

# PENILAIAN HASIL BELAJAR



*Dr. Syahrudin, S.Pd., M.Pd.*



### **Sanksi Pelanggaran Hak Cipta**

Undang-Undang Republik Indonesia No. 28 Tahun 2014 tentang Hak Cipta

#### **Lingkup Hak Cipta**

##### **Pasal 1 :**

Hak Cipta adalah hak eksklusif pencipta yang timbul secara otomatis berdasarkan prinsip deklaratif setelah suatu ciptaan diwujudkan dalam bentuk nyata tanpa mengurangi pembatasan sesuai dengan ketentuan peraturan perundang-undangan.

#### **Ketentuan Pidana**

##### **Pasal 113 :**

1. Setiap Orang yang dengan tanpa hak melakukan pelanggaran hak ekonomi sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf i untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 1 (satu) tahun dan/atau pidana denda paling banyak Rp100.000.000,00 (seratus juta rupiah).
2. Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf c, huruf d, huruf f, dan/atau huruf h untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 3 (tiga) tahun dan/atau pidana denda paling banyak Rp500.000.000,00 (lima ratus juta rupiah).
3. Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf a, huruf b, huruf e, dan/atau huruf g untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 4 (empat) tahun dan/atau pidana denda paling banyak Rp1.000.000.000,00 (satu miliar rupiah).
4. Setiap Orang yang memenuhi unsur sebagaimana dimaksud pada ayat (3) yang dilakukan dalam bentuk pembajakan, dipidana dengan pidana penjara paling lama 10 (sepuluh) tahun dan/atau pidana denda paling banyak Rp4.000.000.000,00 (empat miliar rupiah).

## **PENILAIAN HASIL BELAJAR**

©Dr. Syahrudin, S.Pd., M.Pd, 2023

Diterbitkan Pertama kali oleh

Penerbit Permata Ilmu (Anggota IKAPI), 2023

Jln. Makassar IV Blok A/93 Bumi Sudiang Permai (BSP), Makassar

Tlp: 085396419243

E-mail: penerbitpermatailmu@gmail.com

Instagram: penerbitpermatailmu

website: www.permatailmu.com

### **Penulis**

Dr. Syahrudin, S.Pd., M.d

### **Editor**

Dr. Sakaria, S.S., S.Pd., M.Pd.

Dr. Abdul Haliq, S.Pd., M.Pd.

### **Desain Sampul**

Idho' Lontaral

### **Layout dan Tata Letak**

Ridwan

Perpustakaan Nasional RI: Data Katalog Dalam Terbitan (KDT)

### **PENILAIAN HASIL BELAJAR**

Makassar: Penerbit Permata Ilmu, 2023

vi+132; 14,5 X 21 cm

ISBN: 978-623-98291-4-8

# KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT karena atas izin-Nya penulis dapat menyajikan buku ini kepada para mahasiswa serta pembaca yang menaruh minat terhadap dunia pendidikan.

Tujuan dari penulisan buku ini ialah untuk melengkapi bahan bacaan mahasiswa dalam mengikuti mata kuliah Penilaian Hasil Belajar. Jadi buku ini bukan dimaksudkan sebagai satu-satunya pegangan yang harus dipergunakan dalam mata kuliah *Penilaian Hasil Belajar*. Di samping membaca buku ini, mahasiswa masih perlu untuk membaca literatur lain tentang penilaian hasil belajar.

Buku ini isinya memuat penjelasan mengenai prinsip-prinsip dasar penilaian dalam pendidikan, petunjuk praktis penyusunan soal, teori validitas dan reliabilitas, analisis hasil tes serta analisis soal.

Kami menyampaikan terima kasih kepada semua pihak yang telah membantu dalam penerbitan buku ini. Dan kami menyadari bahwa buku ini masih banyak kelemahan-kelemahan dan kekurangan-kekurangan. Oleh karena itu, kritik dan saran untuk perbaikan dan kesempurnaan buku ini sangat diharapkan dari pembaca.

Penulis,

# DAFTAR ISI

KATA PENGANTAR

DAFTAR ISI

BAB I PENDAHULUAN

- A. Pengertian Pengukuran, Penilaian dan Evaluasi
- B. Penilaian Pendidikan
- C. Makna Penilaian
- D. Fungsi Penilaian

BAB II ALAT PENILAIAN

- A. Tes
- B. Nontes
- C. Sasaran Tes

BAB III PENYUSUNAN TES HASIL BELAJAR

- A. Kisi Tes
- B. Klasifikasi Bentuk Soal

BAB IV KONSTRUKSI ALAT PENILAIAN

- A. Dasar-dasar Penyusunan Tes Hasil Belajar
- B. Konstruksi Soal Uraian/Esei
- C. Konstruksi Soal Objektif

BAB V PENGADMINISTRASIAN TES

- A. Penyusunan Perangkat Tes
- B. Pelaksanaan Tes

BAB VI PENGELOLAAN HASIL TES

- A. Mengolah Tes Hasil Kemampuan Berpikir (Kognitif)

- B. Mengolah Tes Hasil Kemampuan Sikap (Afektif)
- C. Mengolah Tes Hasil Kemampuan Fisik (Psikomotor)

**BAB VII ANALISIS HASILTES**

- A. Penilaian Acuan Patokan (PAP)
- B. Penilaian Acuan Norma (PAN)
- C. Penentuan Nilai Akhir

**BAB VIII ANALISIS SOAL**

- A. Alasan Perlunya Analisis Butir Soal
- B. Karakteristik dan Spesifikasi Butir Soal

**DAFTAR PUSTAKA**

**TENTANG PENULIS**

## PENDAHULUAN

### **A. Pengertian Pengukuran, Penilaian dan Evaluasi**

Banyak orang yang lebih cenderung mengartikan evaluasi, pengukuran dan penilaian sebagai suatu pengertian yang sama sehingga dalam memakainya hanya tergantung dari kata mana yang sedang siap untuk diucapkannya. Ketiga istilah tersebut mempunyai persamaan dan perbedaan. Untuk memahami persamaan, perbedaan, ataupun hubungan dari ketiganya, dapat dipahami melalui contoh-contoh di bawah ini:

- a. *Apabila ada orang yang akan memberi sebatang pensil kepada kita, dan kita disuruh memilih antara dua pensil yang tidak sama panjangnya, maka tentu saja kita akan memilih yang "panjang". Kita tidak akan memilih yang pendek kecuali ada alasan yang sangat khusus.*
- b. *Pasar, merupakan suatu tempat bertemunya orang-orang yang akan menjual dan*

*membeli. Sebelum menentukan barang yang akan dibelinya, seorang pembeli akan memilih dahulu mana barang yang lebih “bagus” menurut ukurannya. Apabila ia ingin membeli jeruk, dipilihnya jeruk yang besar, kuning, kulitnya halus. Semuanya itu dipertimbangkan karena menurut pengalaman sebelumnya, jenis jeruk-jeruk yang demikian ini rasanya akan manis. Sedangkan jeruk yang masih kecil, hijau dan kulitnya agak kasar, biasanya masam rasanya.*

Dari contoh-contoh diatas ini kita dapat simpulkan bahwa sebelum menentukan pilihan, kita mengadakan *penilaian* terhadap benda-benda yang akan kita pilih. Dalam contoh pertama kita memilih mana pensil yang lebih panjang, sedangkan dalam contoh kedua kita menentukan dengan perkiraan kita atas jeruk yang baik, yaitu yang rasanya manis.

Untuk dapat mengadakan penilaian, kita mengadakan *pengukuran terlebih dahulu*. Jika ada penggaris, maka sebelum menentukan pensil mana yang lebih panjang, kita ukur dahulu kedua pensil tersebut. Dan setelah mengetahui berapa panjang masing-masing pensil itu, kita mengadakan penilaian dengan melihat bandingan panjang antara kedua pensil tersebut. Dapatlah kita menyatakan: “ini pensil *panjang* dan ini pensil *pendek*”. Mana pensil yang panjang, itulah yang kita ambil.

Untuk menentukan penilaian mana jeruk yang manis, kita tidak menggunakan “ukuran manis” tetapi menggunakan ukuran *besar, kuning dan halus kulitnya*. Ukuran ini tidak mempunyai

wujud seperti kayu penggaris, tetapi diperoleh berdasarkan pengalaman.

Sebenarnya kita juga mengukur, yakni membandingkan jeruk-jeruk yang ada dengan ukuran tertentu. Setelah itu kita menilai, menentukan pilihan mana jeruk yang paling memenuhi ukuran itulah yang kita ambil.

Dengan demikian kita mengenal dua macam ukuran, yakni ukuran yang terstandar (meter, kilogram, dan sebagainya), ukuran tidak terstandar (depa, jengkal, langkah dan sebagainya) dan ukuran perkiraan berdasarkan hasil pengalaman (jeruk manis adalah yang kuning, besar, halus kulitnya).

Dua langkah kegiatan yang dialui sebelum mengambil barang untuk kita, itulah yang disebut mengadakan evaluasi, yakni mengukur dan menilai. Kita tidak dapat mengadakan penilaian sebelum kita mengadakan pengukuran.

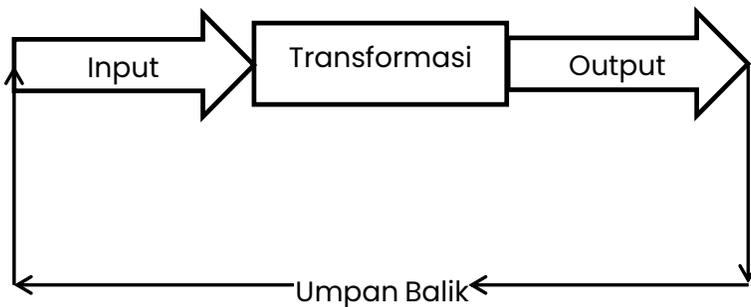
- *Mengukur* adalah membandingkan sesuatu dengan satu ukuran. Pengukuran bersifat *kuantitatif*.
- *Menilai* adalah mengambil suatu keputusan terhadap sesuatu dengan ukuran baik buruk. Penilaian bersifat *kualitatif*.
- *Mengadakan evaluasi* meliputi kedua langkah diatas, yakni mengukur dan menilai. Di dalam istilah asingnya, pengukuran adalah *measurement*, sedang penilaian adalah *evaluation*. Dari kata *evaluation* inilah diperoleh kata Indonesia evaluasi yang berarti menilai (tetapi dilakukan dengan mengukur terlebih dahulu).

## B. Penilaian Pendidikan

Penilaian pendidikan adalah kegiatan menilai yang terjadi dalam kegiatan pendidikan. Guru atau pengelola pengajaran mengadakan penilaian dengan maksud melihat apakah usaha yang dilakukan melalui pengajaran sudah mencapai tujuan.

Apabila sekolah diumpamakan sebagai tempat mengolah sesuatu dan calon siswa diumpamakan sebagai bahan mentah maka lulusan dari sekolah itu dapat disamakan dengan hasil olahan yang sudah siap digunakan. Dalam istilah inovasi yang menggunakan teknologi maka tempat mengolah ini disebut *transformasi*.

Jika digambarkan dalam bentuk diagram akan terlihat sebagai berikut:



- **Input** adalah bahan mentah yang dimasukkan ke dalam transformasi. Dalam dunia sekolah maka yang dimaksud dengan bahan mentah adalah calon siswa yang baru akan memasuki

sekolah. Sebelum memasuki suatu tingkat sekolah (institusi), calon siswa itu dinilai dahulu kemampuannya. Dengan penilaian itu ingin diketahui apakah kelak ia akan mampu mengikuti pelajaran dan melaksanakan tugas-tugas yang akan diberikan kepadanya.

- **Output**

adalah bahan jadi yang dihasilkan oleh transformasi, yakni siswa lulusan sekolah yang bersangkutan. Untuk dapat menentukan apakah seorang siswa berhak lulus atau tidak, perlu diadakan kegiatan penilaian.

- **Transformasi**

adalah mesin yang bertugas mengubah bahan mentah menjadi bahan jadi. Dalam dunia sekolah, sekolah itulah yang dimaksud dengan transformasi. Sekolah itu sendiri terdiri dari beberapa komponen yang menyebabkan berhasil atau gagalnya sebagai transformasi. Bahan jadi yang diharapkan, yang dalam hal ini siswa lulusan sekolah ditentukan oleh beberapa faktor sebagai akibat bekerjanya unsur-unsur yang ada.

Unsur-unsur transformasi sekolah tersebut antara lain:

- a. Guru dan personal lainnya
- b. Bahan pelajaran
- c. Metode mengajar dan sistem evaluasi
- d. Sarana penunjang
- e. Sistem administrasi

- **Umpan Balik (*feed back*)**

adalah segala informasi baik yang menyangkut output maupun transformasi. Umpan balik ini diperlukan sekali untuk

memperbaiki input maupun transformasi. Lulusan yang kurang bermutu atau yang belum memenuhi harapan, akan menggugah semua pihak untuk mengambil tindakan yang berhubungan dengan penyebab kurang bermutunya lulusan.

Penyebab-penyebab tersebut antara lain:

- a. Input yang kurang baik kualitasnya
- b. Guru dan personal yang kurang tepat
- c. Materi yang tidak atau kurang cocok
- d. Metode mengajar dan sistem evaluasi yang kurang memadai
- e. Kurangnya sarana penunjang
- f. Sistem administrasi yang kurang tepat.

Oleh karena itu, penilaian di sekolah meliputi banyak segi: calon siswa, lulusan, dan proses perkuliahan secara menyeluruh.

### **C. Makna Penilaian**

Dalam dunia pendidikan, khususnya dunia persekolahan, penilaian mempunyai makna ditinjau dari berbagai segi.

1. Makna bagi siswa

Dengan diadakannya penilaian, maka siswa dapat mengetahui sejauh mana telah berhasil mengikuti pelajaran yang diberikan oleh guru. Hasil yang diperoleh siswa dari pekerjaan menilai ini ada 2 kemungkinan.

- a) *memuaskan*

Jika siswa memperoleh hasil yang memuaskan, dan hal itu menyenangkan, tentu kepuasan itu ingin diperolehnya lagi pada kesempatan lain waktu. Akibatnya,

siswa akan mempunyai motivasi yang cukup besar untuk belajar lebih giat, agar lain kali mendapat hasil yang lebih memuaskan lagi. Keadaan sebaliknya dapat terjadi, yakni siswa sudah merasa puas dengan hasil yang diperoleh dan usahanya kurang gigih untuk lain kali.

b) *tidak memuaskan*

Jika siswa tidak puas dengan hasil yang diperoleh, ia akan berusaha agar lain kali keadaan itu tidak terulang lagi. Maka ia lalu belajar giat. Namun demikian, keadaan sebaliknya dapat terjadi. Ada beberapa siswa yang lemah kemauannya, akan menjadi putus asa dengan hasil kurang memuaskan yang telah diterimanya.

2. Makna bagi guru

a) Dengan hasil penilaian yang diperoleh guru akan dapat mengetahui siswa-siswa mana yang sudah berhak melanjutkan pelajarannya karena sudah berhasil menguasai bahan, maupun mengetahui siswa-siswanya yang belum berhasil menguasai bahan. Dengan petunjuk ini guru dapat lebih memusatkan perhatiannya kepada siswa-siswa yang belum berhasil. Apalagi jika guru tahu akan sebab-sebabnya, ia akan memberikan perhatian yang memusat dan memberikan perilaku yang lebih teliti sehingga keberhasilan selanjutnya dapat dicapai.

b) Guru akan mengetahui apakah *materi* yang diajarkan sudah tepat bagi siswa sehingga untuk memberikan pengajaran

di waktu yang akan datang tidak perlu diadakan perubahan.

