

# **MATERI EVALUASI PEMBELAJARAN**



Oleh:

**Sri Wening**

**WORKSHOP PENGEMBANGAN PENILAIAN HASIL BELAJAR  
YANG DISELENGGARAKAN PADA TANGGAL 6 – 7 SEPTEMBER 2012**

**FAKULTAS TEKNIK  
UNIVERSITAS NEGERI YOGYAKARTA  
2012**

## A. PENILAIAN HASIL BELAJAR

Kegiatan pengukuran dan penilaian hasil belajar yang dilakukan oleh pendidik untuk mengetahui kemampuan yang dimiliki siswa dan keberhasilan siswa dalam meraih tujuan pembelajaran yang telah ditetapkan. Dalam kehidupan sehari-hari, tes dan pengukuran pada umumnya memang menjadi penentu nasib siswa atau peserta didik. Namun demikian, sebenarnya bukanlah hasil tes atau pengukuran itu sendiri yang menjadi penentu nasib siswa atau peserta tes, akan tetapi interpretasi terhadap hasil pengukuran dan alat pengukuran tersebut. Oleh karena berdasarkan informasi yang sama masing-masing orang akan dapat memberikan interpretasi yang berbeda-beda. Interpretasi yang bermacam-macam inilah yang harus dihindari, karena hal itu akan membahayakan peserta didik. Hal ini akan sama bahayanya dengan apabila informasi yang diberikan tersebut salah akibat dari kesalahan dalam pemakaian tes ataupun pengukuran yang diterapkan. Dari sini dapat disimpulkan betapa pentingnya suatu tes atau pengukuran dalam dunia pendidikan dan pelatihan. Dengan demikian, sebagai pendidik dan pelatih kita tidak dapat mengabaikan pembuatan tes ataupun cara-cara pemakaiannya maupun cara-cara menginterpretasikannya.

Sementara itu, untuk membuat keputusan tentang peserta didik berdasarkan hasil pengukuran dan penilaian, pada dasarnya adalah mudah tetapi sekaligus juga sulit. Secara relatif mudah, jika kita tidak memikirkan akibat-akibat dari keputusan yang kita buat baik terhadap orang lain maupun diri kita sendiri. Namun sebaliknya, hal itu akan terasa sulit manakala kita sadar akan tanggung jawab kita terhadap akibat dari keputusan yang kita buat tersebut.

## B. KLASIFIKASI HASIL BELAJAR

Menurut Bloom, hasil belajar peserta didik dapat dikelompokkan ke dalam 3 (tiga) aspek/ranah, yaitu:

1. Hasil belajar yang berkaitan dengan perkembangan kognitif (pengetahuan).
2. Hasil belajar yang berkaitan dengan perkembangan afektif (sikap dan nilai-nilai).
3. Hasil belajar yang berkaitan dengan perkembangan keterampilan (psikomotorik).  
(Ketiga aspek di atas dapat juga dikelompokkan dengan istilah 3H: Head, Hand, and Heart).

Sementara itu, hasil belajar pada aspek kognitif (pengetahuan), dapat diklasifikasikan ke dalam 6 tingkatan sesuai kompleksitas/tingkatan berpikir, yaitu:

1. Pengetahuan (*knowledge*); mencakup kemampuan dalam mengingat kembali: istilah, fakta-fakta, metode, prosedur, proses, prinsip-prinsip, pola, struktur atau susunan.
2. Pemahaman (*comprehension*); menyangkut kemampuan seseorang dalam: menafsirkan suatu informasi, menentukan implikasi-implikasi, akibat-akibat maupun pengaruh-pengaruh.

3. Aplikasi (*application*); merupakan kemampuan menerapkan abstraksi-abstraksi: hukum, aturan, metoda, prosedur, prinsip, teori yang bersifat umum dalam situasi yang khusus.
4. Analisis (*analysis*); merupakan kemampuan menguraikan informasi ke dalam bagian-bagian, unsur-unsur, sehingga jelas: urutan ide-idenya, hubungan dan interaksi diantara bagian-bagian atau unsur-unsur tersebut,
5. Sintesis (*synthesis*); adalah kemampuan menyusun/memadukan bagian-bagian, unsur-unsur, menjadi struktur atau pola yang baru, yang sebelumnya tidak ada.
6. Evaluasi (*evaluation*); adalah kemampuan untuk menilai ketepatan: teori, prinsip, metoda, prosedur untuk menyelesaikan masalah tertentu.

### C. PRINSIP-PRINSIP PENILAIAN

Penilaian hasil belajar peserta didik harus mendasarkan pada prinsip-prinsip penilaian sebagai berikut.

1. *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.
2. *Terbuka/transparan*, yakni prosedur penilaian, kriteria penilaian, dan dasar pengambilan keputusan diketahui oleh pihak yang terkait, termasuk peserta tes.
3. *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.
4. *Terpadu dengan pembelajaran*, yakni menilai apapun yang dikerjakan peserta didik dalam kegiatan belajar mengajar, baik kognitif, psikomotorik dan afektifnya. Dengan demikian, penilaian tidak hanya dilakukan setelah peserta didik menyelesaikan pokok bahasan tertentu melainkan saat mereka sedang melakukan proses pembelajaran.
5. *Objektif*, yakni tidak terpengaruh oleh pertimbangan atau unsure subjektif penilai.
6. *Sistematis*, yakni penilaian dilakukan secara berencana dan bertahap untuk memperoleh gambaran tentang perkembangan belajar peserta didik sebagai hasil kegiatan belajarnya.
7. *Berkesinambungan*, yakni dilakukan secara terus menerus sepanjang berlangsungnya kegiatan pembelajaran (bukan ad-hoc).
8. *Adil (fair)*, yakni tidak ada peserta didik yang diuntungkan atau dirugikan berdasarkan latar belakang sosial-ekonomi, budaya, agama, bahasa, suku, warna kulit, dan jender.
9. *Menggunakan acuan kriteria*, yakni menggunakan kriteria tertentu dalam menentukan kelulusan peserta didik (KKM).

### D. JENIS PENILAIAN HASIL BELAJAR

Untuk memperoleh data hasil penilaian yang otentik (mampu menggambarkan kompetensi yang sebenarnya), pendidik dianjurkan untuk menerapkan berbagai teknik penilaian secara komplementer (saling melengkapi) sesuai dengan kompetensi yang akan dinilai. Kombinasi penggunaan berbagai teknik penilaian akan memberikan informasi yang lebih akurat tentang kemajuan belajar peserta didik dibanding hanya menggunakan tes sebagai satu-satunya teknik penilaian. Proses pengumpulan informasi oleh pendidik tentang perkembangan dan pencapaian hasil belajar peserta didik melalui berbagai teknik yang mampu mengungkapkan, membuktikan, atau menunjukkan secara tepat bahwa kompetensi tersebut telah benar-benar dikuasai/dicapai, disebut sebagai penilaian kelas atau Penilaian Berbasis Kelas (PBK).

