala-gejala yang nampak. Oleh karena itu, masalah ketepatan alat ukur yang digunakan (valid) menjadi masalah tersendiri**

- Penilaian biasanya didahului dengan kegiatan pengukuran, yg hasilnya berupa skor.
- Hasil penilaian bersifat relatif, artinya hasilnya tidak selalu tetap dari satu waktu ke waktu yang lain, dan sangat tergantung dari banyak faktor: peserta didik, penilai, dan situasi yang terjadi pada saat penilaian berlangsung.
- Dalam kegiatan pengukuran selalu terjadi kesalahan (*error*), yang disebabkan oleh: (1) alat ukurnya (tidak valid dan reliabel); (2) penilai (faktor subyektif, kecenderungan nilai murah atau mahal, kesan pribadi terhadap testee, pengaruh hasil yang lalu, kesalahan menghitung, suasana hati penilai); (3) kondisi fisik dan psikis peserta tes; dan (4) kesalahan akibat suasana ujian (suasana gaduh, pengawasan yang tidak baik dsb).

Prinsip-prinsip penilaian:

- *Terpadu dengan pembelajaran*, yakni menilai apapun yang dikerjakan peserta didik dalam kegiatan belajar mengajar itu dinilai, baik kognitif, psikomotorik dan afektifnya.
- *Objektif*, yakni tidak terpengaruh oleh pertimbangan subjektif penilai.
- *Sistematis*, yakni penilaian dilakukan secara berencana dan bertahap untuk memperoleh gambaran tentang perkembangan belajar peserta didik sebagai hasil kegiatan belajarnya.

Persyaratan Tes :

1. Validitas

- Suatu tes dikatakan valid (sahih) apabila tes tersebut dapat mengukur apa yang seharusnya diukurnya.
- Contoh:

Untuk mengukur tingkat partisipasi peserta didik dalam proses pembelajaran, maka bukan diukur berdasarkan nilai atau prestasi yang diperoleh pada saat mengikuti ujian, akan tetapi akan lebih tepat jika diukur berdasarkan :

- Tingkat kehadirannya
- Terpusatnya perhatian pada pelajaran

- Reliabel berarti dapat dipercaya. Tes dikatakan memiliki reliabilitas, apabila tes tersebut mempunyai sifat dapat dipercaya, dapat memberikan hasil yang tetap (konsisten) apabila diteskan berulang-ulang.

- **Obyektivitas**

Obyektif berarti tidak adanya unsur-unsur pribadi (subyektivitas) yang mempengaruhi. Suatu tes dikatakan memiliki obyektivitas apabila dalam penggunaannya tidak ada faktor subyektif dari pemakainya yang dapat mempengaruhinya, terutama dalam skoring.

Praktikabilitas

- Suatu tes dikatakan memiliki praktikabilitas yang tinggi, apabila tes tersebut bersifat praktis atau mudah digunakan, mudah dlm pemeriksaan dan mudah pula dalam pengadministrasiannya.

Ekonomis

- Yang dimaksud dengan ekonomis adalah bahwa dalam pelaksanaan tes tersebut tidak membutuhkan biaya yang mahal, peralatan yang kompleks dan mahal, tenaga dan waktu yang banyak.

Popham (1995:152) terdapat tiga sumber kesalahan (*sources of error*) dalam *performance assessment*, yaitu:

- *scoring instrument flaws*, instrumen pedoman pensekoran tidak jelas sehingga sukar untuk digunakan oleh penilai, umumnya karena komponen-komponennya sukar untuk diamati (*unobservable*);
- *procedural flaws*, prosedur yang digunakan dalam *performance assessment* tidak baik sehingga juga mempengaruhi hasil pensekoran;
- *teachers personal-bias error*, penskor (*rater*) cenderung sukar menghilangkan masalah personal bias, yakni ada kemungkinan penskor mempunyai masalah *generosity error*, artinya *rater* cenderung memberi nilai yang tinggi-tinggi, walaupun kenyataan yang sebenarnya hasil pekerjaan peserta tes tidak baik atau sebaliknya.