- c) Guru akan mengetahui apakah *metode* yang digunakan sudah tepat atau belum. Jika sebagian besar dari siswa memperoleh angka jelek pada penilaian yang diadakan, mungkin hal ini disebabkan oleh pendekatan atau metode yang kurang tepat. Apabila demikian halnya, maka guru harus mawas diri dan mencoba mencari metode lain dalam mengajar.

### 3. Makna bagi sekolah

- a) Apabila guru-guru mengadakan penilaian dan diketahui bagaimana hasil belajar siswa-siswanya, dapat diketahui pula apakah kondisi belajar yang diciptakan oleh sekolah sudah sesuai dengan harapan atau belum. Hasil belajar merupakan cerminan kualitas suatu sekolah.
- b) Informasi bagi guru tentang tepat atau tidaknya kurikulum untuk sekolah itu dapat merupakan bahan pertimbangan bagi perencanaan sekolah untuk masa-masa yang akan datang.
- c) Informasi hasil penilaian yang diperoleh dari tahun ke tahun, dapat digunakan sebagai pedoman bagi sekolah, yang dilakukan oleh sekolah sudah memenuhi standar atau belum. Pemenuhan standar akan terlihat dari bagusnyanya angka-angka yang diperoleh siswa.  
Secara terperinci dan sesuai dengan urutan kejadiannya, dalam proses

transformasi ini penilaian dibedakan atas tiga jenis, yakni *sebelum*, *selama* dan *sesudah* terjadi proses dalam kegiatan sekolah. Dalam hal ini pelaksana pendidikan selalu berorientasi pada tujuan yang akan dicapai dan tinjauannya selalu diarahkan pada siswa secara perorangan (individual) maupun secara kelompok (per kelas atau per angkatan).

## D. Fungsi Penilaian

Dengan mengetahui makna penilaian ditinjau dari berbagai segi dalam sistem pendidikan, maka dengan cara lain dapat dikatakan bahwa tujuan dan fungsi penilaian ada beberapa hal:

1. Penilaian bersifat selektif  
Dengan cara mengadakan penilaian guru mempunyai cara untuk mengadakan seleksi atau penilaian terhadap siswanya. Penilaian itu sendiri mempunyai beberapa tujuan, antara lain:
  - a) Untuk memilih siswa yang dapat diterima disekolah tertentu
  - b) Untuk memilih siswa yang dapat naik ke kelas atau tingkat berikutnya
  - c) Untuk memilih siswa yang seharusnya mendapat beasiswa
  - d) Untuk memilih siswa yang sudah berhak meninggalkan sekolah dan sebagainya.
2. Penilaian bersifat diagnostik  
Apabila alat yang digunakan dalam penilaian cukup memenuhi persyaratan, maka dengan melihat hasilnya, guru akan mengetahui

kelemahan siswa. Di samping itu diketahui pula sebab-sebab kelemahan itu. Jadi dengan mengadakan penilaian, sebenarnya guru mengadakan *diagnosa* kepada siswa tentang kebaikan dan kelemahannya. Dengan diketahuinya sebab-sebab kelemahan ini, akan lebih mudah dicari cara untuk mengatasinya.

3. Penilaian berfungsi sebagai penempatan  
Setiap siswa sejak lahirnya telah membawa bakat sendiri-sendiri sehingga pelajaran akan lebih efektif apabila disesuaikan dengan pembawaan yang ada. Akan tetapi disebabkan karena keterbatasan sarana dan tenaga, pendidikan, yang bersifat individual kadang-kadang sukar sekali dilaksanakan. Pendekatan yang lebih bersifat melayani perbedaan kemampuan, adalah pengajaran secara kelompok. Untuk dapat menentukan dengan pasti dikelompok mana seorang siswa harus ditempatkan, digunakan suatu penilaian. Sekelompok siswa yang mempunyai hasil penilaian yang sama, akan berada di kelompok yang sama dalam belajar.
4. Penilaian berfungsi sebagai pengukur keberhasilan  
Fungsi keempat dari penilaian ini dimaksudkan untuk mengetahui sejauh mana suatu program berhasil diterapkan. Telah disinggung pada bagian sebelum ini, keberhasilan program ditentukan oleh beberapa faktor yaitu faktor guru, metode mengajar, kurikulum, sarana dan sistem administrasi.

## Bab Dua

# ALAT PENILAIAN

Alat penilaian cukup banyak dan sering disebut teknik penilaian. Berdasarkan sasaran, tujuan dan situasi pelaksanaannya, teknik penilaian terbagi atas tes dan non tes

## A. Tes

Tes adalah suatu pertanyaan atau latihan atau alat lain yang digunakan untuk mengukur keterampilan, pengetahuan, intelegensi, kemampuan atau bakat yang dimiliki oleh individu atau kelompok. Jenis tes beraneka ragam dapat dibagi berdasarkan bentuknya, fungsinya, saat pelaksanaannya, keperluannya dan kebakuannya.

1. Berdasarkan bentuknya, tes terbagi atas tes verbal dan nonverbal
  - a) *Tes verbal*, ialah tes yang menghendaki jawaban siswa dengan kata-kata secara lisan atau tertulis. Tes tertulis dapat pula terbagi atas tes objektif dan esai.

- b) *Tes nonverbal*, ialah tes terhadap keterampilan yang dimiliki siswa. Tes ini dilakukan dengan format penilaian atau lembar pengamatan, misalnya untuk menilai tulisan indah, lukisan, olahraga, menjahit, memasak, mengetik dan sebagainya.
2. Berdasarkan fungsi dan tujuannya, tes terbagi atas tes formatif, subsumatif, tes sumatif, tes belajar tahap akhir, tes seleksi, dan tes penempatan.
  3. Berdasarkan saat pelaksanaannya, kita kenal tes awal dan tes akhir
    - a) Tes awal (*pretest*), yaitu tes yang dilaksanakan sebelum KBM dimulai. Tes ini diperlukan untuk mengetahui sejauh mana pengetahuan dan kemampuan yang dimiliki siswa, untuk mengetahui apakah siswa telah memiliki pengetahuan prasyarat untuk memulai bahan baru dalam KBM, untuk mengetahui darimana sebaiknya guru mulai. Selain itu juga untuk mengetahui sejauh mana keberhasilan guru mengelola KBM (setelah dibandingkan dengan tes akhir)
    - b) Tes akhir (*posttest*), yaitu tes yang dilaksanakan setelah melaksanakan KBM, untuk melihat sejauh mana keberhasilan KBM setelah dibandingkan dengan hasil awal tes.
  4. Berdasarkan keperluan, kita mengenal berbagai macam, yaitu:
    - a) Tes hasil belajar (*achievement test*), yaitu tes yang diperlukan untuk mengetahui hasil belajar siswa. Tes ini sering sekali dilaksanakan di sekolah.
    - b) Tes kejiwaan (*psycho-test*), yaitu yang diperlukan untuk mengetahui sifat, semangat,

kecenderungan, keinginan, atau keadaan kejiwaan seseorang.

- c) Tes kecerdasan (*intelligence test*), yaitu tes yang diperlukan untuk mengetahui kecerdasan (IQ) seseorang, dengan jalan membagi usia mental dengan usia sesungguhnya dikalikan 100. Tes ini sering disebut tes IQ.
- d) Tes kepribadian (*personality test*), yaitu tes yang diperlukan untuk mengetahui sifat, semangat, kecenderungan kehidupan jiwa seseorang dalam aspek yang lebih luas.

Selain itu untuk keperluan-keperluan tertentu kita kenal pula *tes daya tahan tubuh*, *tes kesehatan badan* dan lain sebagainya.

- 5. Berdasarkan kebakuanannya, kita mengenal tes baku dan tes tidak baku.
  - a) Tes baku (*standardized*), yaitu tes yang sudah baku (standar), telah memenuhi syarat-syarat tes yang baik karena telah diujicobakan. Biasanya dibuat oleh instansi tertentu yang khusus mengelola penilaian.
  - b) Tes nonbaku, yaitu tes yang dibuat oleh setiap guru yang hendak melakukan penilaian. Sering butir-butir soal itu dibuat secara cepat dan kurang memperhatikan syarat-syarat penyusunan soal yang baik.

## B. Nontes

Bentuk penilaian yang termasuk nontes adalah sebagai berikut:

- 1. Angket (*questionnaire*)  
Angket adalah instrumen untuk pengumpulan bahan informasi dengan daftar pertanyaan kepada responden (penjawab). Butir pertanyaan

itu disebut kuesioner. Kuesioner ini ada yang tertutup dan ada yang terbuka. Kuesioner tertutup ialah kuesioner yang hanya memerlukan jawaban singkat seperti *ya* atau *tidak* atau memberi tanda cek di tempat yang disediakan. Kuesioner terbuka adalah kuesioner yang memberikan kesempatan kepada responden untuk menjawab dengan kata-kata atau kalimat di tempat yang telah disediakan.

2. Pengamatan (*observation*)  
Pengamatan dilakukan untuk mengetahui perbuatan, tingkah laku, semangat, kesungguhan, kesabaran, kerjasama, tanggung jawab pada waktu siswa melakukan KBM. Yang bersangkutan sebaiknya tidak mengetahui bahwa ia sedang diamati, dan perlu dilakukan beberapa kali agar hasil lebih terpercaya.
3. Wawancara (*interview*)  
Wawancara lebih praktis pelaksanaannya. Kita dapat langsung berhadapan dengan pihak yang diwawancarai. Informasi yang didapatkan akan lebih sesuai dengan bahan yang diwawancarakan. Hanya pelaksanaannya memerlukan waktu yang cukup lama.
4. Skala penilaian (*rating scales*)  
Penilaian dengan skala penilaian dilakukan oleh orang yang mengetahui betul tentang orang yang akan dinilai. Penilaian dilakukan menurut skala yang disediakan. Biasanya aspek yang dinilai menyangkut: sikap, minat, gairah kerja, inisiatif, cara bergaul, kesungguhan, kepemimpinan, sikap bertanggung jawab, ketelitian, kejujuran, kesabaran.

5. Surat pujian (*recommendation*)

Seorang calon karyawan atau calon siswa tertentu diperlukan bahan pertimbangan untuk dapat diterima atau ditolak. **Salah satu bahan pertimbangan untuk dapat diterima atau ditolak.** Salah satu bahan pertimbangan ialah informasi berupa rekomendasi dari sekolah asal atau dari tempat asal kerja. Surat ini sering pula disebut referensi.

## C. Sasaran penilaian

### 1. Ranah Kognitif

Benyamin S. Bloom dengan *Taxonomy of Education Objective* membagi enam tingkat kemampuan pada ranah kognitif. Kemampuan ini perlu dikembangkan dalam pendidikan dan menjadi sasaran dalam penilaian, yaitu:

#### a. Ingatan (*knowledge, recall*) = C1

Ingatan adalah tingkat kognitif yang paling rendah, yaitu kemampuan untuk mengingat kembali bahan yang telah diinformasikan. Kemampuan siswa tingkat ini adalah mengetahui istilah-istilah, fakta-fakta yang khas, cara atau metode, konsep-konsep dasar dan prinsip-prinsip.

Perilaku siswa yang mencerminkan kemampuan mengingat kembali antara lain:

- mendefinisikan
- memberikan label nama-nama
- memerikan
- mengenal
- membuat senarai

- menyebutkan
- memilih
- memadankan
- membuat ikhtisar
- **menyebutkan**

### **b. Pemahaman (*Comprehension*) = C2**

Kemampuan intelektual siswa untuk memahami, mengerti, dapat menafsirkan atau meramalkan hal-hal yang diajarkan/dipelajari, bahan yang diinformasikan atau hal yang dibahas. Kemampuan siswa tingkat ini berupa:

- (a) mengerti tentang fakta, konsep dan prinsip-prinsip
- (b) menafsirkan informasi verbal
- (c) menafsirkan bagan atau diagram
- (d) menerjemahkan informasi verbal ke rumus-rumus matematik
- (e) meramalkan kecenderungan berdasarkan data
- (f) membenarkan suatu cara atau metode

Perilaku siswa yang mencerminkan kemampuan memahami di antaranya adalah:

- (a) mengkonversikan angka-angka (*convert*)
- (b) memperluas informasi (*extend*)
- (c) mempertahankan pendirian (*defend*)
- (d) menjeneralisasikan (*generalizes*)
- (e) membedakan sesuatu (*distinguishes*)
- (f) memberikan contoh (*example*)
- (g) meramalkan sesuatu akibat (*estimate*)
- (h) menulis kembali (*rewrite*)
- (i) menerangkan suatu akibat (*explains*)
- (j) mengintisarikan (*summarizes*)

### c. Penerapan (*Application*) = C3

Penerapan yaitu kemampuan intelektual siswa untuk menerapkan, menggunakan bahan yang telah diinformasikan, diajarkan/dipelajari terhadap situasi baru atau kenyataan yang dihadapi. Misalnya, mampu menggunakan konsep, prinsip, hukum, teori, rumus dan peraturan-peraturan.

Kemampuan siswa tingkat ini berupa:

- (a) menerapkan prinsip-prinsip atau konsep-konsep yang telah dipahami untuk menghadapi situasi baru
- (b) menerapkan hukum (dalil) atau teori-teori
- (c) memecahkan persoalan dengan konsep-konsep yang telah dimiliki
- (d) memecahkan perhitungan-perhitungan matematik
- (e) membuat grafik atau diagram
- (f) menunjukkan penggunaan suatu metode

Perilaku siswa yang mencerminkan kemampuan menerapkan di antaranya ialah:

- (a) mengubah pola (*changes*)
- (b) menghitung/mengklasifikasi (*calculate*)
- (c) memodifikasi (*modifies*)
- (d) mendemonstrasikan penggunaan (*operate*)
- (e) memprediksi (*predict*)
- (f) memecahkan persoalan (*solves*)

### d. Analisis (*Analysis*) = C4

Kemampuan analisis adalah kemampuan tingkat intelektual siswa untuk menjabarkan bahan yang diajarkan/dipelajari menjadi bagian-bagian

yang tetap terorganisasi dan tetap dapat dipahami. Misalnya, kemampuan mengenal bagian-bagian, hubungan antar bagian, dan prinsip-prinsip hubungan antar bagian.

Kemampuan siswa tingkat analisis ini berupa:

- (a) mengenal asumsi-asumsi yang terselubung
- (b) mengenal kesalahan logika penalaran
- (c) membedakan fakta dan kesimpulan
- (d) menguji hubungan antar data
- (e) menganalisis suatu struktur yang terorganisasi

Perilaku siswa yang mencerminkan kemampuan menganalisis diantaranya adalah:

- (a) menjabarkan (*differentiates*)
- (b) menguraikan (*breakdown*)
- (c) membedakan (*discriminate*)
- (d) memisahkan (*separate*)
- (e) menurunkan (*infers*)

### **e. Sintesis (*Synthesis*) = C5**

Sintesis adalah kemampuan intelektual siswa untuk menghimpun bagian-bagian menjadi satu kesatuan yang terorganisasi dan terpadu. Misalnya kemampuan untuk menyusun tema suatu karangan, rencana untuk pemecahan suatu masalah, skem untuk menganalisis suatu informasi.