Meskipun penilaian berbasis kelas (PBK) menekankan penggunaan berbagai teknik penilaian sesuai kompetensi yang akan diukur, namun secara garis besar alat penilaian pada dasarnya dapat digolongkan menjadi dua macam, yaitu : (1) Tes; dan (2) Bukan Tes (Non-Tes).

### 1. Teknik Non-Tes

Alat penilaian yang tergolong teknik non-tes antara lain:

- |                                       |   |
|---------------------------------------|---|
| a. Kuesioner/angket                   | f. Portofolio                                       |
| b. Wawancara ( <i>interview</i> )     | g. Jurnal   |
| c. Daftar Cocok ( <i>check-list</i> ) | h. Inventori  |
| d. Pengamatan atau observasi          | i. Penilaian diri ( <i>self-assessment</i> )        |
| e. Penugasan                          | j. Penilaian oleh teman ( <i>peer assessment</i> ). |

### 2. Teknik Tes

Tes adalah sejumlah pertanyaan yang harus dijawab, atau pernyataan-pernyataan yang harus dipilih, ditanggapi, atau tugas-tugas yang harus dilakukan oleh orang yang diuji untuk waktu tertentu, dengan tujuan untuk mengukur suatu kompetensi tertentu dari orang yang diuji tersebut.

Tes merupakan sejumlah pertanyaan yang memiliki jawaban benar atau salah, pertanyaan yang membutuhkan jawaban, pertanyaan yang harus diberikan tanggapan dengan tujuan mengukur tingkat kemampuan seseorang atau mengungkap aspek tertentu dari orang yang dikenai tes. Dengan demikian, setiap tes menuntut keharusan adanya respons dari orang yang dites.

### E. PERSYARATAN TES

Sebuah tes dapat dikatakan baik sebagai alat pengukur, jika memiliki beberapa persyaratan sebagai berikut : (1) validitas; (2) reliabilitas; (3) obyektivitas; (4) praktikabilitas; dan (5) ekonomis.

## 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 :

- a. Tingkat kehadirannya
- b. Terpusatnya perhatian pada pelajaran
- c. Selalu menjawab pertanyaan yang diajukan oleh guru.

Dengan demikian, nilai yang diperoleh pada saat ujian bukan menggambarkan partisipasi peserta didik sebagaimana yang hendak diukur. Atau dengan kata lain, tes yang digunakan untuk mengukur prestasi belajar tersebut tidak tepat (tidak valid) jika digunakan untuk mengukur tingkat partisipasi peserta didik dalam proses pembelajaran.

## 2. Reliabilitas

Reliabel berarti dapat dipercaya. Maka tes dikatakan memiliki reliabilitas, apabila tes tersebut mempunyai sifat dapat dipercaya. Seseorang dikatakan dapat dipercaya apabila orang tersebut selalu bicara ajeg (konsisten), tidak plin-plan atau berubah-ubah pembicaraannya dari waktu ke waktu. Demikian pula halnya, sebuah tes dikatakan dapat dipercaya (reliabel) jika dapat memberikan hasil yang tetap (konsisten) apabila diteskan berulang-ulang.

## 3. Obyektivitas

Dalam pengertian sehari-hari, 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 skoringnya.

Apabila dikaitkan dengan reliabilitas, maka obyektivitas menetapkan keajegan (konsistensi) pada sistem pemberian skor (skoringnya), sedangkan reliabilitas menekankan keajegan pada hasil tes yang diperoleh oleh setiap peserta tes.

Dalam hal ini, terdapat dua faktor yang mempengaruhi obyektivitas suatu tes, yaitu :

### a. *Bentuk tes*

Tes yang berbentuk uraian (esai), akan memberikan banyak kemungkinan kepada si penilai untuk memberikan banyak penilaian (skoring) menurut caranya sendiri. Dengan demikian, maka hasil yang diperoleh seorang peserta tes, akan dapat berbeda apabila dinilai oleh penilai lainnya. Hal ini menunjukkan bahwa dengan menggunakan tes bentuk uraian (bentuk subyektif), akan memungkinkan masuknya unsur subyektivitas dari si penilai dalam melakukan skoring.

### b. *Penilai*

Dengan menggunakan tes bentuk uraian, faktor subyektivitas dari seorang penilai akan dapat masuk secara lebih leluasa dan mempengaruhi dalam pemberian skor (skoring). Faktor-faktor yang dapat mempengaruhi dalam subyektivitas penilaian tersebut antara lain : kesan penilai terhadap peserta tes (*hallo-effect*), tulisan, bahasa, waktu pelaksanaan penilaian dan sebagainya.

#### 4. Praktikabilitas

Suatu tes dikatakan memiliki praktikabilitas yang tinggi, apabila tes tersebut bersifat praktis atau mudah digunakan dan mudah pula dalam pengadministrasiannya. Adapun tes yang praktis adalah tes yang memiliki ciri-ciri sebagai berikut :

- a. Mudah digunakan, misalnya tidak menuntut peralatan yang kompleks, mudah pengawasannya, dan memberikan kemudahan kepada peserta tes untuk mengerjakan terlebih dahulu bagian-bagian yang dianggapnya mudah.
- b. Mudah pemeriksaannya, artinya bahwa tes tersebut dilengkapi dengan kunci jawaban maupun pedoman pemberian skor (skoringnya). Sebagai contoh : tes bentuk obyektif, pemeriksaannya akan lebih mudah dilakukan jika dikerjakan dalam lembar jawaban tersendiri. Untuk saat ini, pemeriksaan tes bentuk obyektif sudah dapat dilakukan dengan bantuan komputer, sehingga dapat dilakukan dalam waktu yang lebih cepat dan akurat.
- c. Tes tersebut dilengkapi dengan petunjuk-petunjuk yang jelas, sehingga dapat diujikan oleh orang lain.

#### 5. 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.

### F. BENTUK PELAKSANAAN TES

Menurut bentuk pelaksanaannya, secara garis besar dikenal tiga bentuk tes, yaitu: (1) tes lisan; (2) tes bentuk perbuatan; dan (3) tes tertulis.

#### 1. Ujian Lisan

Tes ini pada umumnya berbentuk tanya jawab *face to face*. Penilai memberikan pertanyaan (*interview*) langsung kepada peserta tes. Ujian lisan pada umumnya digunakan untuk mengevaluasi hasil belajar dalam bentuk kemampuan dalam mengemukakan ide-ide dan pendapat-pendapatnya secara lisan. Bagi bidang studi yang menuntut keterampilan-keterampilan untuk berbicara atau bernegosiasi dengan orang lain, maka ujian lisan ini dirasa mempunyai kedudukan yang cukup penting. Namun, karena alasan teknis (kepraktisan), ujian lisan ini pada umumnya jarang digunakan untuk melakukan penilaian kompetensi dalam kegiatan pembelajaran yang rutin.