Penilaian Kinerja :

- Penilaian kinerja: dilakukan dengan meminta (menugaskan) kepada peserta tes untuk menampilkan/ mendemonstrasikan pengetahuan dan atau keterampilan yang dikuasainya ke dalam berbagai macam konteks sesuai kriteria yang diinginkan.
- Agar dapat melakukan penilaian kinerja secara obyektif, guru harus dapat merinci kompetensi yang ingin dicapai melalui kegiatan pembelajaran tersebut, yang dikelompokkan menjadi dua komponen utama, yaitu:
 - (1) Aspek Proses
 - (2) Aspek Hasil

- Penilaian aspek proses, meliputi: langkah kerja, sikap kerja, kebenaran pemakaian alat, keselamatan kerja, dan pemilihan bahan.
- Penilaian terhadap komponen produk atau hasil pekerjaan, meliputi: dimensi/ukuran produk, kekuatan, finishing, dan waktu untuk penyelesaian pekerjaan.
- Penilaian keterampilan juga dapat dilakukan terhadap aspek-aspek personal, seperti: kerjasama, inisiatif, tanggung jawab, kreativitas dsb.

EVALUASI PEMBELAJARAN

PENGUKURAN, PENILAIAN & EVALUASI

- Mengukur adalah menentukan dimensi kuantitatif berdasarkan suatu standar/ukuran/instrumen yang telah ditetapkan sebelumnya.
- Menilai adalah menentukan dimensi kualitatif terhadap suatu hasil pengukuran, berdasarkan kriteria tertentu yang telah ditetapkan.
- Pengertian evaluasi dalam konteks evaluasi hasil belajar adalah sama dengan penilaian.

PENGUKURAN, PENILAIAN, dan EVALUASI

PENGUKURAN

Menurut Allen dan Yen pengukuran adalah penetapan angka terhadap suatu obyek atau peristiwa dengan cara yang sistematis

Karakteristik pengukuran:

1. Penggunaan angka atau skala tertentu
2. Menurut suatu aturan atau formula tertentu

Terdapat 4 skala pengukuran, yaitu:

1. **Skala nominal**, yaitu skala yang bersifat kategorial, misal jika sebutir soal dapat dijawab benar maka ia mendapat nilai satu sedangkan jika menjawab salah diberi angka nol

2. **Skala ordinal**, yaitu angka yang menunjukkan adanya urutan, tanpa mempersoalkan jarak antar urutan tersebut

Misal: angka yang menunjukkan rangking siswa dalam suatu mata pelajaran tertentu

3. **Skala interval**, yaitu angka yang menunjukkan adanya jarak yang sama dari angka yang berurutan dan tidak memiliki nol mutlak, pada data ini walaupun datanya nol tetapi masih mempunyai nilai

Misal: nol derajat celcius ternyata masih ada nilainya

4. **Skala rasio**, yaitu angka yang memiliki semua karakteristik angka atau skala yang terdahulu dan memiliki nol mutlak. Jadi kalau data nol tidak mempunyai apa-apa

Misal: tinggi badan seseorang. Jika ada tinggi badan seseorang 0cm, maka tinggi tersebut tidak mempunyai makna

PENILAIAN

Menurut Popham, penilaian adalah usaha formal untuk menentukan status peserta didik menurut ubahan pendidikan yang diinginkan

Menurut Patrix dan Nix, bahwa penilaian mengacu pada penghargaan yang luas yang mencakup bukti dan aspek dari pengetahuan, pemahaman, keterampilan dan atribut peserta didik

Jadi penilaian adalah usaha menentukan keadaan obyek dengan suatu aturan dan dilakukan melalui pengambilan keputusan terhadap sesuatu dengan berpegang pada ukuran baik atau buruk, pandai atau bodoh, sehat atau sakit

Penilaian memerlukan data yang baik mutunya sehingga perlu didukung oleh proses pengukuran yang baik

EVALUASI

Menurut Griffin & Nix, evaluasi adalah *judgment* terhadap nilai atau implikasi dari hasil pengukuran. Menurut definisi ini, kegiatan evaluasi selalu didahului dengan kegiatan pengukuran dan penilaian