Kemampuan siswa tingkat sintesis ini berupa:

- (a) merumuskan tema sebuah karangan yang terorganisasi
- (b) berpidato menyampaikan bahan yang terorganisasi
- (c) menulis cerita pendek, puisi atau lagu yang gagasannya terorganisasi

- (d) mengusulkan suatu rencana dengan sistematika kegiatannya
- (e) mengintegrasikan hasil belajar dari berbagai bidang studi (mata pelajaran) untuk memecahkan masalah yang dihadapi
- (f) merumuskan skema baru untuk mengklasifikasikan berbagai objek
- (g) merumuskan suatu tujuan yang mencakup berbagai aspek
- (h) merumuskan berbagai kesimpulan yang mencakup berbagai aspek

Perilaku siswa yang mencerminkan kemampuan mensistesisikan diantaranya ialah:

- (a) mengkategorikan (*categorizes*)
- (b) mengkombinasikan (*combines*)
- (c) mengkompilasikan (*compiles*)
- (d) mengkomposisikan (*composes*)
- (e) merakit (*assembles*)
- (f) mengorganisasikan (*organizes*)

## **f. Evaluasi (*Evaluation*) = C6**

Evaluasi merupakan tingkat kemampuan intelektual yang paling tinggi yaitu kemampuan intelektual siswa untuk menilai penerapan sesuatu disertai alasan-alasan yang tepat. Misalnya menilai suatu pernyataan, novel, puisi, laporan penelitian dan sebagainya.

Kemampuan siswa tingkat evaluasi ini berupa:

- (a) menilai ketaatan penalaran
- (b) menilai kelayakan data untuk dasar sesuatu kesimpulan.
- (c) menilai hasil suatu karya/hasil seni

- (d) menilai suatu pertunjukan
- (e) menilai suatu kegiatan dan sebagainya

Perilaku siswa yang mencerminkan kemampuan mengevaluasi ini di antaranya ialah:

- (a) menghargai (*appraise*)
- (b) menilai kelayakan data untuk dasar sesuatu kesimpulan.
- (c) menilai hasil suatu karya/hasil seni
- (d) menilai suatu pertunjukan
- (e) menilai suatu kegiatan dan sebagainya

Perilaku siswa yang mencerminkan kemampuan mengevaluasi ini di antaranya ialah:

- (a) menghargai (*appraises*)
- (b) mengeritik (*critizes*)
- (c) membenarkan (*justifies*)

## 2. *Ranah Afektif*

Ranah afektif ialah ranah yang berkaitan dengan ketajaman perasaan (emosi) siswa. Sasaran penilaian berkaitan dengan keberhasilan kita mengembangkan ketajaman perasaan (emosi) siswa, dari taraf menerima, menaruh minat hingga bersikap taat asas terhadap makna sesuatu hal atau peristiwa berdasarkan hati nurani.

Ranah afektif mengandung lima aspek yang biasa menjadi sasaran penilaian atau umpan balik bagi guru. Lima aspek tersebut ialah:

- a) Menerima nilai  
Menerima nilai ialah kesadaran para siswa akan adanya nilai-nilai tertentu

dalam kehidupan atau dalam masyarakat, mau menerima nilai-nilai itu dan aktif memperhatikan nasihat yang berhubungan dengan nilai tersebut. Misalnya:

- (a) Sadar dan merasakan adanya cara berpakaian sopan, cara berbicara dengan orang tua dan guru, cara bertetangga dan bergaul sehari-hari
  - (b) Mau mendengarkan pembicaraan orang lain, mau menerima adat-istiadat orang lain, peka terhadap lingkungan.
  - (c) Aktif memperhatikan nasihat orang tua atau guru, memperhatikan orang membaca sesuatu yang penting
- b) Menanggapi
- Dalam hal ini yang dimaksud dengan menanggapi adalah kerelaan, kepatuhan dan perasaan puas para siswa menanggapi sesuatu. Misalnya:
- (a) Mau mengikuti tata tertib sekolah, peraturan lalu lintas
  - (b) Rela mengikuti aturan pembuangan sampah, kebersihan, dan aturan hidup sehat
  - (c) Mau aktif menyelenggarakan majalah sekolah, mengikuti ketentuan OSIS, PKS, UKS, dan ketentuan lainnya.
  - (d) Merasa senang membaca buku di perpustakaan.
  - (e) Puas dan senang mengikuti kegiatan kesenian dan olahraga
- c) Menghargai
- Menghargai dalam hal ini ialah sikap para siswa untuk menghargai nilai-nilai yang baik

ditemukan dalam masyarakat. Dari sikap menghargai ini kemudian diwujudkan dalam tingkah laku dan akan tumbuh nilai-nilai ideal yang diyakini dan dijadikan pegangan dalam perbuatan sehari-hari. Misalnya:

- (a) Menerima baik suatu nilai dalam masyarakat
- (b) Menghargai nilai tertentu yang dianggap baik
- (c) Melaksanakan perbuatan yang diyakini mengandung nilai yang baik
- d) Mengorganisasi nilai  
Siswa mengorganisasi nilai-nilai yang diyakini sebagai nilai yang baik. Siswa mengelompokkan beberapa nilai, mencari hubungan yang harmonis dan taat asas. Dapat terjadi pula munculnya nilai yang lebih tinggi. Misalnya:
  - (a) Pekerjaan yang disenangi akan dirasakan sebagai rekreasi
  - (b) Pekerjaan yang benar bagaimanapun akan terbukti benar
- e) Watak yang mantap oleh nilai-nilai kompleks  
Nilai-nilai yang tersusun mantap, matang yang terjadi pedoman hidup seorang dewasa yang telah mapan. Ia telah kukuh berdiri di atas keyakinannya, tidak dapat diombang-ambingkan oleh cobaan atau pengaruh dari luar. Setiap masalah dihadapi secara realistis, objektif, adil, jujur, sabar, tenggang rasa dan tawakkal.

Kognitif ada hubungannya dengan afektif. Untuk mendapatkan pengetahuan (kognitif) harus menerima dan ada perhatian (afektif) terhadap bahan

yang diinformasikan (diajarkan). Kesiapan mengikuti tata tertib dan peraturan diperlukan pemberitahuan dan pemahaman terhadap tata tertib dan peraturan itu. Bahan pelajaran yang senangi (afektif) akan lebih mudah dipahami (kognitif) atau sesuatu yang telah dipahami, dapat menimbulkan rasa puas. Orang dapat meyakini ajaran agama karena telah memahami kebenaran ajaran agama tersebut.

### 3. *Ranah Psikomotor*

Sasaran penilaian untuk ranah ini berkaitan erat dengan pengembangan keterampilan dari tingkat profesional fungsi saraf dan otot (*neuro-mascular*) secara terintegrasi dan terorganisasi. Ranah psikomotor memiliki enam macam, yaitu:

a) Gerak refleks (*reflex movement*)

Gerak refleks ialah gerak yang terjadi secara otomatis dan tiba-tiba terhadap rangsangan dari luar, tanpa kendali olah pusat urat saraf, misalnya:

(a) refleks satu gerak saja (*segmental reflexes*)

(b) refleks berturut-turut (*intersegmental reflexes*)

(c) refleks yang telah melibatkan unsur pikir (*suprasegmental reflexes*), seperti pada waktu kita terpeleset, menjaga keseimbangan di atas titian

Gerak refleks ini belum dikaitkan dengan tujuan pendidikan

b) Gerak dasar

Gerak ini merupakan gerak dasar untuk gerak lebih lanjut, termasuk gerak dasar ini ialah

gerak ditempat (*nonlocomotor movement*), gerak untuk berpindah tempat (*locomotor movement*) dan gerak terkoordinasi (*manipulative movement*), misalnya:

- (a) Jalan di tempat, loncat di tempat, memutar anggota badan, memutar badan, goyang pinggul, menekan, meloncat, berlari, meluncur
  - (b) Merangkak, berguling-guling, berjalan, melompat, berlari, meluncur.
  - (c) Menjahit, menyulam, menggambar, menulis, main musik, melempar dan menangkap bola dalam permainan, memotong sesuatu, menyusun letak sesuatu.
- c) Kemampuan menanggapi (*perceptual abilities*)

Kemampuan menanggapi yaitu kemampuan siswa untuk memahami stimulus (rangsangan) dari luar yang menyebabkan ia melakukan gerak menanggapi stimulus itu, setelah prosesnya melalui pusat urat saraf. Kemampuan menanggapi ini merupakan kemampuan-kemampuan:

- 1) Membedakan gerak (*kinetic discrimination*).

Siswa mampu membedakan gerak pada keadaan tertentu. Misalnya, menyadari bagaimana bagan bergerak, bagaimana hubungan badan dengan benda di sekitarnya. Situasi itu misalnya pada saat berdiri seimbang pada satu kaki sambil melakukan gerak tertentu.

- 2) Membedakan sesuatu yang dilihat (*visual discrimination*)  
Siswa mampu membedakan bentuk benda yang dilihat. Misalnya: membedakan bentuk bola dengan kotak, kerucut, balok, dan sebagainya. Siswa mampu menyaksikan dua layang-layang berada di udara, dua ekor kerbau yang berlaga, atau pertarungan dua pesawat di angkasa.
- 3) Membedakan yang didengar (*auditory discrimination*)  
Siswa mampu membedakan berbagai bunyi dan berbagai suara, membedakan bunyi yang mendekat dan menjauh. Mampu membedakan suara A dengan suara B, mampu membedakan bunyi suling dan klarinet.
- 4) Membedakan rabaan (*tactile discrimination*)  
Siswa mampu membedakan benda-benda yang halus, kasar, licin. Kasap, lunak, keras, tumpul, tipis, tebal, kaku, kenyal, dan sebagainya.
- 5) Terkoordinasi (*coordinated abilities*)  
Siswa mampu mengkoordinasikan beberapa macam gerak. Misalnya melihat teks surat sambil jari-jari dengan licin bergerak menekan huruf-huruf mesin ketik, mata melihat not-not balok sambil jari-jari tangan menekan tut-tut organ/piano dan kaki kanan berpindah-pindah menekan pedal bunyi bas pada organ/piano tersebut.

- d) Kemampuan fisik (*physical abilities*)
- 1) Daya tahan (*endurance*), misalnya kemampuan fisik mengikuti lari marathon, larilintasalam, menyelam dan sebagainya.
  - 2) Daya lentur (*flexibility*), misalnya melakukan kayang, mencium lutut dalam keadaan kaki lurus. Demikian juga kelenturan-kelenturan yang dilakukan para penari balet.
  - 3) Kekuatan (*strength*), misalnya kekuatan yang dimiliki oleh seorang yudoka, pegulat, petinju, angkat berat dan sebagainya.
  - 4) Ketangkasan, kegesitan (*agility*), misalnya lari 100 meter (*sprint*), menyerang atau menangkis secara tiba-tiba dalam olahraga silat.
- e) Gerak yang terampil (*skilled movement*)
- Gerak terampil didapatkan melalui latihan yang berulang-ulang sampai betul-betul efisien. Gerak ini terdiri atas tiga macam, yaitu:
- 1) Gerak sederhana yang sesuai (*simple adaptive skill*), misalnya gerak yang sederhana, menarik dan mendorong untuk menggergaji papan.
  - 2) Gerak sederhana yang tergabung (*compound adaptive skills*), misalnya gabungan beberapa gerak dasar pada waktu bermain ping-pong atau bermain badminton
  - 3) Gerak sederhana yang kompleks (*complex adaptive skills*), misalnya gerakan akrobatik, senam lantai, senam kuda-kuda, atau berbagai jenis keterampilan gerak pada pertunjukan sirkus.

f) Gerak untuk komunikasi (*nondiscrusive communication*)

Gerak yang bersifat komunikatif dengan aneka ragam variasi, gerak jenis ini dibagi atas dua macam, yaitu:

- 1) Gerak yang bersifat ekspresif (*expressive movement*), yaitu gerak yang bersifat pengungkapan perasaan melalui komunikasi verbal, misalnya memberi tekanan tertentu pada kata-kata yang diucapkan

Gerak yang bersifat interpretasi (*interpretative movement*), yaitu gerak estetis dan kreatif untuk mengungkapkan interpretasi tertentu yang mengandung kesan dan pesan secara artistik. Biasanya gerak ini terpadu dengan bunyi musik yang mengiringi.

## Bab **Tiga**

# PENYUSUNAN TES HASIL BELAJAR

Apabila untuk suatu tindak evaluasi berupa tes hasil belajar telah tersedia alat-alat atau instrumennya, maka kita tinggal memilih saja alat yang sesuai untuk digunakan. Tetapi apabila alat-alat itu belum tersedia, maka kita harus membuat dan menyusun sendiri alat-alat yang akan digunakan itu. Rangkaian kegiatan suatu tes hasil belajar diawali dengan penyusunan *kisi tes*

### **A. Kisi Tes**

Kisi tes adalah suatu format berupa matriks yang memuat pedoman untuk menulis soal atau merakit soal menjadi suatu tes. Penyusunan kisi tes merupakan langkah penting yang harus dilakukan sebelum penulisan soal, karena didalam kisi tes terdapat penjelasan isi dan bentuk soal yang diterima.

Kisi tes prestasi belajar harus memenuhi beberapa persyaratan yaitu:

1. Mewakili isi kurikulum/kemampuan yang akan diujikan
2. Komponen-komponennya rinci, jelas dan mudah dipahami
3. Soal-soalnya dapat dibuat sesuai dengan indikator dan bentuk soal yang ditetapkan
4. Komponen-komponen yang biasa dimuat dalam penyusunan kisi tes prestasi belajar adalah sebagai berikut:
  - a. Jenis sekolah/ jenjang sekolah
  - b. Tingkat sekolah
  - c. Bidang studi/mata pelajaran
  - d. Tahun ajaran
  - e. Kurikulum yang diacu/dipergunakan
  - f. Jumlah soal
  - g. Bentuk soal
  - h. Tujuan kelas/tujuan pembelajaran atau merupakan hal yang diukur
  - i. PB/SPB atau Tema/Subtema
  - j. Materi yang akan diberikan/dijadikan soal
  - k. Indikator
  - l. Nomor urut soal (jika diperlukan)

*Contoh format kisi-kisi soal*

NO URUT	TUJUAN PEMBELAJARAN	PB/ PSB	MATERI	INDI- KATOR	NO URUT SOAL

## **B. Klasifikasi Bentuk Soal**

### **1. Tes Uraian/Esei**

Tes uraian/esei adalah butir soal yang mengandung pertanyaan atau tugas yang jawaban atau pengerjaan soal tersebut harus dilakukan dengan cara mengekspresikan pikiran peserta tes. Peserta tes bebas untuk menjawab pertanyaan yang diajukan. Setiap peserta tes dapat memilih, menghubungkan, dan menyampaikan gagasannya dengan menggunakan kata-katanya sendiri.

#### *1) Kebaikan Tes Uraian/Esei*

Soal uraian mempunyai beberapa kelebihan yang tidak dapat atau sukar diperoleh melalui penggunaan tipe butir soal lain. Kelebihan itu antara lain:

- a. Mudah disiapkan dan disusun
- b. Tidak memberi banyak kesempatan untuk berspekulasi atau untung-untungan
- c. Mendorong siswa untuk berani mengemukakan pendapat serta menyusun dalam bentuk kalimat yang bagus
- d. Memberi kesempatan kepada siswa untuk mengutarakan maksudnya dengan gaya bahasa dan caranya sendiri
- e. Dapat diketahui sejauh mana siswa mendalami suatu masalah yang ditekankan.