## 2. Ujian Perbuatan

Tes bentuk perbuatan ini pada umumnya dilakukan dengan cara menyuruh peserta tes untuk melakukan sesuatu pekerjaan yang bersifat fisik (praktik). Tes bentuk perbuatan ini sangat cocok untuk melakukan penilaian dalam pelajaran praktik/keterampilan atau praktikum di laboratorium. Alat yang digunakan untuk melakukan penilaian pada umumnya berupa lembar pengamatan (lembar observasi). Tes bentuk perbuatan ini pada umumnya dapat digunakan untuk menilai proses maupun hasil (produk) dari suatu kegiatan praktik.

## 3. Ujian Tertulis (*paper and pencil test*)

Ujian tertulis ini biasanya dilakukan secara berkelompok dengan mengambil tempat di suatu ruangan tertentu. Ujian tertulis, yang awalnya dikenal sebagai *paper and pencil test*, karena adanya perkembangan teknologi komputer, saat ini sudah mulai digantikan oleh tes yang berbasis komputer (*Computer-Based Peserta tesng* atau CBT), yang tidak lagi membutuhkan kertas dan alat tulis. Dalam ujian tertulis dikenal dua bentuk tes, yaitu tes esai (uraian) dan tes obyektif.

### a. Soal Tes Bentuk Uraian (Essai)

Pertanyaan yang diajukan dalam soal tes bentuk esai (uraian) hendaknya benar-benar merupakan soal-soal yang memerlukan pemikiran untuk dapat memberikan jawabannya. Tes ini umumnya memerlukan jawaban yang berbentuk bahasan. Ciri-cirinya selalu diawali dengan kata-kata "Bagaimana, Mengapa, Berikan alasan, Uraikan, Jelaskan, Bandingkan, Simpulkan, Tunjukkan, Bedakan" dan sebagainya. Soal tes bentuk esai ini mempunyai dua bentuk, yaitu esai terbatas dan esai bebas.

Mengingat untuk dapat memberikan jawaban soal tes bentuk esai ini melibatkan tingkat berpikir yang tinggi dan kemampuan berpikir abstrak, maka soal tes ini tentunya belum sesuai untuk digunakan bagi peserta didik di tingkat dasar, seperti: kelas 1, 2, atau 3 SD. Soal tes bentuk esai ini jika disusun dengan baik akan memiliki beberapa keunggulan yang tidak ditemui pada tes obyektif. Keunggulan-keunggulan tes bentuk esai tersebut antara lain :

- 1) Jawaban harus disusun sendiri oleh peserta tes (melatih dalam pemilihan kata-kata dan menyusun kalimat)
- 2) Tidak ada kemungkinan menebak;
- 3) Dapat mengukur kemampuan yang kompleks;
- 4) Dapat digunakan untuk mengembangkan penalaran peserta tes;
- 5) Proses penyusunan soalnya relatif mudah; dan
- 6) Proses berpikir peserta tes dapat dilacak dari jawabannya.

Namun demikian, tes bentuk esai juga mempunyai beberapa kelemahan, antara lain:

- 1) Jumlah soal sangat terbatas, sehingga cakupan materi (validitas isi) lemah;
- 2) Tingkat kebenaran jawaban dan penilaiannya subyektif;
- 3) Jawaban peserta tes kadang tidak relevan dengan pertanyaan;

- 4) Pemeriksaannya sulit, hanya dapat dilakukan oleh penyusunnya;
- 5) Skor tes umumnya kurang reliabel;
- 6) Kualitas jawaban tergantung pada kemampuan peserta tes dalam memilih kata-kata dan menyusun kalimat; dan
- 7) Banyak dijumpai soal-soal tes uraian yang hanya mengungkap pengetahuan yang dangkal.

#### **b. Soal Tes Obyektif**

Soal tes obyektif pada umumnya tepat digunakan untuk mengevaluasi hasil belajar yang berupa kemampuan-kemampuan dalam: mengingat kembali fakta-fakta (*knowledge*), memahami hubungan antara dua hal atau lebih (*comprehension*), dan kemampuan dalam mengaplikasikan prinsip-prinsip (*application*). Pada saat ini, penggunaan tes bentuk obyektif (terutama bentuk pilihan ganda) sudah sangat berkembang, sehingga juga dapat dikembangkan untuk mengukur kemampuan-kemampuan dalam menganalisis, mensintesis dan mengevaluasi.

Namun, soal tes obyektif juga banyak dikritik karena dianggap tidak mampu mengembangkan daya nalar siswa. Beberapa kelemahan tes obyektif antara lain:

- 1) Tidak melatih peserta tes untuk mengemukakan ide-idenya secara tertulis;
- 2) Kemungkinan menebak besar sekali, dan sulit dilacak;
- 3) Memungkinkan untuk saling menyontek;
- 4) Sulit untuk membuat soal yang baik, dan sering hanya mengukur kemampuan yang dangkal;
- 5) Banyak waktu yang tersita untuk membaca soal dan jawabannya.
- 6) Beberapa kemampuan tertentu, seperti: kemampuan dalam mengemukakan pendapat, ide-ide dan sebagainya tak mungkin diukur dengan tes bentuk obyektif.

Namun demikian, soal tes obyektif juga memiliki beberapa keunggulan yang tidak ditemukan pada soal-soal tes bentuk esai. Keunggulan-keunggulan tersebut adalah :

- 1) Jumlah soal banyak, sehingga dapat mencakup semua isi mata pelajaran (representatif → validitas isi baik);
- 2) Penilaiannya mudah (bisa diwakilkan atau dengan alat scanner), dan obyektif;
- 3) Tugas yang harus dilakukan peserta tes jelas, sehingga tidak ada kemungkinan bagi peserta tes untuk mengemukakan hal-hal yang tidak relevan dengan pertanyaan;
- 4) Hasil tes dapat diinformasikan lebih cepat;
- 5) Reliabilitas skor tinggi; dan
- 6) Memungkinkan penyelenggaraan tes bersama pada wilayah yang luas (UNAS, UASBN, UAS, UUB dsb).

#### **c. Bagaimana Menetapkan Bentuk Tes yang Sesuai?**



Tes obyektif mempunyai tujuan yang berbeda dengan tes esai. Oleh karena itu, di antara kedua tes tersebut tidak dapat dibandingkan karena mengemban misi yang berbeda. Kedua tes tersebut selain masing-masing memiliki kelebihan juga memiliki kelemahan.