Menurut Tyler, evaluasi adalah proses penentuan sejauhmana tujuan pendidikan telah tercapai

Tujuan Penilaian Pembelajaran

Menurut Djemari (2004), terdapat beberapa tujuan yang ingin dicapai dalam melakukan penilaian hasil belajar. Tujuan yang penting adalah untuk:

- 1) mengetahui tingkat kemampuan peserta didik
- 2) mengetahui pertumbuhan dan perkembangan peserta didik
- 3) mendiagnosis kesulitan belajar peserta didik
- 4) mengetahui hasil belajar
- 5) mengetahui pencapaian kurikulum

Fungsi Penilaian Pembelajaran.

Secara khusus bila ditinjau dari segi didaktik, bagi pendidik memiliki fungsi:

- 1) memberikan pedoman untuk mencari dan menemukan jalan keluar bagi peserta didik yang memang memerlukan (fungsi bimbingan),
- 2) memeriksa pada bagian-bagian manakah peserta didik pada umumnya mengalami kesulitan dalam mengikuti proses pembelajaran, selanjutnya dicari jalan keluarnya (fungsi diagnostik)
- 3) memberikan informasi yang sangat berguna untuk mengetahui posisi masing-masing peserta didik ditengah-tengah kelompoknya

4) memberikan bahan yang penting untuk memilih kemudian menetapkan status peserta didik (fungsi selektif)

5) memberikan petunjuk tentang sejauh manakah program pengajaran yang telah ditentukan dapat dicapai (fungsi instruksional)

Ciri-ciri penilaian (pengukuran) dlm bidang pendidikan

- **Pengukuran (penilaian) tidak dapat dilakukan secara langsung, tetapi hanya didasarkan pada indikator-indikator atau gejala-gejala yang nampak. Oleh karena itu, masalah ketepatan alat ukur yang digunakan (valid) menjadi masalah tersendiri**
- **Penilaian biasanya didahului dengan kegiatan pengukuran, yg hasilnya berupa skor.**

- Hasil penilaian bersifat relatif, artinya hasilnya tidak selalu tetap dari satu waktu ke waktu yang lain, dan sangat tergantung dari banyak faktor: peserta didik, penilai, dan situasi yang terjadi pada saat penilaian berlangsung.
- Dalam kegiatan pengukuran selalu terjadi kesalahan (*error*), yang disebabkan oleh: (1) alat ukurnya (tidak valid dan reliabel); (2) penilai (faktor subyektif, kecenderungan nilai murah atau mahal, kesan pribadi terhadap testee, pengaruh hasil yang lalu, kesalahan menghitung, suasana hati penilai); (3) kondisi fisik dan psikis peserta tes; dan (4) kesalahan akibat suasana ujian (suasana gaduh, pengawasan yang tidak baik dsb).

Prinsip-prinsip pengukuran dalam pendidikan

- **Mendidik**, yakni mampu memberikan sumbangan positif terhadap peningkatan pencapaian belajar peserta didik. Hasil penilaian harus dapat memberikan **umpan balik** dan **memotivasi** peserta didik untuk lebih giat belajar.
- **Terbuka/transparan**, yakni prosedur penilaian, kriteria penilaian, dan dasar pengambilan keputusan diketahui oleh pihak yang terkait, termasuk peserta tes.
- **Menyeluruh**, yakni meliputi berbagai aspek kompetensi yang akan dinilai. Penilaian yang menyeluruh meliputi ranah pengetahuan (kognitif), keterampilan (psikomotor), sikap dan nilai (afektif) yang direfleksikan dalam kebiasaan berfikir dan bertindak.

- *Terpadu dengan pembelajaran*, yakni menilai apapun yang dikerjakan peserta didik dalam kegiatan belajar mengajar itu dinilai, baik kognitif, psikomotorik dan afektifnya.
- *Objektif*, yakni tidak terpengaruh oleh pertimbangan subjektif penilai.
- *Sistematis*, yakni penilaian dilakukan secara berencana dan bertahap untuk memperoleh gambaran tentang perkembangan belajar peserta didik sebagai hasil kegiatan belajarnya.