#### *2) Kelemahan Tes Uraian/Esei*

Tes uraian juga mengandung kelemahan yang serius. Beberapa kelemahan pokok tersebut adalah:

- a. Kadar validitas dan realibilitas rendah karena sukar diketahui segi-segi mana dari

- pengetahuan siswa yang betul-betul telah dikuasai
- b. Kurang representatif dalam hal mewakili seluruh bagian bahan pelajaran yang akan dites karena soalnya hanya beberapa saja (terbatas)
  - c. Cara memeriksanya banyak dipengaruhi oleh unsur-unsur subjektif
  - d. Pemeriksaannya lebih sulit sebab membutuhkan pertimbangan individual lebih banyak dari penilai
  - e. Waktu untuk koreksinya (pemeriksaan) lebih lama dan tidak dapat diwakilkan kepada orang lain.
- 3) *Penggunaan Tes Uraian/Esei*
- a) Bila jumlah siswa atau peserta ujian terbatas maka soal uraian dapat digunakan karena masih mungkin bagi guru untuk dapat memeriksa/menskor hasil ujian tersebut secara baik
  - b) Bila waktu yang dimiliki guru untuk mempersiapkan soal sangat terbatas, sedangkan ia mempunyai waktu yang cukup untuk memeriksa hasil ujian, maka soal uraian dapat dipergunakan.
  - c) Bila tujuan instruksional yang ingin dicapai adalah kemampuan mengekspresikan pikiran dalam bentuk tertulis, menguji kemampuan menulis dengan baik, atau kemampuan penggunaan bahasa secara tertib, maka haruslah menggunakan tes uraian.
  - d) Bila guru ingin memperoleh hasil pengalaman belajar siswanya, maka tes uraian merupakan salah satu bentuk yang paling cocok untuk mengukur pengalaman belajar tersebut.

4) *Klasifikasi Tes Uraian/Esei*

Tes uraian secara umum dapat dibagi menjadi dua jenis, yaitu tes uraian bebas (*extended responses*) dan tes uraian terbatas (*restricted responses*). Perbedaan kedua jenis tes uraian ini adalah besarnya kebebasan yang diberikan kepada peserta tes untuk mengorganisasikan, menulis, dan menyatakan pikiran dan gagasannya

a) *Tes Uraian Bebas (extended responses)*

Dalam soal tes bebas hampir-hampir tidak ada pembatasan terhadap tes dalam memberikan jawabannya. Peserta tes memiliki kebebasan yang luas sekali untuk mengorganisasikan dan mengekspresikan pikiran dan gagasannya dalam menjawab soal tersebut. Dengan demikian, jawaban siswa bersifat terbuka, fleksibel, dan tidak terstruktur. Contoh butir soal uraian bebas:

- *Uraikanlah peranan pemuda dalam perjuangan mencapai kemerdekaan Indonesia sejak tahun 1908 sampai dengan tahun 1928. Dalam uraian Anda hendaknya terdapat contoh-contoh organisasi pemuda yang ada pada masa itu beserta pemimpinnya. Uraian anda hendaknya tidak melebihi 2 halaman.*

b) *Tes Uraian Terbatas (restricted responses)*

Dalam menjawab tes uraian terbatas, peserta tes lebih dibatasi oleh berbagai rambu-rambu yang ditentukan dalam butir soal. Keterbatasan itu mencakup format, isi, dan ruang lingkup jawaban. Jadi soal tes uraian terbatas ini harus menentukan batas

jawaban yang dikehendaki. Batas itu meliputi konteks jawaban yang diinginkan, jumlah butir jawaban yang diharapkan, keluasan uraian jawaban, arah dan luas jawaban yang diminta. Misalnya:

- *Sebutkanlah tiga cara mencegah erosi di lahan kritis!*
- *Bandingkanlah taktik dan strategi perjuangan mencapai kemerdekaan antara Ir. Soekarno dan Drs. Moh. Hatta!*

Ada beberapa ragam tes uraian terbatas, antara lain ragam tes melengkapi dan ragam tes jawaban singkat.

#### (1) Ragam Tes Melengkapi

Yang dimaksud dengan ragam tes melengkapi adalah butir soal yang meminta atau memerintah peserta untuk melengkapi suatu kalimat dengan satu kata, satu frase, satu angka atau satu formula.

Contoh:

- *Kebiasaan merokok akan menyebabkan penyakit ....*
- *Seorang anak umur 12 tahun sebaiknya setiap hari minum susu murni ... gelas.*

#### (2) Ragam Tes Jawaban Singkat

Yang dimaksud dengan tipe butir soal jawaban singkat ialah butir soal berbentuk pertanyaan yang dapat dijawab dengan satu kata, satu frase, satu angka atau satu formula.

Contoh:

- *Siapakah pendiri kerajaan Majapahit? \_\_\_\_\_*
- *$x + 2 = 4$  Berapa  $x$ ? \_\_\_\_\_*

- Apakah nama ibu kota provinsi berikut?
- Irian Jaya -----
- Sulawesi Barat -----

## 2. Tes Objektif

Yang dimaksud dengan tes objektif adalah butir soal yang telah mengandung kemungkinan jawaban yang harus dipilih atau dikerjakan oleh peserta tes. Peserta hanya memilih jawaban yang dianggap benar dari kemungkinan jawaban yang telah disediakan.

- a. Kebaikan-kebaikan tes objektif, antara lain adalah:
  - 1) Lebih representatif mewakili isi dan luas bahan, lebih objektif dalam penskoran
  - 2) Lebih mudah dan cepat cara memeriksanya karena dapat menggunakan kunci tes bahkan alat-alat teknologi
  - 3) Pemeriksaannya dapat diserahkan kepada orang lain
  - 4) Dalam pemeriksaan, tidak ada unsur subjektif yang mempengaruhi
- b. Kelemahan-kelemahan tes objektif antara lain adalah:
  - 1) Persiapan untuk menyusunnya jauh lebih sulit daripada tes esei karena soalnya banyak dan harus teliti untuk menghindari kelemahan-kelemahan yang lain.
  - 2) Soal-soalnya cenderung untuk mengungkapkan ingatan dan daya pengenalan kembali saja, dan sukar untuk mengukur proses mental yang tinggi.
  - 3) Banyak kesempatan untuk main untung-untungan.

- 4) “kerjasama” antar siswa pada waktu mengerjakan soal tes lebih terbuka.

Secara umum, ada tiga tipe tes objektif, yaitu:

- a. benar salah (*true false*)
- b. menjodohkan (*matching*)
- c. pilihan ganda (*multiple choice*)

Tes objektif pilihan ganda dapat dimodifikasi kedalam 5 (lima) ragam berikut:

- a. pilihan ganda biasa
- b. pilihan ganda analisis hubungan antar hal
- c. pilihan ganda analisis kasus
- d. pilihan ganda kompleks
- e. pilihan ganda yang menggunakan diagram, grafik, tabel atau gambar

### **3. Perbedaan Tes Uraian dengan Tes Objektif**

Bila dibandingkan antara tes uraian dengan tes objektif, terdapat berbagai perbedaan. Perbedaan kedua bentuk tes tersebut dapat terlihat dari tabel berikut ini.

Perbedaan antara tes objektif dengan tes uraian

<b>TES OBJEKTIF</b>	<b>TES URAIAN</b>
Baik untuk mengukur pengetahuan ingatan, pemahaman, aplikasi, dan analisa. Kurang tepat untuk mengukur sintesa dan evaluasi	Kurang baik untuk mengukur ingatan, baik untuk mengukur pemahaman, aplikasi, analisa dan paling baik untuk mengukur sintesa dan evaluasi

Dapat mengukur lebih banyak sampel pertanyaan sehingga benar-benar mewakili materi yang diajarkan	Hanya dapat menanyakan beberapa pertanyaan sehingga kurang mewakili materi yang diajarkan
Menyusun pertanyaan yang baik sulit dilakukan dan memakan waktu yang banyak	Menyusun pertanyaan yang baik sulit tetapi lebih mudah dibandingkan pertanyaan objektif
Pengolahan objektif, sederhana, dan ketetapannya (reabilitas) tinggi	Pengolahannya sangat subjektif, sukar dan ketepatannya (reabilitas) rendah
Hasil kemampuan siswa dapat terganggu oleh kemampuan membaca dan menerka	Hasil kemampuan siswa dapat terganggu oleh kemampuan menulis dan bernalar
Mendorong siswa untuk lebih banyak mengingat, membuat interpretasi dan menganalisa ide orang lain	Mendorong siswa untuk mengorganisasikan, menghubungkan, dan menyatakan ide sendiri secara tertulis
Penyelesaian tes oleh siswa dan pengolahan tes oleh guru memerlukan waktu singkat	Penyelesaian tes oleh siswa dan pengolahan tes oleh guru memerlukan waktu yang cukup banyak

## Bab **Empat**

### **KONSTRUKSI ALAT PENILAIAN**

Konstruksi adalah peninjauan kembali untuk diperbaiki terhadap sesuatu. Konstruksi alat penilaian adalah peninjauan kembali butir soal agar dapat mengukur hasil belajar dengan baik dan sesuai dengan tujuan yang ingin dicapai. Dengan pengkonstruksian butir soal yang baik, maka alat penilaian tersebut akan dapat mengukur secara akurat, apapun bentuk dan tipe soalnya. Yang perlu dihayati ialah kemampuan menyusun butir soal yang baik tidak hanya bersifat *pengetahuan dan pemahaman*, tetapi lebih berupa *keterampilan*. Bahkan untuk mencapai tahapan mahir dalam kemampuan mengkonstruksi butir soal, maka aspek *kiat* akan mempunyai peranan yang penting. Guna mencapai kemampuan konstruksi butir soal yang mahir dengan kiatnya itu dibutuhkan pelatihan yang terus-menerus.

## **A. Dasar-dasar Penyusunan Tes Hasil Belajar**

Tes Hasil Belajar (THB) adalah salah satu alat ukur yang paling banyak digunakan untuk menentukan keberhasilan seseorang dalam suatu proses belajar-mengajar atau untuk menentukan keberhasilan suatu program pendidikan. Alat ukur lainnya yang penggunaannya sangat terbatas antara lain pedoman wawancara, pedoman observasi, angket, skala sikap, dan daftar isian. Kesemua alat ukur ini mempunyai peranan tersendiri, namun alat ukur yang satu dengan lainnya dapat saling mendukung dalam pengukuran hasil belajar.

Adapun dasar-dasar penyusunan tes hasil belajar adalah sebagai berikut:

1. THB harus dapat mengukur apa-apa yang dipelajari dalam proses belajar mengajar sesuai dengan tujuan instruksional yang tercantum di dalam kurikulum yang berlaku
2. THB yang disusun, benar-benar mewakili bahan yang telah di pelajari
3. Pertanyaan THB hendaknya disesuaikan dengan aspek-aspek tingkat belajar yang diharapkan.
4. THB hendaknya disusun sesuai dengan tujuan penggunaan tes itu sendiri, karena tes dapat disusun untuk keperluan: pre-tes, post-test, tes diagnostik, tes prestasi belajar, tes formatif, dan tes sumatif.
5. THB disesuaikan dengan pendekatan pengukuran yang dianut, apakah mengacu pada kelompok ataukah mengacu pada patokan tertentu
6. THB hendaknya dapat digunakan untuk memperbaiki proses belajar-mengajar

## B. Konstruksi Tes Uraian/Esei

Untuk memperoleh butir soal yang baik, maka beberapa petunjuk berikut ini diharapkan membantu:

- (1) Butir soal mengukur hasil belajar yang penting saja. Hasil belajar yang remeh (*trivial*) tidak perlu diujikan. Misalnya:

Lemah : jumlah bayi yang meninggal sebelum mencapai usia enam tahun di kecamatan ini tahun lalu adalah ...

Lebih baik : Di Kecamatan ini jumlah bayi yang meninggal sebelum mencapai usia lima tahun dalam dua tahun terakhir adalah ... untuk setiap seribu penduduk

- (2) Butir soal mengandung permasalahan yang bersifat spesifik. Butir soal itu haruslah menjamin bahwa hanya peserta tes yang menguasai isi pelajaran yang dapat menjawab soal itu dengan baik. Misalnya:

Lemah : Daun tembakau mengandung ...

Lebih baik : Bahan yang berbahaya bagi kesehatan yang terdapat pada daun tembakau adalah ....

- (3) Butir soal mengharuskan peserta memberi jawaban yang secara faktual benar

Lemah : Orang merokok akan ...

Lebih baik: Kebiasaan merokok akan menyebabkan penyakit ...

- (4) Butir soal dengan menggunakan bahasa yang jelas dan tidak mengandung arti yang mendua. Misalnya;

Lemah : Ibukota Kuwait yang diduduki Irak adalah

Lebih baik : <sup>....</sup> Ibukota Kuwait adalah ....

- (5) Bila yang ditanyakan menyangkut angka atau jumlah dan satu satuan tertentu, maka sebaiknya nyatakan satuan dalam soal. Misalnya:

Lemah : Seorang anak usia 12 tahun sebaiknya setiap hari minum susu ...

Lebih baik : Seorang anak usia 12 tahun sebaiknya setiap hari minum susu murni ... gelas

- (6) Setiap butir soal sebaiknya hanya berisi satu jawaban yang harus dikerjakan oleh peserta tes. Misalnya:

Lemah : Suatu provinsi terbagi menjadi beberapa ..., yang selanjutnya dibagi lagi menjadi beberapa ..., dan kemudian dibagi lagi menjadi beberapa ..., dan akhirnya unit terkecil disebut ....

Lebih baik : Provinsi Jawa Barat dibagi menjadi ... kabupaten dan kotamadya

- (7) Pergunakanlah kata-kata yang menuntut jawaban yang singkat dan tertentu. Jawaban itu haruslah satu kata, satu frasa, sebuah angka, atau sebuah simbol. Misalnya:

Lemah : Disebut apakah binatang pemakan binatang lain dan tumbuh-tumbuhan?

Lebih baik : Termasuk klasifikasi apakah binatang pemakan binatang lain dan tumbuh-tumbuhan?

(8) Jangan menggunakan kalimat yang langsung diambil dari buku atau dari catatan. Penggunaan kalimat yang langsung diambil dari buku atau catatan cenderung mendorong peserta didikan menghafal “mati”, tanpa berusaha memahami apa yang dipelajarinya. Untuk menghindari kelemahan itu maka sebaiknya bahan ajaran yang diambil dari buku tersebut disusun kembali dalam kalimat yang mudah dipahami oleh siswa.

(9) Jangan sampai pertanyaan yang diajukan menjadi tes bahasa, sedangkan maksudnya untuk menguji materi pelajaran lain. Misalnya:

Lemah : Apakah istilah yang digunakan untuk menyatakan Colombus ke Benua Amerika tahun 1492?

Lebih baik : Siapakah yang menemukan benua Amerika tahun 1492?

(10) Untuk menanyakan istilah atau defenisi sebaiknya digunakan kalimat tanya secara langsung. Kalimat yang lain mendahului kalimat tanya, yang dimaksudkan untuk menjelaskan pertanyaan, namun dapat menyebabkan pertanyaan menjadi kabur. Misalnya:

Lemah : Setiap siswa harus mentaati peraturan sekolah. Ketaatan kepada aturan sekolah itu dalam P4 apa namanya?

Lebih baik : Sikap yang menganjurkan memelihara kebersihan kelas sesuai dengan sila keberapa dari Pancasila?

- (11) Dalam menanyakan masalah perhitungan, guru harus menentukan tingkat ketepatan, terutama untuk angka desimal. Apakah cukup angka bulat saja, atau berapa angka di belakang koma. Misalnya:

Lemah : Berapakah 10:6?

Lebih baik : Berapakah 10:6 (bulatkan sampai dua angka di belakang koma)?

- (12) Sebaiknya hanya satu jawaban untuk satu pertanyaan. Misalnya:

Lemah : Siapakan presiden Republik Indonesia?

Lebih baik : Siapakan presiden pertama Republik Indonesia?