Dalam hal ini, untuk menetapkan sesuai tidaknya bentuk tes yang digunakan dalam pengukuran (penilaian) akan sangat tergantung pada beberapa pertimbangan antara lain:

- a. Indikator/kompetensi dasar yang akan diukur.
- b. Jumlah peserta tes, bila jumlah peserta tes banyak (seperti UAS, UUB, UASBN atau ujian semester), maka pilihan untuk menggunakan bentuk tes obyektif adalah lebih tepat.
- c. Ruang lingkup materi yang akan diujikan. Untuk tes yang mencakup ruang lingkup materi yang luas, maka pemakaian tes bentuk obyektif dipandang lebih sesuai daripada tes esai, karena tes obyektif akan dapat mencakup materi yang luas.
- d. Tingkat perkembangan kognitif peserta didik. Untuk siswa SD kelas I, II, III, maka tes obyektif lebih sesuai dibanding tes esai yang menuntut tingkat berpikir tinggi.

### **G. 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)**

Untuk mendapatkan tes yang tidak menyimpang dari materi pokok (bahan) ajar serta aspek tingkah laku yang akan diukur melalui tes tersebut, maka perlu dibuat tabel spesifikasi (kisi-kisi). Tabel spesifikasi (kisi-kisi) adalah sebuah tabel yang memuat: rincian

tentang materi pokok yang akan diteskan, perilaku (yang menunjukkan tingkatan kognitif) yang akan diukur melalui tes tersebut dan rincian mengenai jumlah butir tes dari tiap aspek tersebut. Tiap kotak atau sel diisi dengan jumlah butir soal yang akan disusun.

Penentuan jumlah butir soal pada tiap pokok materi yang akan diteskan, jumlah butir soal pada setiap sel didasarkan pada perkiraan (*judgement*) berdasarkan rumusan indikator, keluasan materi serta kedalaman serta karakteristik dari setiap pokok materi yang akan diteskan.

**TABEL SPESIFIKASI (KISI-KISI)**

Materi Pokok:	Indikator:	Aspek yang akan diukur						Bentuk Tes	Jumlah butir
Materi Pokok 1									
Materi Pokok 2									
Materi Pokok 3									

#### H. PENYUSUNAN TES BENTUK URAIAN (ESSAI)

Tes uraian adalah perangkat tes yang butir soalnya mengandung pertanyaan atau tugas yang jawaban atau pengerjaan soalnya harus dilakukan dengan cara mengekspresikan pikiran peserta tes. Ciri khas tes uraian adalah bahwa jawaban soal tidak disediakan oleh penyusun tes, tetapi harus disusun sendiri oleh peserta tes. Peserta tes bebas menjawab pertanyaan yang diajukan. Setiap peserta tes dapat memilih, menghubungkan dan menyampaikan gagasannya dengan menggunakan kata-katanya sendiri.

Mutu soal bentuk uraian sebagai alat untuk mengukur hasil belajar dapat diperoleh dengan cara menyusun soal-soal tes tersebut sesuai dengan kaidah-kaidah yang benar. Adapun kaidah-kaidah penyusunan soal tes bentuk uraian adalah sebagai berikut :

- a. Soal harus sesuai dengan indikator.
- b. Batasan pertanyaan dan jawaban yang diharapkan harus dinyatakan secara jelas.
- c. Tingkat kesulitan materi yang ditanyakan sesuai dengan kemampuan berpikir peserta didik, yang ditunjukkan oleh jenjang pendidikan dan tingkat kelas.
- d. Menggunakan kalimat tanya atau perintah yang menuntut jawaban uraian.
- e. Ada petunjuk yang jelas mengenai cara mengerjakan/cara menjawab soal.

- f. Ada pedoman penskorannya (*scoring*).
- g. Rumusan kalimat soal komunikatif (mudah dipahami peserta tes).
- h. Butir soal menggunakan bahasa Indonesia yang baik dan benar.
- i. Tidak menggunakan kata/ungkapan yang menimbulkan penafsiran ganda atau salah pengertian.
- j. Tidak menggunakan bahasa yang berlaku setempat/tabu.

## I. PENYUSUNAN TES BENTUK OBYEKTIF

Tes bentuk objektif adalah perangkat tes yang butir-butir soalnya mengandung alternatif jawaban yang harus dipilih oleh peserta tes. Alternatif jawaban telah disediakan oleh penyusun butir soal. Dalam hal ini, peserta tes hanya tinggal memilih jawaban yang benar atau paling benar dari alternatif jawaban yang telah disediakan.

Pada dasarnya, ada empat bentuk tes obyektif, yaitu : (1) Bentuk Benar-Salah atau B-S; (2) Bentuk jawaban singkat atau isian singkat; (3) Bentuk menjodohkan; dan (4) bentuk pilihan ganda (*multiple choice*). Namun karena alasan kualitas soal yang dihasilkan dalam mengukur tingkat berpikir siswa, maka bentuk pilihan gandalah yang akhir-akhir ini banyak dikembangkan, sedangkan bentuk tes obyektif lainnya tidak banyak dikembangkan kecuali untuk soal-soal penilaian hasil belajar pada kelas-kelas bawah (kelas 1, 2 dan 3 SD).

### 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 yang disebut OPTIONS. Opsi jawaban, terdiri atas: satu jawaban BENAR, yaitu kunci jawaban, dan beberapa alternatif jawaban yang disebut pengecoh (distraktor).

Tes bentuk pilihan ganda ini memiliki dua tipe, yaitu: (1) tipe pertanyaan langsung (*direct question*), dan (2) tipe pernyataan yang tidak lengkap (*incomplete statement*). Tipe pertanyaan langsung, biasanya diakhiri dengan tanda " ? ", sedangkan tipe pernyataan yang tidak lengkap, biasanya diakhiri dengan tanda " .... ". Bagi siswa yang usianya lebih muda (siswa SD dan SMP), menurut Norman akan lebih menyukai tes pilihan ganda dengan tipe pertanyaan langsung daripada tipe pernyataan tidak lengkap.

### 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.

## J. ANALISIS TES

Analisis butir soal dapat dilakukan baik sebelum soal diujikan maupun sesudahnya. Jika analisis dilakukan sebelum soal diujikan, maka analisis butir soal ditujukan untuk mengkaji seberapa jauh butir-butir soal yang bersangkutan sudah memenuhi persyaratan, baik dari aspek materi, konstruksi maupun segi kebakasaannya. Dengan demikian, jika ada kekurang-tepatan, butir soal tersebut dapat segera diperbaiki.

### 1. Telaah Tes Bentuk Pilihan Ganda

Untuk menganalisis butir soal bentuk pilihan ganda sebelum butir soal tersebut diujikan, maka dapat dilakukan dengan bantuan lembaran telaah, sebagai berikut :

ASPEK	KRITERIA	Nomor Soal			
		1	2	...	n
Materi	1. Butir soal sudah sesuai dengan indikator 2. Hanya ada satu jawaban benar. 3. Penggunaan istilah dari segi keilmuan sudah benar.				