- *Berkesinambungan*, yakni dilakukan secara terus menerus sepanjang berlangsungnya kegiatan pembelajaran
- *Adil (fair)*, yakni tidak ada peserta didik yang diuntungkan atau dirugikan berdasarkan latar belakang sosial-ekonomi, budaya, agama, bahasa, suku, warna kulit, dan gender.
- *Menggunakan acuan kriteria*, yakni menggunakan kriteria tertentu dalam menentukan kelulusan peserta didik (KKM).

Hubungan Mengajar dan Penilaian Pencapaian Hasil Belajar

1. Tujuan Instruksional

Tujuan pendidikan dalam arti sempit berorientasi pada diri peserta didik dari pada diri guru

Manfaat tujuan instruksional:

1. Guru dapat mengevaluasi sejauhmana ketercapaian tujuan instruksional
2. Memberikan arah bagaimana proses belajar mengajar tersebut berlangsung

Tujuan instruksional yang baik dan lengkap mengandung 4 elemen:

- a. Audience* (peserta didik yang belajar)
- b. Behavior* (tingkah laku yang diharapkan)
- c. Condition* (prasyarat yang harus disediakan untuk kegiatan PBM)
- d. Degree* (kriteria penilaian untuk penampilan optimum yang diharapkan dari tujuan instruksional)

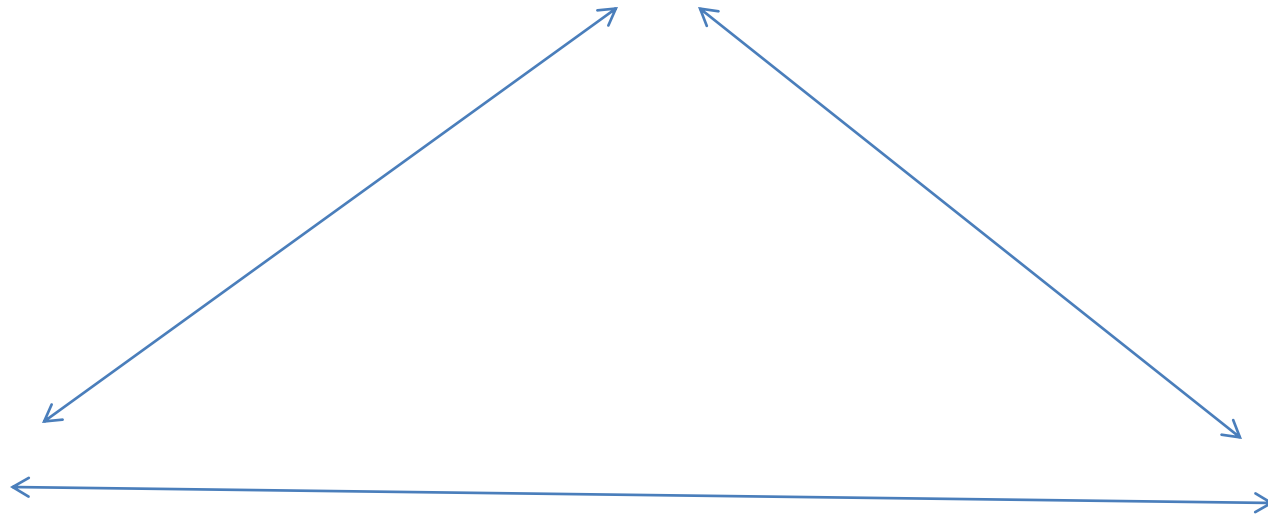
2. Pemilihan Strategi Mengajar

Kendali terbesar dalam pelaksanaan pengajaran adalah pada latar belakang peserta didik, kemampuan dan kemauan yang berbeda. Oleh karena itu guru harus dapat menggunakan pendekatan multi media: menggunakan metode mengajar yang bervariasi dengan media pengajaran yang ragam dan komunikatif

3. Penilaian

Melalui pengukuran dan penilaian, kemampuan peserta didik setelah selesai pengajaran dapat diukur keberhasilan tujuan instruksional dan pengalaman belajarnya.