### C. Konstruksi Tes Objektif

#### 1) Butir soal benar-salah

Ada lima persyaratan yang harus dipenuhi oleh setiap butir tipe soal benar-salah untuk dapat dikatakan sebagai butir soal yang baik, yaitu:

- a) Setiap butir tes harus menguji atau mengukur hasil belajar peserta tes yang penting dan bermakna, tidak menanyakan hal yang remeh (*trivial*). Misalnya:

Lemah : B-S Bung Hatta dilahirkan di Bukittinggi

Lebih baik : B–S Pemikiran Bung Hatta tentang hak asasi manusia telah diabadikan dalam pasal-pasal UUD 1945

Mempersoalkan tempat kelahiran Bung Hatta tidak ada sangkut pautnya dengan peranan bung Hatta dalam perjuangan mencapai kemerdekaan Indonesia. Sedangkan pikiran dan pendapat Bung Hatta tentang hak asasi manusia mempunyai arti sejarah yang amat penting bagi Bangsa Indonesia

- b) Setiap butir soal haruslah menguji pemahaman, tidak hanya pengukuran terhadap daya ingat. Butir soal tidak dianjurkan untuk menguji kemampuan mengingat kata atau frasa yang terdapat dalam buku ajar atau bacaan lainnya. Misalnya:

Lemah : B–S Hukum Newton I menyatakan bahwa setiap benda akan bergerak lurus beraturan atau diam, jika tidak ada tekanan gaya yang bekerja pada benda itu

Lebih baik : B–S Penumpang bis yang duduk tenang dalam bis yang berjalan dengan kecepatan 80 km per jam akan terdorong kedepan bila bis diberhentikan secara tiba-tiba

Dalam butir soal yang pertama di atas jelas hanya menguji kemampuan menghafal bunyi hukum Newton I (Hukum Kelembaman), yang bila dapat dijawab oleh peserta tes dengan benar, penguji belum dapat kepastian bahwa siswa

mengerti akan hukum kelembaman tersebut. Tetapi dengan soal berikutnya penguji akan dapat memastikan bahwa siswa tahu tentang hukum kelembaman dan dapat memahaminya dengan baik.

- c) Kunci jawaban yang ditentukan haruslah benar. Patokan ini kedengarannya remeh, tetapi seringkali kita jumpai bahwa kunci yang ditentukan oleh pembuat soal berbeda dari yang diartikan oleh peserta tes atau perview tes. Hal itu seringkali disebabkan adanya subjektifitas atau bias penyusun tes masuk kedalam soal dan dalam penentuan kunci jawaban. Misalnya:

Lemah : B-S Sebelum dilakukan pernikahan calon pengantin laki-laki diharuskan melamar calon pengantin wanita.

Lebih baik : B-S Dalam masyarakat patrinal, pihak calon pengantin pria diharapkan lebih mengambil inisiatif dari pada pihak pengantin wanita

Dalam soal yang pertama, kunci yang diharapkan adalah B. tetapi, kunci tersebut tidak berlaku bagi masyarakat matrinal. Karena itu haruslah secara eksplisit dinyatakan dalam butir soal tentang kondisi yang menentukan kunci butir soal tersebut.

- d) Butir soal yang baik haruslah jelas jawabannya bagi seorang peserta tes yang belajar, dan jawaban yang salah kelihatan lebih seakan-akan benar bagi peserta tes yang tidak belajar dengan baik. Jadi butir soal tersebut dapat secara jelas

membedakan orang yang belajar dan orang yang tidak belajar. Misalnya:

B-S Makanan kaleng lebih mahal harganya daripada makanan segar. (S)

B-S Bahasa ilmiah yang digunakan di pesantren di Jawa Barat pada awal abad ke-20 adalah bahasa Arab dan Bahasa Jawa. (B)

Bagi peserta tes yang tidak pernah belajar tentang prinsip ekonomi tentu akan menyangka bahwa makanan kaleng akan lebih mahal dari makanan segar, karena harga kemasan itu sendiri yang telah menyebabkan harga bahan makanan tersebut menjadi tinggi. Sedangkan karena jumlah permintaan makanan kaleng, menyebabkan harga makanan segar menjadi lebih mahal. Demikian pula halnya dengan bahasa ilmiah di pesantren di Jawa Barat. Bagi orang yang tidak tahu sejarah pendidikan dengan baik akan menyangka bahwa bahasa Sunda-lah yang digunakan dalam buku-buku pelajaran utama di pesantren di daerah Sunda. Jadi jawaban itu akan menarik bagi orang yang tidak belajar.

- e) Pernyataan dalam butir soal harus dinyatakan secara jelas dan menggunakan bahasa yang baik dan benar. Soal yang jelas itu bila dalam soal tersebut hanya mempersoalkan satu gagasan saja. Selain itu pernyataan satu gagasan itu haruslah disusun dalam tata kalimat yang baik dan benar dan tidak mengandung pengertian mendua. Karena tidak boleh mengandung pengertian mendua itu, maka butir soal itu

harus dinyatakan dalam kalimat yang sesingkat mungkin. Jangan menggunakan kalimat beranak yang tidak jelas. Untuk dapat menyatakan butir soal itu secara jelas maka sebaiknya tidak menggunakan kata-kata negatif yang mengaburkan arti. Misalnya:

Lemah : B-S Kekalahan Jerman terhadap Sekutu dalam Perang Dunia II bukan disebabkan oleh ketidakmampuan Jerman dalam strategi memenangkan pertempuran tetapi lebih disebabkan oleh kelemahan semangat perang rakyat Jerman.

Lebih baik : B-S Hilangnya semangat perang rakyat Jerman adalah penyebab utama kekalahan Jerman terhadap Sekutu dalam Perang Dunia II

Untuk lebih meningkatkan mutu butir soal tipe benar-salah, beberapa pertimbangan berikut sebaiknya digunakan:

- a) Jumlah butir soal yang kuncinya S (salah) sebaiknya lebih banyak dari butir soal yang kunci jawabannya B (benar)
- b) Susunlah kalimat soal sedemikian rupa sehingga logika sederhana akan cenderung mengarah ke jawaban yang salah. Misalnya:

B - S Bilasebuahbendaberatnya1kgterapung setengahnya di atas permukaan air akan tenggelam bila seandainya benda itu beratnya 1,5 kg

Dengan logika sederhana, maka bila suatu benda terapung setengahnya, untuk tenggelam akan membutuhkan tambahan berat setengah kali berat yang sekarang. Jadi dibutuhkan berat 1,5 kg. Sedangkan yang benar adalah dibutuhkan tambahan berat 1 kg lagi.

- c) Susunlah jawaban salah sesuai dengan anggapan umum yang salah tentang suatu kenyataan. Misalnya:

B – S Semua benda yang sama beratnya akan jatuh dengan kecepatan yang sama dari ketinggian yang sama.

- d) Pergunakan rujukan untuk beberapa buah soal, misalnya dengan menggunakan teks atau gambar sebagai rujukan untuk butir soal

2) Butir soal menjodohkan

Beberapa prinsip konstruksi butir soal menjodohkan:

- a) Pernyataan di bawah kolom pertama dan dibawah kolom kedua masing-masing haruslah terdiri dari kelompok yang homogen. Misalnya:

Lemah:

Kolom Pertama

Kolom Kedua

1. Ir. Soekarno

A. Pemerataan hasil pembangunan

2. Pertanian

B. Makhluk laut yang melahirkan

3. Persaingan bebas

C. Proklamator kemerdekaan

4. Ikan Paus

D. Demokrasi liberal

E. Prioritas pembangunan

Lebih baik:

Kolom Pertama

Kolom Kedua

1. Ir. Soekarno

A. Bapak Koperasi Indonesia

2. Suharto

B. Bapak Palang Merah

3. Moh. Hatta

C. Bapak Pramuka Indonesia

4. Sultan

D. Bapak Pembangunan

Hamengkubuwono

E. Bapak Revolusi Indonesia

b) Pernyataan di bawah kolom kedua harus lebih banyak dari pernyataan dibawah kolom pertama. Untuk memudahkan penyediaan lembaran jawaban yang seragam, maka dianjurkan supaya jumlah pernyataan dibawah kolom pertama berkisar antara 3 atau 4 buah. Sedangkan pernyataan di bawah kolom kedua adalah 5. Dengan demikian lembaran jawaban akan seragam dengan bentuk butir soal pilihan ganda lainnya.

3) Butir soal pilihan ganda biasa

Berikut ini dikemukakan beberapa prinsip pokok dalam konstruksi butir soal tipe pilihan ganda biasa.

a) *Saripati permasalahan harus ditempatkan pada pokok soal (stem)*. Inti permasalahan dalam butir soal tersebut harus dicantumkan dalam rumusan pokok soal, sehingga dengan membaca pokok soal, mahasiswa sudah dapat menentukan jawaban sebelum dilanjutkan membaca pilihan jawaban. Persyaratan ini tidak berlaku bagi pengembangan butir soal kesusastaan.

*Contoh:*

Yang kurang baik,

Pulau Jawa adalah pulau yang....

- A. menghasilkan banyak minyak
- B. penduduknya terpadat
- C. dijadikan objek wisata
- D. menghasilkan banyak rempah-rempah

Yang lebih baik,

Pulau yang terpadat penduduknya di Indonesia adalah pulau ...

- A. Sumatera
- B. Jawa
- C. Kalimantan
- D. Sulawesi

Contoh yang kedua lebih baik dari contoh pertama karena dengan membaca pokok soal, siswa sudah dapat membuat jawaban sebelum membaca pilihan A, B, C dan D.

b) *Hindari pengulangan kata-kata yang sama dalam pilihan*

Peniadaan pengulangan kata berarti menyangkut waktu menulis dan membaca serta menghemat tempat.

*Contoh:*

Yang kurang baik,

Penulisan contoh 2 diatas diubah menjadi:

Pulau yang terpadat penduduknya di Indonesia adalah ....

- A. pulau Sumatera
- B. pulau Jawa
- C. pulau Kalimantan
- D. pulau Sulawesi

c) *Hindari rumusan kata yang berlebihan.* Tidak selalu penjelasan terinci mempermudah

pengertian, justru dapat membingungkan dan mengaburkan pengertian. Yang penting rumusan yang baik yang berisi, padat dan jelas tanpa kata-kata “kembang”.

*Contoh:*

Kalau butir soal 2 ditambah rumusannya menjadi:

Pulau yang terpadat penduduknya di Indonesia sehingga sukar untuk meningkatkan produksi pangan adalah pulau ....

- A. Sumatera
- B. Jawa
- C. Kalimantan
- D. Sulawesi

Tambahan kata-kata “sehingga sukar untuk meningkatkan produksi pangan” membuat pengertian pernyataan menjadi kabur dan kalimat yang harus dibaca menjadi lebih panjang.

- d) Kalau pokok soal merupakan pernyataan yang belum lengkap, maka kata atau kata-kata yang melengkapi harus diletakkan pada ujung pernyataan, bukan di tengah-tengah kalimat.

*Contoh:*

Yang kurang baik,

Menurut De Bakey, ... adalah penyebab penyakit penyempitan pembuluh darah.

- A. kolesterol
- B. kelebihan berat
- C. merokok
- D. tekanan batin

Yang lebih baik,

Menurut De Bakey, penyakit penyempitan pembuluh darah disebabkan oleh ....

- A. kolesterol
- B. kelebihan berat

- C. merokok  
D. tekanan batin
- e) *Susunan alternative jawaban dibuat teratur dan sederhana.* Cara menyusun alternatif jawaban dibuat berderet dari atas ke bawah. Kalau yang dideretkan itu terdiri dari satu kata, urutan ke bawah dibuat berdasarkan alfabet, kalau yang dideretkan bilangan, urutan ke bawah berdasarkan bilangan yang makin bertambah besar atau makin menurun, atau diurutkan panjang kalimat.
- f) *Hindari penggunaan kata-kata teknis atau ilmiah atau istilah yang aneh dan mentereng.* Perlu diingat bahwa tes yang dikembangkan bertujuan untuk mengukur materi pelajaran, kalau materi tersebut tidak menyangkut perbendaharaan, janganlah menggunakan istilah teknik dan aneh.  
*Contoh:*  
Yang kurang baik,  
Apakah kritik utama ahli psikologi terhadap tes?  
A. tes menimbulkan *anxiety*  
B. tes selalu disertai *cultural bias*  
C. tes hanya mengukur hal-hal yang *trivial*  
D. tes tergantung pada kemampuan kognitif guru  
Yang lebih baik,  
Apakah kritik utama ahli psikologi terhadap tes?  
A. tes menimbulkan rasa cemas  
B. tes sangat bergantung pada nilai budaya tertentu  
C. tes mengukur hasil belajar yang tidak penting  
D. tes sangat ditentukan oleh pengetahuan guru
- g) *Semua pilihan jawaban harus homogen dan dimungkinkan sebagai jawaban benar.* Ciri

husus pilihan ganda dari tes objektif lain adalah pada pilihan ganda semua alternatif jawaban ada kemungkinan sebagai jawaban yang benar, sehingga siswa terpaksa membaca dan memikirkan semua pilihan dan menentukan mana yang paling tepat untuk menjawab pertanyaan tersebut. Hindari pengecoh yang dengan melihat sepintas, siswa sudah dapat menentukan pengecoh tersebut tidak ada sangkutannya dengan pokok soal atau pengecoh tersebut adalah jawaban yang tidak masuk akal.

*Contoh:*

Yang kurang baik,

Siapakah pahlawan wanita dari Aceh?

- A. R.A Kartini
- B. Imam Bonjol
- C. Cut Nya Dien
- D. Teuku Umar

Yang lebih baik,

Siapakah pahlawan wanita dari Aceh?

- A. R.A Kartini
- B. Dewi Sartika
- C. Cut Nya Dien
- D. Cut Keke

*h) Hindari keadaan dimana jawaban yang benar selalu ditulis lebih panjang dari jawaban yang salah. Ada kecenderungan siswa memilih jawaban yang lebih panjang dan yang lebih terinci sebagai jawaban yang benar. Oleh karena itu, penulis soal berusaha agar pengecoh dan jawaban yang benar ditulis sama panjang dengan rincian yang sama pula.*

*i) Hindari adanya petunjuk/indikator pada jawaban yang benar.*

*Contoh:*

Yang kurang baik,

Agar air panas dalam teko tidak cepat dingin, maka teko tersebut dibungkus dengan ....

- A. kain
- B. seng
- C. tembaga
- D. timah

Pilihan B, C dan D termasuk logam, A bukan logam.

Yang lebih baik,

Air panas akan bertahan panas jika disimpan dalam bejana yang terbuat dari ....

- A. aluminium
- B. keramik
- C. plastic
- D. seng

j) *Hindari menggunakan pilihan yang berbunyi "semua yang di atas benar" atau "tidak satupun diatas yang benar".* Adanya pilihan semacam ini sebenarnya mengurangi jumlah alternatif pilihan karena kalau siswa sudah mengenal satu atau dua di antara empat pilihan sebagai jawaban ketiga, siswa tersebut akan memilih "semua yang diatas benar". Hal yang sama berlaku untuk "tidak satupun yang di atas benar".

k) *Gunakan tiga atau lebih alternatif pilihan.* Kalau hanya ada dua pilihan, bentuk ini sama dengan bentuk salah-benar. Dua pilihan berarti tebakannya tinggi sedangkan kalau lima pilihan faktor tebakan menurun yaitu 20 persen. Banyaknya pilihan yang disediakan sangat ditentukan oleh usia peserta tes dan juga tergantung dari sifat bahan yang disajikan.