ASPEK	KRITERIA	Nomor Soal			
		1	2	...	n
	4. Pengecoh benar-benar berfungsi. 5. Pengecoh benar-benar homogen dari segi materi keilmuan.				
Konstruk	1. Pokok soal tidak memberi petunjuk ke arah jawaban kunci. 2. Penyertaan grafik, gambar ataupun tabel pada soal benar-benar berfungsi. 3. Tidak menggunakan kata negatif ganda 4. Panjang alternatif pilihan jawaban relatif sama. 5. Untuk soal hitungan, jawaban sudah diurutkan berdasarkan nilainya. 6. Tidak menggunakan alternatif jawaban "tidak ada jawaban yang benar" atau semua benar. 7. Pengecoh benar-benar masuk akal dan tidak terlalu kentara kesalahannya. 8. Pengecoh tidak menggiring ke arah jawaban kunci.				
Bahasa	1. Tidak menggunakan kata-kata atau istilah yang mendua-arti. 2. Kalimat lugas (kalimat efisien) 3. Kalimat informatif / komunikatif (menurut pemahaman peserta tes). 4. Memperhatikan persyaratan ejaan yang disempurnakan. 5. Menggunakan istilah baku (bebas dari istilah lokal).				

Keterangan: V : jika Sesuai  
 -- : jika Belum Sesuai

## 2. Telaah Tes Bentuk Uraian

### a. Aspek Materi:

- 1) Soal harus sesuai dengan indikator.
- 2) Batasan pertanyaan dan jawaban yang diharapkan (ruang lingkup) harus jelas.
- 3) Isi materi yang ditanyakan sudah sesuai dengan jenjang, jenis sekolah, atau tingkat kelas.

### b. Aspek Konstruksi

- 4) Rumusan kalimat soal atau pertanyaan harus menggunakan kata-kata tanya atau perintah yang menuntut jawaban uraian, seperti: mengapa, uraikan, jelaskan, bandingkan, hubungkan, tafsirkan, buktikan, dan hitunglah.
- 5) Ada petunjuk yang jelas tentang cara mengerjakan soal.
- 6) Ada pedoman penskoran, yang berupa rincian dan bobot komponen yang akan dinilai, serta kriteria penskorannya.

- 7) Hal-hal lain yang menyertai tes seperti tabel, gambar, grafik, peta, atau yang sejenisnya harus disajikan dengan jelas dan terbaca.

### c. Aspek Bahasa

- 8) Rumusan kalimat tes harus komunikatif.
- 9) Butir tes menggunakan bahasa Indonesia yang baik dan benar.
- 10) Rumusan tes tidak menggunakan kata-kata/kalimat yang menimbulkan penafsiran ganda atau salah penafsiran.
- 11) Tidak menggunakan bahasa yang berlaku setempat, jika tes akan digunakan untuk daerah lain atau nasional.

### 3. Analisis Secara Empiris

Analisis butir soal juga sangat penting dilakukan setelah butir-butir soal tersebut diujikan. Hal ini bertujuan untuk mencari informasi tentang kualitas setiap butir dalam tes, yang dilakukan dengan cara mengkaji terhadap jawaban-jawaban peserta tes pada setiap butir soal.

Dalam analisis butir soal tes ini biasanya dibahas mengenai empat hal, yaitu :

- a. Indeks kesukaran butir
- b. Daya pembeda atau indeks diskriminasi butir
- c. Korelasi skor butir dengan skor total tes
- d. Analisis distraktor

Pada uraian berikut akan dibahas mengenai tiga hal, yaitu indeks kesukaran butir, daya pembeda atau indeks diskriminasi butir dan analisis distraktor.

#### a. Indeks Kesukaran Butir

Indeks kesukaran butir adalah proporsi dari peserta tes yang menjawab benar pada butir tersebut terhadap jumlah peserta tes secara keseluruhan. Dengan demikian, jika semua peserta tes dapat menjawab dengan benar pada butir tertentu, maka berarti indeks kesukaran butir tersebut adalah 100% atau 1,00. Sedangkan jika semua peserta tes menjawab salah pada butir tersebut, maka berarti indeks kesukaran butirnya adalah 0% atau 0,00. Dalam hal ini, berarti bahwa indeks kesukaran butir ini akan berkisar antara 0,00 sampai dengan 1,00.

Jika  $P$  = Indeks kesukaran butir

$B$  = Jumlah peserta tes yang menjawab benar pada butir tersebut

$N$  = jumlah seluruh peserta tes

Maka : 
$$P = \frac{B}{N} \times 100 \%$$

Pada dasarnya indeks kesukaran butir ( $p$ ) merupakan nilai rata-rata dari kelompok subyek yang dites pada butir tersebut, sehingga indeks kesukaran tersebut juga disebut indeks kesukaran rata-rata.

Indeks kesukaran rata-rata ini paling umum digunakan untuk menentukan taraf kesukaran butir. Indeks kesukaran rata-rata ini sebenarnya lebih tepat jika disebut sebagai indeks kemudahan (*esainess*), karena mempunyai hubungan yang terbalik dengan taraf kesukaran butir. Sebagai contoh, bahwa makin tinggi taraf kesukaran butir, maka akan makin rendah nilai indeksnya, dan demikian pula sebaliknya.

Mengenai berapa besar nilai indeks kesukaran butir tersebut ternyata tidak ada patokan yang pasti. Akan tetapi, sebagai ilustrasi dapat diberikan klasifikasi sebagai berikut :

Indeks Kesukaran	Klasifikasi Butir
> 0,90	Butir terlalu mudah
0,70 – 0,90	Mudah
0,83 – 0,69	Agak sukar
< 0,30	Sangat sukar

#### b. Indeks Diskriminasi (Daya Pembeda)

Indeks diskriminasi atau daya pembeda butir adalah kemampuan butir tersebut dalam membedakan antara peserta tes yang berkemampuan rendah dengan peserta tes yang berkemampuan tinggi. Butir tes dikatakan mempunyai daya pembeda yang tinggi apabila dijawab benar oleh semua atau sebagian besar peserta tes dari kelompok berkemampuan tinggi, dan tidak dapat dijawab dengan benar oleh semua atau sebagian besar peserta tes dari kelompok berkemampuan rendah. Makin besar perbedaan proporsi antara penjawab benar dari kelompok tinggi dengan penjawab benar dari kelompok rendah, maka berarti akan makin tinggi pula daya pembeda dari butir tersebut.