Tujuan (umun dan khusus)



Pengalaman
belajar

Proses evaluasi

Gambar. Proses Pengajaran

Aspek-aspek Hasil Belajar

Prinsip dasar yang harus diperhatikan dalam mengevaluasi hasil belajar adalah evaluator dituntut untuk mengevaluasi secara menyeluruh terhadap peserta didik baik dari segi pemahaman materi (kognitif), segi penghayatan (afektif), dan pengalamannya (psikomotor)

1. Ranah Kognitif (Benyamin S Bloom, 1956)

- a. Pengetahuan
- b. Pemahaman
- c. Penerapan
- d. Analisis
- e. Sintesis
- f. Evaluasi

2. Ranah Afektif (Krathwohl, dkk 1964)

- a. Kesiediaan menerima
- b. Memberi tanggapan
- c. Menilai
- d. Organisasi
- e. Karakteristisasi

RELIABILITAS UNTUK SKOR SKALA INTERVAL/ESSAY

Rumus koefisien Alfa Cronbach:

$$r_i = \frac{k}{(k-1)} \left\{ 1 - \frac{\sum si^2}{st^2} \right\}$$

dimana: k = mean kuadrat antara subyek

$\sum si^2$ = mean kuadrat kesalahan

S_t^2 = varians total

Rumus varians total dan varians item:

$$S_t^2 = \frac{\sum xt^2}{n} - \frac{(\sum xt)^2}{n^2}$$

$$S_i^2 = \frac{JK_i}{n} - \frac{JK_s}{n^2}$$

dimana: JK_i = jumlah kuadrat seluruh skor item

JK_s = jumlah kuadrat subjek

No	No Item										Xt	Xt ²
	1	2	3	4	5	6	7	8	9	10		
1	1	4	3	3	4	4	3	2	1	3	28	784
2	3	4	3	2	3	3	3	3	2	3	29	841
3	2	3	4	3	3	4	3	4	2	2	30	900
4	4	4	4	4	3	4	3	4	4	4	38	1444
5	4	3	3	4	2	4	3	4	2	1	30	900
6	3	4	4	3	2	4	3	3	3	3	32	1024
7	4	3	3	3	3	4	4	2	3	3	32	1024
8	4	4	4	2	3	3	4	3	3	3	33	1089
9	3	3	3	4	3	3	2	4	4	2	31	961
10	2	4	4	3	3	3	2	3	3	3	30	900
	30	36	35	31	29	36	30	32	27	27	313	9867
	900	1296	1225	961	841	1296	900	1024	729	729	9901	
	100	132	125	107	87	132	94	108	81	79		

$$S_i^2 = \frac{9867}{10} - \frac{(313)^2}{100} = 7,01$$

$$S_t^2 =$$

Multiple-choice test

Cara mengolah skor: $S = R - \frac{W}{O - I}$

S = skor terakhir

R = jawaban yang benar

W = jawaban yang salah

O = banyaknya option

Tes Objektif Bentuk Benar-Salah (*True-False Test*)

Adalah butir soal yang terdiri dari pernyataan yang disertai alternatif jawaban yaitu menyatakan pernyataan tersebut benar atau salah.

Kebaikan:

1. Perangkat soal dapat mewakili seluruh pokok bahasan
2. Mudah penyusunannya
3. Mudah diskor
4. Baik untuk mengukur fakta dan hasil belajar terutama dengan ingatan
5. Dapat digunakan berkali-kali

Kelemahan:

1. Mendorong testee untuk menebak jawaban
2. Terlalu menekankan ingatan
3. Meminta respon testee berbentuk penilaian absolut
4. Banyak masalah yang tidak dapat dinyatakan hanya dengan dua kemungkinan benar atau salah

Penyusunan Tes B-S:

1. Butir soal harus hasil belajar testee yang penting dan bermakna
2. Butir soal harus menguji pemahaman, tidak hanya daya ingat
3. Kunci jawaban yang ditentukan harus benar
4. Butir soal harus dapat secara jelas membedakan testee yang belajar dan tidak belajar
5. Butir soal harus dinyatakan secara jelas dan menggunakan bahasa yang baik dan benar