- l) Pokok soal diusahakan tidak menggunakan ungkapan atau kata-kata yang bermakna tidak tentu. Misalnya kebanyakan, seringkali, kadang-kadang dan sejenisnya.

Contoh:

Yang kurang baik.

Kebanyakan hewan hidupnya di dalam air, bernafas dengan ...

- A. insang
- B. kulit
- C. paru-paru
- D. insang dan paru-paru

Yang lebih baik

Belut bernafas dengan ....

- A. insang
- B. kulit
- C. paru-paru
- D. insang dan paru-paru

- m) Pokok soal sedapat mungkin dalam pernyataan atau pertanyaan **positif**. Jika terpaksa menggunakan pertanyaan negatif, maka kata negative tersebut digarisbawahi atau ditulis tebal atau ditulis dengan huruf miring atau *italic*.

Contoh:

Yang kurang baik,

Pada semua tumbuhan yang berdaun hijau, fotosintesis *tidak* akan terjaditanpa ....

- A. udara, tanah dan air
- B. cahaya, udara dan tanah
- C. air, cahaya dan udara
- D. air, tanah, dan cahaya

Yang lebih baik,

Pada semua tumbuhan yang berdaun hijau, fotosintesis akan terjadi bila terdapat ....

- A. udara, tanah dan air
- B. cahaya, udara dan tanah
- C. air, cahaya dan udara
- D. air, tanah, dan cahaya

4) Pilihan ganda analisis hubungan antar hal

Butir soal hubungan antar hal terdiri dari dua pernyataan. Kedua pernyataan ini dihubungkan oleh kata "SEBAB". Jadi ada dua kemungkinan hubungan kedua pernyataan tersebut yaitu *ada* hubungan sebab akibat atau *tidak ada* hubungan sebab akibat. Supaya kedua pernyataan ini termasuk pilihan ganda maka harus dicari variabel lain yang dapat mengukur kemampuan siswa. Variabel tersebut adalah kualitas pernyataan yaitu apakah pernyataan pertama *benar* atau *salah* dan apakah pernyataan kedua *benar* atau *salah*. Dengan adanya berbagai hal yang harus dinilai dari dua pernyataan tersebut maka dapatlah dikembangkan tes bentuk Hubungan Antar Hal dengan petunjuk penyelesaian sebagai berikut:

Untuk soal-soal berikut ini, pilihlah:

- A. Jika kedua pernyataan benar dan merupakan hubungan sebab akibat
- B. Jika pernyataan pertama dan kedua benar tetapi tidak menunjukkan hubungan sebab akibat
- C. Jika salah satu dari kedua pernyataan tersebut salah
- D. Jika kedua pernyataan salah

Contoh soal

Frekuensi detak nadi seseorang yang baru berlari cepat akan naik

SEBAB

Pada waktu lari cepat denyut jantung bertambah cepat

5) Butir soal pilihan ganda analisis kasus

Siswa dihadapkan pada suatu kasus. Kasus ini disajikan dalam bentuk ceritera, peristiwa dan sejenisnya. Kepada siswa tersebut diajukan beberapa pertanyaan. Setiap pertanyaan dibuat dalam bentuk Melengkapi Pilihan.

Bentuk soal tes analisis kasus menggunakan petunjuk yaitu untuk menjawab butir soal berikut ini disediakan satu kasus. Anda diminta memahaminya secara cermat, kemudian jawablah soal-soal berikutnya.

Contoh:

*"Kadit Lantas Polda Jatim Letkol Pol. Dra. Watie Soemarsono menjelaskan jumlah kecelakaan lalu lintas di Jatim pada bulan Januari – November 1986 adalah sebanyak 7090 kasus atau meningkat 4,87 persen disbanding tahun 1985 periode yang sama. Meningkatnya kecelakaan lalu lintas itu antara lain karena terhentinya Operasi Zebra menjadi operasi rutin lalu lintas. Di samping itu, pengguna jalan hanya berdisiplin jika ada petugas".*

- (1) Meningkatnya kecelakaan lalu lintas di Jatim bukan hanya disebabkan oleh terhentinya Operasi Zebra tetapi juga disebabkan oleh ....
- A. pengawas lalu lintas yang tidak pernah kendor

- B. volume kendaraan di jalan makin bertambah
  - C. angkutan yang terlibat dalam pengaturan lalu lintas dikurangi jumlahnya
  - D. potensi polisi lalu lintas belum dikerahkan secara maksimal
- (2) Dalam periode yang sama di tahun 1986 kasus kecelakaan lalu lintas di Jatim naik 4,87 persen dibandingkan dengan tahun 1985. Dari data ini ....
- A. dapat dibuat kesimpulan bahwa kecelakaan lebih banyak terjadi pada tahun 1985 dibandingkan dengan tahun 1986.
  - B. belum dapat dibuat kesimpulan tentang jumlah kecelakaan karena kasus kecelakaan bulan Desember belum dilaporkan
  - C. sudah dapat dipastikan bahwa kasus kecelakaan selama tahun 1956 lebih banyak terjadi daripada tahun 1985.
  - D. tidak dapat dibuat generalisasi karena banyak kasus kecelakaan yang tidak dilaporkan yang bersangkutan kepada pihak polisi lalu lintas.

Dari contoh ini dapat dikatakan bahwa dari satu kasus dapat dibuat beberapa pertanyaan dan aspek berpikir yang diukur lebih tinggi dari pemahaman.

## 6) Pilihan ganda kompleks

Bentuk soal pilihan ganda kompleks ini biasa juga disebut Melengkapi Berganda atau Asosiasi Pilihan Ganda. Struktur pertanyaan ini sama dengan Melengkapi Pilihan. Perbedaannya adalah, kalau pada

Melengkapi Pilihan hanya ada satu jawaban yang benar atau paling benar, tetapi pada Melengkapi Berganda justru jawaban yang benar dapat lebih dari satu, mungkin 2, 3 atau 4. Jadi pada ragam melengkapi berganda diperbolehkan menuliskan keempat alternatif pilihan jawaban yang benar, tidak ada pengecoh. Dengan kata lain jika *semua alternatif pilihan benar, janganlah* dimasukkan dalam ragam melengkapi pilihan tetapi harus dimasukkan ke dalam ragam melengkapi berganda. Adapun petunjuk untuk melengkapi berganda adalah ....

Untuk soal berikut pilihlah:

- A. Jika (1) dan (2) benar
- B. Jika (1) dan (3) benar
- C. Jika (2) dan (3) benar
- D. Jika semuanya benar

*Contoh:*

Salah satu vitamin yang larut dalam lemak adalah vitamin A yang terdapat dalam ....

- (1) minyak ikan dan telur
- (2) bayam, ikan dan telur
- (3) air susu dan wortel

7) Pilihan ganda menggunakan diagram, gambar, grafik atau tabel

Bentuk soal ini mirip dengan Analisis Kasus baik struktur maupun pola pertanyaannya. Bedanya dalam tes bentuk ini tidak disajikan kasus dalam bentuk cerita atau peristiwa tetapi kasus tersebut berupa diagram, gambar, grafik dan tabel.

Contoh:

Tabel X di bawah ini menggambarkan rata-rata suhu dan curah hujan di kota X

Udara (°C)	Jan	Feb	Mar	Apr	Mei	Jun	Jul	Agt	Sep	Okt	Nop	Des
Suhu Udara (°C)	28,9	29,9	31,1	29,9	29,1	28,6	27,9	28,1	28,9	28,7	28,4	28,6
Curah hujan (mm)	1,0	4,0	23,0	86,0	27,0	0,0	0,0	1,0	2,0	42,0	38,0	8,0

Pertanyaan: Manakah yang benar untuk kota X?

- A. bulan yang terpanas suhu udaranya adalah bulan yang sedikit curah hujannya
- B. setiap tahun selalu turun hujan di kota X
- C. terjadi dua kali musim hujan dalam setahun di kota X
- D. waktu yang paling baik untuk menanam padi di kota X adalah pada bulan Juni

Tes pilihan ganda dengan semua variasinya selalu dapat menggunakan diagram, gambar, grafik dan tabel baik sebagai pokok soal (*stem*) maupun sebagai alternatif pilihan. Biasanya butir soal yang menggunakan diagram, gambar, grafik dan tabel dapat mengukur aspek proses berpikir yang lebih tinggi dari aspek ingatan



## PENGADMINISTRASIAN TES

Yang dimaksud dengan pengadministrasian tes dalam kaitan pemabahasan ini adalah pelaksanaan tes mulai dari proses penyuntingan naskah tes, sampai dengan mengerjakan tes. Langkah-langkah itu meliputi penyuntingan tes, pengadaan naskah tes dan pelaksanaan tes.

### **A. Penyusunan Perangkat Tes**

Penyusunan suatu perangkat tes yang akan digunakan harus mempertimbangkan dua hal utama:

#### **1. Penyuntingan Naskah Tes**

Suatu naskah tes terdiri dari beberapa butir soal. Penyusunan butir tes tersebut sehingga menjadi suatu perangkat tes haruslah mempertimbangkan beberapa hal yang memungkinkan peserta tes dapat mengerahkan kemampuan terbaiknya dalam

mengerjakan tes tersebut. Hal ini berarti bahwa pertimbangan utama dalam penyuntingan tes adalah peserta tes. Untuk itu hendaklah diperhatikan hal-hal sebagai berikut:

- a. Tes bentuk objektif sebaiknya tidak dilaksanakan secara lisan
- b. Butir tes sebaiknya disusun mulai dari pokok bahasan yang dibahas paling awal ke pokok bahasan yang dibahas terakhir.
- c. Tingkat kesukaran disusun mulai yang termudah meningkat terus sampai kepada yang sukar, dalam arti bahwa butir soal yang mudah diletakkan pada awal naskah sedangkan butir soal yang sukar diletakkan pada akhir naskah.
- d. Butir tes yang setipe hendaknya di kelompokkan dalam satu kelompok. Jadi jangan sampai ada satu tipe tes tersebar di beberapa kelompok. Misalnya tes pilihan ganda biasa dicampurkan dengan pilihan ganda sebab akibat, dan sebaiknya. Setiap bagian tes haruslah berisi satu tipe tes dan dikumpulkan dalam satu kelompok.
- e. Tulislah petunjuk pengerjaan tes secara jelas, sehingga tidak seorangpun perlu bertanya lagi tentang cara mengerjakan tes tersebut atau bertanya tentang apa yang harus dilakukan. Petunjuk tes ini sangat besar perannya bagi keberhasilan peserta tes. Selain itu petunjuk tes juga akan turut menentukan apakah pengukuran hasil belajar peserta tes dapat dilakukan secara akurat atau tidak. Suatu petunjuk tes minimal harus mencantumkan: (1) apa yang harus dilakukan oleh peserta tes, (2) bagaimana peserta tes harus mengerjakan tes tersebut, dan (3) dimana jawaban tes harus ditulis. Karena

itu setiap perangkat tes harus mempunyai: (1) petunjuk umum yang menjadi pedoman mengerjakan keseluruhan perangkat tes, dan (2) petunjuk khusus yang merupakan pedoman untuk mengerjakan satu kelompok tes tertentu. Dalam petunjuk itu, sebaiknya disertakan juga contoh mengerjakan tes tersebut (terutama untuk tes yang penting, yaitu tes yang akan menentukan keputusan bagi individu). Di samping itu, dalam petunjuk tes itu harus tercantum tentang cara peserta tes untuk mengubah jawabannya. Misalnya bila ia pada mulanya memilih jawaban tertentu, tetapi kemudian ia ingin mengubah jawaban tersebut, haruslah dijelaskan cara yang harus ditempuhnya.

- f. Penyusunan butir tes tersebut hendaknya diatur sedemikian rupa sehingga tidak menimbulkan kesan berdesak-desak. Setiap butir tes hendaklah diatur sehingga memudahkan peserta tes untuk membacanya.
- g. Susunlah setiap butir tes sehingga *stem* dan seluruh optionnya terletak dalam satu halaman yang sama
- h. Letakkanlah wacana atau petunjuk soal yang digunakan sebagai rujukan bagi suatu atau beberapa butir tes di atas butir tes yang bersangkutan
- i. Hindarilah meletakkan kunci jawaban dalam suatu pola tertentu

## **2. Penggandaan Naskah Tes**

Setelah naskah tes selesai disunting, langkah berikutnya adalah penggandaan naskah tes. Tentu saja prosedur penggandaan ini harus dapat

menjamin kerahasiaan naskah tes, di samping adanya jaminan hasil penggandaan tersebut tidak akan mengganggu konsentrasi peserta tes dalam melaksanakan tes.

Penggandaan tes sebaiknya terpisah antara lembaran jawaban tes dari lembaran jawaban. Pemilihan ini akan lebih memudahkan peserta tes dalam menentukan tempat menjawab tes yang diberikan. Sedangkan bagi guru hal itu sangat memudahkan dalam penskoran, karena tidak perlu digalaukan oleh lembaran tes yang dapat mengganggu dalam penskoran

Beberapa petunjuk praktis dalam penggandaan tes, sebagai berikut:

- a. Antar butir tes harus cukup tersedia ruang atau batas, sehingga tidak kesan berdesak-desak.
- b. Angka atau huruf yang disediakan di depan alternatif jawaban hendaknya sepenuhnya sama dengan angka atau huruf yang digunakan dalam lembaran jawaban
- c. Untuk tipe tes menjodohkan, maka kedua kolom yang berisi tes atau alternatif jawaban itu harus terletak dalam satu halaman yang sama.
- d. Butir tes yang menggunakan wacana, maka butir tes yang berhubungan dengan wacana tersebut harus terletak dalam halaman yang sama.
- e. Semua wacana, grafik, diagram dan gambar yang digunakan sebagai landasan bagi butir tes harus terjamin kejelasannya, keakuratannya, dan keterbacaannya
- f. Kalau naskah digandakan dalam jumlah yang banyak, maka harus terjamin setiap naskah sama jelasnya. Jangan sampai terjadi sebagian

peserta tes mendapat naskah yang kurang baik dibandingkan dengan peserta tes lainnya.

## **B. Pelaksanaan Tes**

Pengadministrasian tes juga harus mempertimbangkan berbagai cara pelaksanaan tes. Cara pelaksanaan tersebut antara lain: (1) tes catatan tertutup atau tes catatan terbuka, (2) tes diumumkan atau tes dirahasiakan (mendadak), (3) tes lisan atau tertulis, dan (4) tes tindakan (praktek). Dalam bagian ini hanya akan dibahas kekuatan dan keterbatasannya.

### **1. Catatan Terbuka dan Catatan Tertutup (*open books dan close books*)**

Dalam melaksanakan tes hasil belajar seorang pengajar mempunyai hak penuh untuk menentukan apakah para peserta tes boleh melihat buku/catatan dan menggunakan berbagai alat belajar seperti tabel, kamus, kalkulator, dan sebagainya atau tidak. Boleh ataupun tidak boleh, keduanya mempunyai keuntungan/kekuatan dan kelemahan/keterbatasannya.

Kekuatan-kuatannya jika peserta tes diizinkan menggunakan buku atau catatan atau alat-alat belajar yang lain adalah:

- a. Para siswa tidak terlalu tegang pikirannya pada soal, menghadapi atau pada saat melaksanakan ujian
- b. Para siswa akan “bertanya” kepada buku atau catatan atau alat belajar lain yang dimilikinya ketimbang menyontek pekerjaan temannya.

- c. Para siswa akan terbiasa membuat catatan yang sebaik-baiknya dan atau memiliki buku-buku atau alat belajar lainnya karena mengetahui betul manfaatnya.
- d. Para siswa akan terbiasa membaca buku atau catatan atau berlatih menggunakan tabel, kalkulator dan sejenisnya karena terasa benar-benar manfaatnya, yakni kelak kalau tes tidak akan mengalami kesulitan dalam menggunakannya.