Daya beda butir tes dinyatakan dengan  $d$ , dan besarnya dapat dicari dengan rumus sebagai berikut :

$$d = \frac{n_T}{N_T} - \frac{n_R}{N_R}$$

Dalam hal ini:

$d$  = daya pembeda butir atau indeks diskriminasi butir

$n_T$  = jumlah penjawab benar pada kelompok tinggi

$N_T$  = jumlah peserta tes dari kelompok berkemampuan tinggi

$n_R$  = jumlah penjawab benar pada kelompok rendah

$N_R$  = jumlah peserta tes dari kelompok berkemampuan rendah

Untuk menghitung indeks diskriminatif atau daya beda butir tes maka harus disusun terlebih dahulu berdasarkan ranking hasil skor keseluruhan tes, dari skor yang tertinggi sampai dengan skor yang terendah. Dalam prakteknya, perhitungan tersebut dilakukan sebagai berikut :

- 1) Untuk jumlah peserta tes  $N > 90$  orang, maka diambil 27% N sebagai kelompok atas dan 27% N sebagai kelompok bawah. Sisanya yang di tengah-tengah, yakni sebanyak 46% tidak diikutsertakan dalam perhitungan.
- 2) Jika jumlah peserta tes  $< 90$  orang, maka diambil 50% N sebagai kelompok atas dan 50% sebagai kelompok bawah. Hal ini berarti bahwa semua peserta tes diikuti dalam perhitungan.

Nilai indeks diskriminasi atau daya pembeda ( $d$ ) ini terletak dalam rentang antara -1,00 sampai dengan +1,00. Nilai positif menunjukkan bahwa proporsi siswa pada kelompok tinggi yang menjawab benar adalah lebih besar daripada proporsi siswa pada kelompok rendah. Demikian pula sebaliknya, nilai negatif menunjukkan bahwa proporsi siswa pada kelompok tinggi yang dapat menjawab benar adalah lebih rendah daripada proporsi siswa dari kelompok rendah.

Cara perhitungan indeks diskriminasi sebagaimana diuraikan di atas merupakan cara perhitungan yang paling banyak dilakukan, karena cara perhitungannya mudah dan penafsirannya dapat dilakukan secara langsung. Selain cara perhitungan di atas, indeks diskriminasi atau daya pembeda butir tes dapat pula ditentukan berdasarkan analisis korelasi sebagaimana dilakukan untuk mencari validitas butir.

Butir tes yang mempunyai nilai  $d$  negatif jelas sangat tidak dikehendaki dan harus diganti. Indeks kesukaran butir ternyata mempunyai kaitan yang erat dengan indeks diskriminasinya. Dalam arti butir-butir tes yang mempunyai taraf kesukaran menengah (0,25 – 0,75) akan mempunyai daya pembeda yang baik. Butir-butir yang mempunyai indeks kesukaran terlalu tinggi atau terlalu rendah, tidak akan mempunyai daya pembeda yang baik.

Berdasarkan pengalaman praktis, Ebel (1965) memberikan pedoman dalam penafsiran nilai  $d$  sebagai berikut :

Daya Pembeda ( $d$ )	Kriteria Butir
$> 0,40$	Butir sangat baik, dan dapat berfungsi dengan baik.
$0,30 - 0,39$	Butir memerlukan revisi kecil, atau tidak sama sekali.
$0,20 - 0,29$	Butir berada dalam batas diterima dan disisihkan, sehingga memerlukan revisi.
$< 0,19$	Butir jelek, harus disisihkan atau direvisi secara total.

### c. Analisis Keberfungsian Distraktor

Tujuan dari analisis butir antara lain adalah untuk mengetahui mengenai distribusi jawaban dari peserta tes di antara alternatif jawaban yang tersedia. Dengan demikian, akan dapat diketahui apakah semua alternatif jawaban yang disediakan tersebut ada subyek pemilihnya, atautkah pilihan jawaban subyek hanya mengelompokkan pada beberapa alternatif jawaban saja, sehingga alternatif jawaban yang lain sama sekali tidak



ada pemilihnya. Dari sini juga akan dapat diketahui berapa banyak subyek yang memilih kunci jawaban dan berapa banyak subyek yang memilih distraktor atau pengecoh.

Berdasarkan penyelidikan terhadap distribusi jawaban dari suatu butir tersebut, akan ditentukan :

1. Berapa banyak subyek yang memilih jawaban kunci (subyek yang jawabannya benar)
2. Distraktor atau pengecoh mana yang bagi sebagian besar subyek terlalu menyokong kesalahannya, sehingga tidak ada subyek yang memilih jawaban yang benar.
3. Distraktor mana yang merupakan distraktor yang menyesatkan.
4. Distraktor mana yang mempunyai daya tarik bagi subyek yang kurang pandai, dan tidak atau kurang mempunyai daya tarik bagi subyek yang pandai.

Untuk dapat mengetahui efektivitas dari masing-masing distraktor tersebut perlu ditentukan proporsi yang pasti dari responden yang memilih masing-masing distraktor, baik dari kelompok atas maupun bawah. Hal ini dapat dilakukan dengan cara menghitung jumlah peserta tes baik dari kelompok atas maupun kelompok bawah yang memilih pada masing-masing alternatif jawaban.

Dalam kaitannya dengan analisis distraktor ini, maka distraktor atau jawaban bukan kunci dapat dianggap fungsional atau efektif sebagai penggiur atau pengecoh apabila :

1. Dipilih oleh minimal 5% dari masing peserta tes; artinya bahwa distraktor yang dibuat tersebut mengundang respon atau menggiur bagi peserta tes.
2. Mempunyai daya tarik bagi subyek yang kurang pandai, dan tidak atau kurang mempunyai daya tarik bagi subyek yang pandai. Dengan demikian, diharapkan bahwa jumlah pemilih dari masing-masing distraktor tersebut sebagian besar adalah kelompok bawah (kurang pandai), dan hanya sebagian kecil saja yang berasal dari kelompok atas (kelompok pandai).
3. Tidak bersifat menyesatkan, baik bagi kelompok bawah dan terlebih lagi bagi kelompok atas. Artinya bahwa jumlah pemilih pada masing-masing distraktor hendaknya tidak melebihi jumlah pemilih pada jawaban kunci (terutama bagi kelompok atas).
4. Masing-masing distraktor hendaknya mengandung daya tarik (*response-elicitation*) yang kurang lebih seimbang.

## **K. ACUAN PENILAIAN**

Dilihat dari penafsiran hasil tes, pengukuran dalam bidang pendidikan dapat didasarkan pada acuan norma/relatif atau acuan kriteria/patokan. Kedua acuan tersebut menggunakan asumsi yang berbeda tentang kemampuan seseorang. Penafsiran hasil tes antara kedua acuan itu juga berbeda, sehingga menghasilkan informasi yang berbeda maknanya. Pemilihan acuan ditentukan oleh karakteristik mata pelajaran yang akan diukur dan tujuan yang akan dicapai.

Penilaian acuan norma berasumsi bahwa kemampuan orang berbeda dan dapat digambarkan menurut distribusi normal. Perbedaan itu harus ditunjukkan oleh hasil

pengukuran, misalnya setelah mengikuti pembelajaran selama satu semester, peserta didik dites. Hasil tes seorang peserta didik dibandingkan dengan kelompoknya, sehingga dapat diketahui posisi (rangking) peserta didik tersebut di kelas itu.