Cara mengolah skor:

$$S = R - W$$

S = skor terakhir

R = jumlah soal yang dijawab benar (*Right*)

W = jumlah soal yang dijawab salah (*Wrong*)

Tes Objektif Bentuk Menjodohkan (*Matching-test*)

Butir soal ditulis dalam dua kolom. Kolom pertama adalah pokok soal atau steam. Kolom kedua adalah kolom jawaban

Kelebihan:

1. Menguji hasil belajar yang berhubungan dengan pengetahuan : istilah, definisi, peristiwa, penanggulangan
2. Menguji kemampuan yang berhubungan langsung atau tidak langsung
3. Mudah dikonstruksi, sehingga diperoleh sejumlah butir soal yang cukup untuk menguji pokok bahasan tertentu
4. Lebih merata untuk pokok bahasan dan sub pokok bahasan
5. Mudah diskor

Kekurangan/kelemahan:

1. Terlalu mengandalkan aspek ingatan
2. Karena mudah disusun, sering dijadikan pelarian pengajar apabila tidak sempat membuat tes bentuk lain
3. Karena jawaban pendek, kurang baik untuk mengevaluasi pengertian atau interpretasi

Petunjuk penyusunan tes :

1. Kolom pertama dan kedua haruslah terdiri atas kelompok yang homogen
2. Pernyataan kolom kedua lebih banyak dari kolom pertama
3. Kelompok soal dan jawaban berada pada satu halaman
4. Petunjuk mengerjakan dibuat seringkasan dan setegas mungkin

Contoh:

Kolom pertama

1. Soekarno
2. Suharto
3. Moh Hatta
4. Sultan Hamengkubuwono

Kolom kedua

- A. Bapak koperasi Indonesia
- B. Bapak palang merah
- C. Bapak pramuka Indonesia
- D. Bapak Pembangunan
- E. Bapak Revolusi Indonesia

Cara mengolah skor:

$$S = W$$

Tes objektif bentuk isian (*fill in*)

Bentuk tes biasanya berupa cerita atau karangan. Kata-kata penting dalam cerita beberapa dikosongkan, tugas *testee* mengisi bagian yang dikosongkan

Kebaikan:

1. Masalah yang diujikan tertuang keseluruhan
2. Mengungkap pengetahuan *testee* secara bulat mengenai suatu bidang atau hal
3. Cara penyusunan itemnya mudah

Kelemahan:

1. Cenderung lebih banyak aspek pengetahuan
2. Banyak memakan tempat
3. Kurang komprehensif (hanya mengungkap sebagian saja dari bahan yang seharusnya diteskan)
4. Memberi peluang *testee* bermain tebak terka

Petunjuk Penyusunan Tes:

1. Jawaban yang harus diisi testee dituliskan pada lembar jawaban secara terpisah
2. Ungkapan cerita bahan tes disusun secara ringkas dan padat
3. Pengujian soal dapat berupa gambar, peta dan sebagainya sehingga dapat dipersingkat
4. Selain pengetahuan dapat mengungkap kompetensi yang lain

Contoh:

Pada pembuatan busana, terlebih dahulu harus menentukan.....(1), karena berguna untuk membuat(2), kemudian dapat sebagai pedoman dalam merancang bahan dan harga

Cara mengolah skor : $S = R$

Tes Objektif Bentuk Melengkapi (*completion*)

Tes yang terdiri dari susunan kalimat yang bagian-bagiannya sudah dihilangkan diganti titik-titik dan titik-titik ini dilengkapi oleh *testee*

Kebaikan:

1. Mudah penyusunannya
2. Menghemat tempat
3. Bahan yang disajikan cukup banyak
4. Dapat mengukur berbagai taraf kompetensi

Kelemahan:

1. Cenderung mengungkap daya ingat atau hafalan saja
2. Butir-butir tes kurang relevan untuk diujikan
3. Karena penyusunannya mudah, tester sering kurang hati-hati dalam menyusun kalimatnya

Contoh:

1. Sumber ide yang tepat untuk pembuatan busana modifikasi adalah
2. Busana yang dibuat tanpa melakukan penjahitan disebut....

Cara mengolah skor: $S = R$