Sebaliknya, keterbatasan-keterbatasannya dapat berupa:

- a. Para siswa mungkin saja malas membaca buku atau catatan dengan alasan dalam ujian akan bebas melihat buku atau catatan.
- b. Mereka yang jarang membuka buku/catatan akan habis waktu ujiannya untuk mencari/membolak-balik lembaran buku untuk mendapatkan jawaban.
- c. Ada kecenderungan para siswa malas berpikir, hal yang sangat mudah pun dicari jawabannya di dalam buku.
- d. Bagi siswa yang alat kelengkapan belajarnya minimal akan dirugikan.

Tes yang tidak mengizinkan para peserta tes membuka buku atau catatan mempunyai keuntungan/kekuatan antara lain:

- a. Membiasakan para siswa untuk memahami isi buku atau catatan yang dimilikinya sebab jika tidak, akan tidak berhasil dalam ujian
- b. Membiasakan para siswa untuk berpikir sendiri, bukan menggantungkan diri kepada buku atau catatan yang ada.

- c. Membiasakan para siswa membuat rangkuman mengenai isi buku atau catatan yang dipelajarinya.

Keterbatasan dari tes yang tidak mengijinkan membuka buku atau catatan antara lain:

- a. Mendorong siswa akan melihat pekerjaan temannya (nyontek) apabila sudah betul-betul tak berhasil menemukan jawabannya
- b. Siswa belum tentu terlatih menggunakan buku atau catatan sebagai sumber belajar
- c. Kaburnya prinsip bahwa buku itu untuk digunakan, bukan untuk dihafal. Bahkan dalam kehidupan nyata kelak buku itu memang untuk digunakan. Seorang jaksa akan mempelajari dan membukabuka KUHP menjelang mengajukan tuntutan; seorang insinyur sipil akan melihat tabel-tabel perhitungan konstruksi baja pada saat harus menghitung konstruksi sebuah jembatan; seorang dokter akan melihat buku patologi pada saat akan menentukan diagnosis pasien yang ditanganinya, dan masih banyak contoh-contoh lain.

## **2. Tes Diumumkan dan Dirahasiakan**

Pelaksanaan/pengadministrasian tes dapat dilakukan dengan memberi pengumuman lebih dahulu atau tanpa pemberitahuan sebelumnya. Pada umumnya para ahli psikologi belajar atau psikologi pendidikan tidak dapat menyetujui adanya tes yang pelaksanaannya tidak diumumkan sebelumnya (dirahasiakan). Tetapi pelaksanaan tes yang dirahasiakan ini masih dapat memenuhi tujuan tes tertentu, karena ia mempunyai beberapa kelebihan, seperti:

- a. Dapat mengukur pengetahuan siap yang dimiliki oleh siswa
- b. Dapat memotivasi/meningkatkan usaha belajarnya secara terus-menerus, karena jadwal tes tidak diketahui dengan pasti waktu pelaksanaannya.
- c. Dapat digunakan sebagai alat peningkatan disiplin belajar.

Sifat rahasia atau terbuka itu tidak hanya diterapkan pada pengadministrasian tes tetapi juga pada hasil tes. Kekuatan dan keterbatasan hasil tes yang diumumkan itu antara lain:

- a. Peserta tes yang lulus, apalagi kalau nilainya bagus akan menjadi bangga karena diketahui oleh teman-temannya. Dan hal ini akan memacu untuk belajar lebih baik lagi atau sekurang-kurangnya mempertahankannya.
- b. Terjadi semacam dilayani secara layak dan perasaan dihargai
- c. Tumbuh kepercayaan para siswa kepada lembaga pendidikan di mana mereka belajar, khususnya kepada guru yang bersangkutan bahwa tes beserta penilainnya dilakukan secara objektif. Kepercayaan ini makin kuat lagi apabila hasil pekerjaan tes dikembalikan kepada mereka, dan guru yang bersangkutan memberi kesempatan kepada para siswa untuk membahas kunci jawabannya dan mempersoalkan nilai mereka
- d. Pihak guru tentu akan mengoreksi dan memberi nilai kepada setiap pekerjaan peserta tes dengan cermat karena tidak ingin kepercayaan siswa kepada dirinya hilang karena kecerobohnya.

Keterbatasan-keterbatasan hasil tes yang diumumkan adalah:

- a. Membuat malu siswa yang tidak lulus atau nilainya rendah yang pada gilirannya akan menghapuskan motivasi belajarnya
- b. Kesempatan untuk demokratis seperti yang diutarakan pada keuntungan butir c di atas dapat saja cenderung menjadi protes-protes
- c. Guru yang terlambat atau tidak dapat mengumumkan tepat waktu akan merasa mempunyai beban mental yang berat dan memang dapan menjerumuskan kepada cemoohan oleh para siswa
- d. Memerlukan kemampuan administrasi yang prima dan memerlukan fasilitas serta dana tambahan

Jika hasil tes tidak diumumkan (dirahasiakan), maka kekuatan-kekuatannya antara lain adalah:

- a. Tidak menuntut kemampuan administratif yang prima dan mahal
- b. Tidak akan terjadi protes-protes dari pihak peserta tes yang akan merepotkan para guru maupun lembaga pendidikan yang bersangkutan
- c. Jika dipandang perlu, nilai seorang peserta tes dapat diputuskan dengan mengikutsertakan faktor-faktor non-tes, kerajinan misalnya.

Keterbatasan-keterbatasan hasil tes yang tidak diumumkan (dirahasiakan) adalah:

- a. Tes itu tidak ada atau kurang berguna karena tidak komunikatif dengan para siswa atau orang tua siswa yang bersangkutan. Padahal hasil tes belajar itu berfungsi dan bermanfaat

jika dikomunikasikan dengan pihak-pihak yang bersangkutan.

- b. Dapat saja terjadi seorang guru itu “main hakim sendiri” tanpa diketahui oleh siapapun.
- c. Para siswa tidak merasakan hasil jerih payahnya padahal hasil yang diperoleh ini memberi motivasi yang sangat penting dalam proses belajar.

### **3. Tes Lisan dan Tertulis**

Kekuatan tes tertulis antara lain:

- a. Kemampuan memilih kata-kata, kekayaan informasi, kemampuan berbahasa, kemampuan memilih ataupun memadukan ide-ide dan proses berpikir peserta tes dapat dilihat dengan nyata.
- b. Kemampuan-kemampuan seperti yang disebutkan pada butir a di atas dapat dibandingkan antara yang satu dengan yang lain.
- c. Dalam waktu yang relatif terbatas dapat dilaksanakan tes yang terdiri atas sejumlah besar peserta tes sehingga ekonomis
- d. Memungkinkan dikoreksi oleh lebih dari seorang korektor (jika tesnya esei) sehingga lebih objektif

Sedangkan keterbatasan-keterbatasannya antara lain adalah:

- a. Khusus untuk tes bentuk esei, tes tertulis itu menuntut tugas peserta tes yang terlalu berat.
- b. Dalam hal bentuk tes esei khususnya, maka ketunabahasaan akan merugikan peserta tes yang bersangkutan apabila masalah bahasa diperhitungkan di dalam memberi nilai.
- c. Yang bersifat massal itu biasanya kurang baik dibandingkan dengan yang individual

- d. Siswa cenderung menuliskan jawabannya berpanjang-panjang sehingga jawaban tersebut malah menyimpang dari persoalannya, hal ini tidak dapat dikontrol oleh guru. Di samping itu karena asyiknya terpaku pada salah satu butir, akhirnya siswa lupa waktu sehingga pada saat waktu tes habis peserta tes yang bersangkutan belum beranjak ke butir soal tes yang lain

Bagi tes lisan kekuatan-kekuatannya antara lain:

- a. Dapat dilaksanakan secara individual sehingga lebih cermat dan dapat dilakukan "*probing*" sehingga penguji dapat mengetahui persis dimana posisi hasil belajar peserta didik yang bersangkutan.
- b. Kemampuan-kemampuan seperti yang ada pada tes tertulis yang telah diuraikan di atas dapat dipantau secara langsung oleh guru yang mengujinya
- c. Dengan tes lisan memungkinkan terjadinya komunikasi dua arah dan dialog aktif. Hal ini mendorong siswa menyiapkan diri sebaik-baiknya
- d. Siswa dapat mengungkapkan argument-argumennya secara lebih bebas sehingga guru yang menguji mengetahui persis jalan pikiran siswa

Sementara itu keterbatasan-keterbatasan tes lisan antara lain:

- a. Tidak ekonomis
- b. Jika yang melaksanakannya hanya seorang, jadi satu lawan satu maka dapat terjadi subyektivitas yang sukar di kontrol
- c. Memungkinkan guru "*main hakim sendiri*", bahkan dendam pribadi dapat dilampiaskan disitu

- d. Bagi peserta tes yang “gagap” atau “groggi” dirugikan oleh sistem ujian lisan ini

#### **4. Tes Tindakan atau Praktek**

Bagi tes tindakan atau praktek, kekuatan-kekuatannya antara lain:

- a. Terjadinya pengecekan terhadap terbentuk atau tidaknya keterampilan yang dirumuskan di dalam TIK
- b. Membuat pergantian suasana sehingga kejenuhan dapat dikurangi atau dihilangkan

Sementara itu keterbatasan-keterbatasan tes tindakan atau praktek antara lain:

- a. Tidak semua bahan dapat diujipraktekan
- b. Mahal dan guru dituntut lebih mampu dari siswanya yang hal ini mungkin tidak dapat dipenuhi, terutama dalam bidang olahraga
- c. Jika praktiknya tidak dalam keadaan yang sesungguhnya maka siswa cenderung main-main, atau kalau mereka juga sungguh-sungguh, maka kurang manfaatnya karena praktek dalam kehidupan sehari-hari tidak sama dengan situasi praktek “tiruan”. Dalam praktek tiruan ini siswa umumnya justru kikuk; jadi tidak berlangsung wajar.

## **PENGOLAHAN HASIL TES**

Pengolahan hasil tes merupakan kegiatan lanjutan pengadministrasian ujian, yaitu memeriksa hasil ujian dengan mencocokkan jawaban peserta dengan kunci jawaban untuk tes kognitif dan tes keterampilan.

Dalam uraian berikut ini akan dibicarakan bagaimana mengolah hasil tes kognitif dan tes keterampilan sehingga diperoleh skor setiap peserta tes. Dari skor inilah diturunkan nilai mereka.

### **A. Mengolah Hasil Tes Kemampuan Berpikir**

Untuk mengukur kemampuan berpikir dapat dikembangkan tes bentuk objektif atau tes bentuk uraian baik terbuka maupun uraian terbatas. Mengolah tes bentuk objektif lebih cepat, sederhana, dan mudah dibandingkan dengan tes bentuk uraian. Berikut diuraikan prosedur memeriksa hasil tes.

### **a. Tata Cara Memeriksa Hasil Tes Objektif**

Hasil tes objektif dapat diperiksa dengan menggunakan lembar kunci jawaban, sedangkan lembar kunci jawaban itu ada beberapa jenis. Beberapa jenis lembar kunci jawaban dan cara menggunakannya akan diuraikan di bawah ini.

#### **(1) Kunci Jawaban Berdampingan**

Format lembar kunci jawaban ini sama benar dengan lembar jawaban, artinya untuk lembar kunci jawaban digunakan blangko lembar jawaban yang digunakan siswa. Kemudian blangko lembar jawaban itu diisi dengan kunci-kunci setiap soal, yang lazimnya tersusun dalam lajur dari atas ke bawah. Lembar kunci jawaban itu diletakkan berdampingan dengan lembar jawaban siswa, sehingga dengan cepat akan diketahui nomor-nomor soal yang dijawab benar atau salah, yaitu dengan mencocokkan jawaban tiap nomor soal pada lembar jawaban dengan kunci jawaban. Nomor-nomor soal yang dijawab salah dibubuhi tanda silang, sehingga untuk lembar jawaban itu jumlah soal yang dijawab benar dapat dihitung. Jumlah soal yang dijawab benar itu merupakan skor mentah

#### **(2) Kunci Berkarbon**

Format lembar jawaban untuk pemeriksaan dengan kunci berkarbon ini berbeda dengan kunci berdampingan. Dengan kunci berkarbon ini di sebelah kanan tiap nomor soal disediakan kotak-kotak sebanyak alternatif jawaban yang tersedia. Siswa diminta membubuhkan tanda silang di dalam kotak di bawah pilihan yang diminati. Format lembar kunci jawaban telah terisi

lingkaran-lingkaran yang merupakan kunci dalam kotak setiap nomor soal. Lembar kunci jawaban itu diletakkan diatas lembar jawaban dengan disisipi karbon diantaranya, kemudian ketiganya dijepit dengan penjepit kertas. Kemudian lingkaran-lingkaran pada kunci jawaban itu ditindis sampai tindasnya itu dapat membekas pada lembar jawaban. Setelah satu lembar jawaban selesai diperiksa seperti itu, penjepit kertas dilepaskan. Tampaklah pada lembar jawaban, nomor-nomor soal yang dijawab benar tanda silangnya terlingkari dengan lingkaran hasil tindasan karbon, sedangkan nomor-nomor soal yang dijawab salah lingkarannya akan membekas di kotak yang berbeda. Dengan demikian jumlah jawaban benar pada lembar jawaban itu dapat dihitung.

(3) Kunci Tusukan

Format lembar jawaban untuk pemeriksaan dengan kunci tusukan ini sama dengan kunci berkarbon. Bedanya adalah dengan kunci tusukan ini pemeriksaan tidak dilakukan dengan karbon. Pada lembar kunci jawaban tidak dibubuhi lingkaran pada kotak jawaban yang benar, tetapi dibubuhi silang pada lembar jawaban. Lembar kunci jawaban diletakkan diatas lembar jawaban, kemudian dijepit dengan penjepit kertas. Kotak-kotak yang bertanda silang pada kunci jawaban ditusuk dengan jarum sampai tusukan itu membekas pada lembar jawaban. Jawaban-jawaban yang benar pada lembar jawaban akan tertusuk, sedangkan jawaban-jawaban yang salah letak tusukannya jatuh di kotak yang berbeda.

## (4) Kunci Berjendela

Format lembar jawaban untuk pemeriksaan dengan kunci berjendela ini sama pula dengan kunci berkarbon. Lembar kunci jawaban dibuat dari blangko lembar jawaban. Kunci setiap nomor soal dilubangi pada kotak yang sesuai. Lembar kunci jawaban itu diletakkan diatas lembar jawaban, kemudian dijepit dengan penjepit kertas. Melalui lubang-lubang (jendela) pada kunci jawaban itu dibubuhkan garis datar yang membekas pada lembar jawaban. Setelah pemeriksaan satu lembar jawaban selesai, penjepit kertas dilepaskan, maka pada lembar jawaban tampak garis-garis datar yang terletak sekotak dengan tanda silang dan ada pula yang tidak sekotak. Garis-garis datar yang sekotak dengan tanda silang itu menunjukkan jawaban yang benar, sedangkan yang tidak sekotak adalah jawaban yang salah. Kemudian jumlah jawaban yang benar dihitung untuk menentukan skornya. Berikut ini diuraikan masing-masing cara memberi skor pada tes objektif.

## (a) Soal Benar-Salah

Apabila penskoran menggunakan sistem denda atau jawaban yang salah diperhitungkan dalam skor, maka cara penskorannya adalah:

$$S = R - W$$

dengan pengertian:

S = Skor yang diperoleh

R = *Right* (jawaban yang benar)

W = *Wrong* (jawaban yang salah)

*Contoh:*

Jumlah soal tes = 20 buah

Si A menjawab betul 16 buah dan salah 4 buah.