Penilaian acuan kriteria berasumsi bahwa hampir semua orang dapat belajar apa saja, meskipun dengan waktu yang berbeda. Dalam acuan kriteria, penafsiran skor hasil tes selalu dibandingkan dengan kriteria yang telah ditetapkan. Bagi peserta didik yang telah mencapai kriteria yang telah ditetapkan (standar) dinyatakan lulus (tuntas), dan diberi pelajaran tambahan yang biasa disebut pengayaan, sedangkan bagi peserta didik yang belum mencapai standar diberi remedi sampai mencapai kriteria ketuntasan.

Ketuntasan belajar untuk setiap indikator ditetapkan dalam suatu kompetensi dasar atau kriteria ketuntasan minimal (KKM). Satuan pendidikan harus menentukan kriteria ketuntasan minimal dengan mempertimbangkan tingkat kemampuan rata-rata peserta didik dan pendukung penyelenggaraan pembelajaran.

#### **L. PENILAIAN PEMBELAJARAN KETERAMPILAN**

Hal yang penting dalam pembelajaran keterampilan adalah diperolehnya penguasaan keterampilan praktis, serta pengetahuan dan perilaku yang berhubungan langsung dengan keterampilan tersebut. Sehubungan dengan itu, maka para ahli telah mengembangkan berbagai metode pembelajaran keterampilan yang berbeda-beda, tergantung pada sasaran atau maksud yang hendak dicapai di dalam pembelajaran tersebut.

Model yang sederhana untuk pembelajaran keterampilan kerja adalah metode empat tahap TWI (*Training Within Industry*). Tahap-tahap tersebut meliputi :

a. Persiapan

Dalam hal ini, pendidik atau instruktur mengutarakan sasaran-sasaran latihan kerja, menjelaskan arti pentingnya latihan, membangkitkan minat para peserta pelatihan (peserta didik) untuk menerapkan pengetahuan yang sudah dimilikinya dalam situasi yang riil.

b. Peragaan

Pada tahap ini, instruktur memperagakan keterampilan yang dipelajari oleh peserta didik, menjelaskan cara kerja dan proses kerja yang benar. Dalam hal ini, instruktur harus mengambil posisi sedemikian rupa sehingga para peserta pelatihan akan dapat mengikuti demonstrasi mengenai proses kerja dengan baik.

c. Peniruan

Pada tahap ini, peserta pelatihan menirukan aktivitas kerja yang telah diperagakan oleh instruktur. Dalam hal ini, instruktur mengamati peniruan yang dilakukan oleh peserta pelatihan, menyuruh melakukannya secara berulang-ulang dan membantu serta mendorong para peserta pelatihan agar dapat melakukan pekerjaannya dengan benar.

d. **Praktik**

Setelah instruktur yakin bahwa peserta pelatihan telah dapat melakukan tugas pekerjaan dengan cara kerja yang benar, maka selanjutnya instruktur memberikan tugas kepada peserta pelatihan untuk melakukan tugas pekerjaannya. Dalam hal ini, peserta pelatihan mengulangi aktivitas kerja yang baru saja dipelajarinya sampai keterampilan tersebut dapat dikuasai sepenuhnya. Instruktur melakukan pengamatan untuk melakukan penilaian baik terhadap aktivitas atau cara kerja peserta pelatihan maupun hasil-hasil pekerjaan atau produk yang dihasilkannya.

Metode empat tahap ini mempunyai keterbatasan, karena hanya cocok untuk pembelajaran keterampilan yang bertujuan membuat barang (fabrikasi), sedangkan pembelajaran keterampilan yang memiliki karakteristik yang berbeda (seperti: *trouble shooting*, layanan/jasa) tidak tepat menggunakan langkah-langkah pembelajaran tersebut.

### **1. Karakteristik Pembelajaran Keterampilan**

Skill atau keterampilan merupakan integrasi dari gerakan-gerakan atau perbuatan-perbuatan yang teratur dengan baik dan dapat menyesuaikan dengan kondisi-kondisi yang bervariasi dan kompleks untuk mencapai hasil sesuai yang diinginkan. Skill adalah kemampuan menghasilkan, membuat, melakukan gerakan yang dapat diamati yang dilakukan oleh seseorang pada waktu melakukan tugasnya. Skill adalah kemampuan praktis atau perubahan perilaku yang dihasilkan dari latihan dan pengalaman.

Di dalam mempelajari skill, petunjuk-petunjuk praktis adalah jauh lebih baik daripada mencoba-coba (*trial and error*). Skill harus dipelajari/dibentuk dengan berpraktek atau berlatih (*by training, by doing*), walaupun juga diperlukan adanya pengetahuan.

Syarat dasar dalam mempelajari skill adalah adanya umpan balik atau (*feedback*). Feedback menekankan pada kegunaannya sebagai informasi yang berguna untuk perbaikan-perbaikan atau perubahan-perubahan. Dengan demikian, feedback adalah informasi yang disediakan bagi siswa yang memungkinkan membandingkan hasil belajar skill yang dicapai dengan suatu standar hasil belajar yang telah ditetapkan.

Hampir semua macam skill atau hampir setiap skill mempunyai dasar *knowledge* dan biasanya akan paling efisien dipelajari bila peserta didik menguasai pengetahuan yang relevan sebelum mempraktekannya.

Untuk melaksanakan suatu tugas memerlukan pengetahuan dan skill. Pengetahuan tersebut meliputi :

- a. Pengetahuan mengenai situasi pekerjaan dan faktor-faktor keselamatan dan kesehatan kerja.
- b. Pengetahuan tentang material (bahan-bahan), alat-alat, mesin-mesin.
- c. Pengetahuan mengenai standar kinerja, baik kualitas hasil pekerjaan maupun produktivitas, dan bagaimana standar tersebut harus dicapai.

Di samping itu, pengaruh kemampuan skill yang telah dimiliki (*entering behavior*) dalam mempelajari skill yang baru sangat penting. Fitts menyatakan bahwa mempelajari skill baru harus dibentuk oleh atau didasarkan pada skill yang lama yang telah dimilikinya. Belajar skill baru pada umumnya terdiri dari koordinasi skill-skill lama yang kurang kompleks yang telah dikuasai seseorang ke dalam pola yang baru tertentu. Pada tingkat permulaan/dasar, metodenya harus benar, karena jika salah pada tingkat selanjutnya akan sukar dibetulkan. Pada tingkat selanjutnya, dengan tetap mempertahankan cara/metode yang benar dan ketelitian, mulailah berlatih menekankan pada kecepatan/kuantitas.

## 2. Penilaian Keterampilan

Untuk menilai kemampuan skill yang dimiliki oleh seseorang, maka hanya ada satu bentuk tes yang paling tepat, yaitu tes perbuatan (*performance test*). Artinya orang yang akan dinilai kemampuan skillnya harus menampilkan atau melakukan skill yang dimilikinya di bawah persyaratan-persyaratan kerja yang berlaku.

Di industri, standard kinerja tersebut dikenal dengan standar pekerja berpengalaman atau *experienced worker's standar* (EWS), dan standar kinerja tersebut ditentukan berdasarkan tingkatan kuantitas dan kualitas hasil pekerjaan dari pekerja berpengalaman.

Agar penilaian *skill* dapat dilakukan secara obyektif, maka skill tersebut harus dirinci ke dalam elemen-elemen skill yang masing-masing dapat dinilai atau diukur sesuai dengan indikatornya.

Sementara itu, standar hasil yang dikehendaki, dalam hal ini dapat diukur berdasarkan:

- a. Waktu yang diperlukan untuk menyelesaikan job, yang dapat dihitung berdasarkan waktu rata-rata yang digunakan oleh sejumlah siswa untuk menyelesaikan job yang sama.
- b. Kualitas
  - 1) Kecocokan terhadap kebutuhan
  - 2) Kekuatan produk
  - 3) Penyelesaian, kerapihan, kebersihan, nampak, dan sebagainya
  - 4) Untuk skill pembuatan barang-barang yang terukur, maka dimensi atau ukuran menjadi faktor yang penting untuk diperhatikan.

Berkaitan dengan uraian di atas, maka penilaian keterampilan (*skill*) pada dasarnya dapat dilakukan dengan dua cara, yaitu :

- a. Penilaian secara subyektif; dan
- b. Penilaian secara obyektif

Penilaian secara subyektif adalah penilaian yang didasarkan atas kesan umum guru terhadap hasil kerja siswanya, yang biasanya dibandingkan dengan hasil kerjanya

guru sendiri atau hasil kerja berdasarkan standar industri. Sementara itu, penilaian secara obyektif didasarkan atas skema penilaian yang memuat semua komponen skill yang perlu dinilai. Elemen-elemen skill dapat dikelompokkan menjadi dua, yaitu aspek skill (keterampilan) dan aspek personal.

Aspek skill (keterampilan), meliputi :

1) Kualitas penyelesaian job

Dalam hal ini, yang dinilai adalah kesempurnaan, nampak, kelurusan, kesikuan, kerataan, kekuatan, ukuran dan sebagainya.

2) Keterampilan dalam menggunakan alat dengan benar.

Sebagai indikasinya adalah efisien pemakaian, kebenaran pemakaian, kelayakan, pemeliharaan keselamatan kerja dan sebagainya.

3) Kecepatan melaksanakan pekerjaan (produktivitas)

4) Kemampuan mengambil keputusan-keputusan berdasarkan aplikasi informasi yang diberikan.

Aspek personal, meliputi : sikap kerja, usaha/inisiatif dan sebagainya.

Dengan demikian, secara garis besar penilaian pembelajaran keterampilan pada dasarnya dapat dilakukan terhadap dua hal, yaitu : (1) proses pelaksanaan pekerjaan, yang mencakup : langkah kerja dan aspek personal; dan (2) produk atau hasil pekerjaan.

Penilaian terhadap aspek proses umumnya lebih sulit dibanding penilaian terhadap produk atau hasil kerja. Penilaian proses hanya dapat dilakukan dengan cara pengamatan (observasi), dan dilakukan seorang demi seorang. Penilaian proses pada umumnya cenderung lebih subyektif dibanding penilaian produk, karena tidak ada standar yang baku. Namun demikian, penilai dapat lebih meningkatkan obyektivitas penilaiannya dengan cara analisis tugas (*analisis skill*).

Sementara itu, penilaian produk pada umumnya lebih mudah dilakukan daripada penilaian proses, karena dapat dilakukan dengan menggunakan instrumen yang lebih valid dan reliabel, seperti alat-alat ukur mikrometer, meteran dan sebagainya. Dalam penilaian produk, karakteristik yang digunakan sebagai standar biasanya adalah berhubungan dengan kemanfaatan, kesesuaian dengan tujuan, dimensi, nampak luar, tingkat penyimpangan, kekuatan dan sebagainya.

## DAFTAR PUSTAKA

- Badan Standar Nasional Pendidikan (2006). *Panduan Penyusunan Kurikulum Tingkat Satuan Pendidikan Jenjang Pendidikan Dasar dan Menengah*. Jakarta: BSNP.
- Erickson, R.C. & Wentling, T.L. (1988), *Measuring Student Growth : Techniques and Procedures for Occupational Education*. Illinois : Griffon Press.
- Grondlund, N.E. (1982). *Constructing Achievement Test* (3rd. ed.). Englewood Cliffs, NJ: Prentice Hall Inc.

•  
•  
• *Peraturan Pemerintah Nomor 19 Tahun 2005, tentang Standar Nasional Pendidikan.* Jakarta: Setjen Depdiknas.

*Peraturan Menteri Pendidikan Nasional Nomor 20 Tahun 2007, tentang Standar Penilaian Pendidikan.* Jakarta: Setjen Depdiknas.

•  
• Popham, W.J. (1995). *Classroom assessment: What teachers need to know.* Boston, MA: Allyn & Bacon, Inc.

•  
*Undang-undang Nomor 20 Tahun 2003, tentang Sistem Pendidikan Nasional.* Jakarta: Setjen Depdiknas.

# **EVALUASI HASIL BELAJAR**



**Oleh :**

**Sri Wening**

**Makalah Disampaikan Pada Pembekalan Pemenuhan  
Guru Produktif SMK, Tanggal 23 Juli 2013**

**FAKULTAS TEKNIK  
UNIVERSITAS NEGERI YOGYAKARTA  
2013**



KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN  
UNIVERSITAS NEGERI YOGYAKARTA  
FAKULTAS TEKNIK

Alamat : Kampus Karangmalang Yogyakarta, 55281 Telp. (0274) 586168 psw 276.289 (0274) 586734 Fax. (0274) 586734  
Website : <http://ft.uny.ac.id> e-mail : [ft@uny.ac.id](mailto:ft@uny.ac.id) ; [teknik@uny.ac.id](mailto:teknik@uny.ac.id)



Certificate No: QSC 00002

# Sertifikat

No : 1023/UN34.15/PP/2012

Diberikan Kepada :

Dr. Sri Wening

Sebagai Tutor

pada kegiatan

“WORKSHOP PENGEMBANGAN PENILAIAN HASIL BELAJAR”

Yang Diselenggarakan Oleh Fakultas Teknik Universitas Negeri Yogyakarta pada tanggal 6 sampai dengan 7 September 2012  
di Gedung KPLT Fakultas Teknik - UNY.

Yogyakarta, 19 September 2012  
Dekan,

Dr. Moch Bruri Triyono  
NIP.19560216 198603 1 003