Maka skor untuk *Si A* adalah:

$$16 - 4 = 12$$

Dengan menggunakan rumus seperti ini maka ada kemungkinan seorang siswa memperoleh skor negatif.

Apabila penskoran dilakukan tanpa pembobotan, artinya setiap soal diberi bobot sama dan tanpa denda maka cara penskoran yang paling sederhana adalah:

$$S = R$$

yang dihitung hanya yang betul

(untuk soal yang tidak dikerjakan dinilai 0)

Kadang-kadang untuk beberapa soal tertentu diberikan bobot yang berbeda dengan yang lain.

Yang dimaksudkan dengan bobot di sini adalah salah satu faktor pengganda yang dikenakan pada soal-soal yang diberi bobot (*B*) itu. Oleh karena itu, untuk soal-soal yang berbobot, cara memberi skornya adalah:

$$S = B \times R \text{ atau } S = B \times (R - W)$$

### (b) Soal Menjodohkan

Untuk macam soal ini dapat ditempuh cara penskoran seperti bentuk soal diatas, yaitu cara  $S = R$  atau  $S = R - W$ . Namun jika faktor terkaan diperhitungkan, maka untuk suatu kelompok soal penskorannya menjadi:

$$S = R - \frac{R}{(n_1 - 1)(n_2 - 1)}$$

Dengan  $n_1$  menyatakan banyaknya pernyataan di lajur sebelah kiri dan  $n_2$  menyatakan banyaknya pilihan jawaban di sebelah kanan

**(c) Soal Pilihan Ganda**

Apabila penskorannya hanya didasarkan atas jawaban yang benar, maka cara yang paling sederhana adalah:

$$S = R$$

Namun kadang-kadang dijumpai pula penskoran yang didasarkan atas jawaban yang benar dan yang salah. Untuk ini penskorannya menjadi:

$$S = R - W$$

Ada pula cara penskoran yang memperhitungkan faktor terkaan. Untuk ini penskorannya adalah:

$$S = R - \frac{W}{O - 1}$$

S = skor yang diperoleh

R = jawaban yang betul

W = jawaban yang salah

O = banyaknya opsi

1 = bilangan tetap

*Contoh:*

Seorang murid menjawab 17 betul dari 20 soal. Soal bentuk *multiple choice* ini menggunakan option sebanyak 4 buah.

$$\text{Skor} = 17 - \frac{3}{4-1} = 16$$

### **b. Tata Cara Memeriksa Hasil Tes Uraian/Esei**

Pada waktu pengembangan butir tes uraian dianjurkan agar pertanyaan dibatasi atau dibuat

tertutup sehingga jawaban untuk soal tersebut tidak memberi kemungkinan jawaban yang bermacam-macam. Dengan demikian, tes uraian hendaknya dilengkapi dengan pedoman penilaian atau kunci jawaban sehingga penilaiannya lebih mengarah pada penilaian objektif. Pedoman penilaian yang dibuat penyusun butir soal inilah yang akan digunakan sebagai acuan dalam memeriksa lembar jawaban ujian uraian. Namun sebelum menggunakan pedoman penilaian tersebut perlu diperiksa lebih dahulu apakah jawaban yang diminta oleh penulis soal sudah sesuai dengan tingkat penguasaan peserta tes. Pengalaman menunjukkan bahwa tidak jarang jawaban yang diharapkan penulis soal terlalu banyak, atau terlalu sulit atau keduanya. Bila terlalu banyak konsep yang diharapkan penulis dalam jawaban dan pedoman penilaian tidak ditinjau kembali dapat mengakibatkan banyak peserta tes tidak lulus. Di samping itu, jika terlalu sulit arti konsep yang diharapkan (sebagai konsekuensi dari proses berpikir yang terlalu berat) maka pedoman penilaian perlu diujicobakan.

Dalam pemberian skor tes esei, dikenal ada dua cara yaitu *metode analisis* dan *metode klasifikasi atau memilah*. Pada metode analisis terlebih dahulu dipersiapkan suatu model jawaban bagi setiap butir soal. Pada model jawaban itu jawaban diuraikan beberapa langkah atau unsur, yang setiap unsur jawabannya disediakan skor tertentu. Dengan model jawaban itu pekerjaan tiap siswa dibandingkan dan setiap unsur jawaban diberi skor sesuai dengan tingkat kebenarannya.

Pada metode klasifikasi tidak perlu dipersiapkan model jawaban. Pekerjaan tiap siswa dibaca dan ditelaah secara keseluruhan. Setelah jawaban untuk satu soal selesai ditelaah, lembar pekerjaan itu dikumpulkan dan diklasifikasikan menjadi tumpukan *baik sekali*, *baik*, *sedang*, *kurang* dan *kurang sekali*. Lembar pekerjaan untuk nomor soal yang telah ditelaah itu dapat diberi skor, misalnya 95 untuk tumpukan baik sekali, 75 untuk yang baik, 55 untuk yang sedang, 35 untuk yang kurang dan 15 untuk yang kurang sekali. Setelah itu pemeriksaan dapat dilanjutkan untuk nomor soal berikutnya, dan dilakukan proses yang sama dengan nomor soal sebelumnya. Demikian seterusnya sampai semua soal tertelaah. Jadi dengan metode klasifikasi ini lembar pekerjaan siswa akan diperiksa nomor soal demi nomor soal, bukan siswa demi siswa.

Selain dari kedua cara penskoran di atas, pedoman berikut ini perlu mendapat perhatian.

- 1) Ambil lembar jawaban ujian (LJU) peserta sekitar 10%. Pengambilan dilakukan secara acak. Bilamana Anda telah mengetahui kemampuan peserta misalnya siswa yang Anda ajar sendiri, sehingga Anda mengetahui benar siapa yang pandai di kelas itu, siapa saja yang sedang tingkat kemampuannya, maka pemilihan secara acak tidak diperlukan, tetapi pilih pekerjaan dari salah seorang peserta tes yang pandai, dua pekerjaan dari peserta yang kemampuannya sedang dan pekerjaan dari seorang yang kemampuannya kurang

- 2) Sebelum mulai memberi skor, persiapkan terlebih dahulu model jawaban untuk setiap butir soal, kemudian tentukanlah skor yang akan diberikan untuk setiap butir soal. Jika menggunakan metode analisis, tentukan juga skor untuk setiap langkah tiap butir soal. Jika menggunakan metode klasikal, tentukanlah skor untuk setiap tingkat klasifikasi tiap butir soal.
- 3) Setiap lembar pekerjaan hendaknya diperiksa tanpa memperhatikan identitas siswa, kalau perlu lembar pekerjaan itu diberi sandi. Sebab dengan mengetahui identitas siswa dapat menurunkan keobjektifan pemberian skornya
- 4) Periksalah lembar pekerjaan itu nomor soal demi nomor soal. Misalnya suatu tes esei terdiri atas sepuluh nomor soal, maka periksalah nomor satu dulu untuk seluruh siswa, kemudian baru dilanjutkan dengan soal nomor dua untuk seluruh siswa, dan seterusnya sampai pada soal nomor sepuluh. Cara ini dimaksudkan untuk meningkatkan keobjektifan penskoran, atau menurunkan kesubjektifan.

## **B. Mengolah Hasil Tes Kemampuan Sikap (Afektif)**

Pengukuran ranah afektif tidaklah semudah mengukur ranah kognitif. Pengukuran ranah afektif tidak dapat dilakukan setiap saat (dalam arti pengukuran formal) karena perubahan tingkah laku siswa tidak dapat berubah sewaktu-waktu. Pengubahan sikap seseorang memerlukan waktu yang relative lama. Demikian juga pengembangan minat dan penghargaan terhadap nilai-nilai. Adapun tujuan penilaian afektif adalah:

- a. Untuk mendapatkan umpan balik (*feedback*) baik bagi guru maupun siswa sebagai dasar untuk memperbaiki proses belajar mengajar dan mengadakan program perbaikan (*remedial program*) bagi anak didiknya.
- b. Untuk mengetahui tingkat perubahan tingkah laku anak didik yang dicapai yang antara lain diperlukan sebagai bahan bagi perbaikan tingkah laku anak didik, pemberian laporan kepada orang tua, dan penentuan lulus tidaknya anak didik.
- c. Untuk menempatkan anak didik dalam situasi belajar mengajar yang tepat, sesuai dengan tingkat pencapaian dan kemampuan serta karakteristik anak didik.
- d. Untuk mengenal latar belakang kegiatan belajar dan kelainan tingkah laku anak didik.

Sehubungan dengan tujuan penilaiannya ini, maka yang menjadi sasaran penilaian kawasan afektif adalah *perilaku* anak didik, bukan *pengetahuannya*. Sebagai contoh, siswa bukan dituntut untuk mengetahui sebab-sebab terbentuknya BPUPKI, tetapi bagaimana sikapnya terhadap pembentukan BPUPKI tersebut.

Pertanyaan afektif tidak menuntut jawaban benar atau salah, tetapi jawaban yang khusus tentang dirinya mengenai minat, sikap, dan internalisasi nilai. Sebelum melakukan penilaian terhadap aspek afektif, sama halnya dengan mengukur aspek kognitif, guru diharapkan mendaftarkan materi yang dicakup dihubungkan dengan TIU dan TIK-nya.

Ada beberapa bentuk skala yang dapat digunakan untuk mengukur sikap, antara lain:

1) *Skala Likert*

Skala ini disusun dalam bentuk suatu pernyataan dan diikuti oleh lima respons yang menunjukkan tingkatan. Misalnya:

- SS = Sangat Setuju
- S = Setuju
- TB = Tidak berpendapat
- TS = Tidak Setuju
- STS = Sangat Tidak Setuju

Pertanyaan:

"Amir tidak mau berhenti bekerja sebelum pekerjaan selesai	SS	S	TB	TS	STS
------------------------------------------------------------	----	---	----	----	-----

2) *Skala Pilihan Ganda*

Skala ini bentuknya seperti soal bentuk pilihan ganda yaitu suatu pernyataan yang diikuti sejumlah alternatif pendapat.

*Contoh:*

Dalam suatu upacara bendera:

- a) Setiap peserta harus dengan khidmat mengikuti jalannya upacara tanpa kecuali
- b) Peserta diperbolehkan berbicara asal dalam batas-batas tertentu dan tidak mengganggu jalannya upacara
- c) Dalam keadaan terpaksa peserta boleh berbicara tetapi hanya dengan berbisik
- d) Peserta boleh (bebas) berbicara asal tertib

### 3) Skala Thurstone

Skala Thurstone merupakan skala mirip skala buatan Likert karena merupakan suatu instrumen yang jawabannya menunjukkan tingkatan

1	2	3	4	5	6	7	8	9	10
A	B	C	D	E	F	G	H	I	J
very favourable			neutral				very unfavourable		

Pernyataan yang diajukan kepada responden disarankan oleh Thurstone kira-kira 10 butir, tetapi tidak kurang dari 5 butir.

### 4) Skala Guttman

Skala ini mengandung tiga atau empat buah pernyataan yang masing-masing harus dijawab “ya” atau “tidak”. Pernyataan-pernyataan tersebut menunjukkan tingkatan yang berurutan sehingga bila responden setuju pernyataan nomor 2, diasumsikan setuju juga nomor 1.

Contoh:

1. Pengguguran kandungan dapat diterima bila ibu dalam keadaan bahaya
2. Pengguguran kandungan dapat diterima bila ternyata bayi dalam kandungan mengalami cacat serius.
3. Pengguguran kandungan dapat diterima bila sebagai perkosaan
4. Pengguguran kandungan dapat diterima bila untuk alasan KB
5. Pengguguran kandungan dapat diterima apapun alasannya

Bila responden setuju dengan pernyataan nomor 3 misalnya, maka keseluruhan pernyataan nomor 1 dan 2 juga dianggap disetujui. Dan dianggap pasti tidak setuju dengan pernyataan nomor 4 dan nomor 5.

5) *Semantic differential*

Instrumen ini mengukur konsep-konsep untuk tiga dimensi. Dimensi-dimensi yang ada diukur dalam kategori: baik-tidak baik, kuat-lemah dan cepat-lambat atau aktif-pasif, atau dapat juga berguna-tidak berguna.

Contoh:

Main musik								
Baik	1	2	3	4	5	6	7	Tidak baik
Berguna	1	2	3	4	5	6	7	Tidak berguna
Aktif	1	2	3	4	5	6	7	pasif

cara ini dapat digunakan untuk mengetahui minat dan pendapat siswa mengenai sesuatu kegiatan atau topik dari suatu mata pelajaran.

6) *Pengukuran Minat*

Di samping skala seperti yang dicontohkan di atas, minat juga dapat diukur dengan cara seperti di bawah ini:

A. Mengunjungi perpustakaan : SS S TS STS

B. Sandiwara : SS S TS STS

Pilihan: *senang sampai dengan sangat tidak senang dapat ditentukan sendiri seberapa suka*

### **C. Mengolah Hasil Tes Kemampuan Keterampilan Fisik (Psikomotor)**

Pengukuran kemampuan keterampilan fisik atau psikomotorik dilakukan terhadap hasil-hasil belajar yang berupa penampilan. Namun demikian biasanya pengukuran kemampuan (ranah) ini disatukan atau dimulai dengan pengukuran ranah kognitif sekaligus. Misalnya penampilannya dalam menggunakan thermometer diukur mulai dari pengetahuan mereka mengenai alat tersebut, pemahaman tentang alat dan penggunaannya (aplikasi), kemudian cara menggunakannya dalam bentuk keterampilan. Untuk pengukuran yang terakhir ini harus diperinci antara lain: cara memegang, cara meletakkan/menyelipkan ke dalam ketiak atau mulut, cara membaca angka, cara mengembalikan ke dalam tempatnya dan sebagainya. Ini semua tergantung dari kehendak kita, asal tujuan pengukuran dapat tercapai.

Instrumen yang digunakan mengukur keterampilan biasanya berupa matriks. Perhatikan contoh berikut:

*Instrumen untuk mengamati keterampilan  
praktek memasak (dalam skala 5)*

Nama: ..... Kelas: .....

No	Keterampilan	Skor				
		1	2	3	4	5
1.	Terampil menyiapkan alat					
2.	Tekun dalam bekerja					X
3.	Menggunakan waktu dengan efektif			X		
4.	Mampu bekerja sama					
5.	Memperhatikan keselamatan kerja		X	X		
6.	Memperhatikan kebersihan			X		
7.	Hasil makanan enak				X	X

Keseluruhan hasil sesuai dengan skor yang diperoleh.  
Untuk A ini skornya adalah:

$$\frac{5 + 3 + 2 + 3 + 3 + 5}{7} = \frac{25}{7} = 3,57$$

# Bab Tujuh

## ANALISIS HASIL TES

Setelah lembar jawaban atau lembar pekerjaan siswa selesai diperiksa, maka setiap lembar jawaban itu telah mendapatkan skor. Skor-skor itu masih perlu dianalisis lebih lanjut, oleh karena itu skor itu lazim disebut *skor mentah*.

Pada uraian bab ini banyak dipergunakan istilah *skor* (skor mentah) dan *nilai* (jadi). Istilah skor dan nilai dipergunakan dalam pengertian yang berbeda. Skor adalah jumlah jawaban betul siswa yang diperoleh langsung dari kerja koreksi. Sebaliknya, nilai adalah angka ubahan dari skor yang diperoleh dari kerja analisis atau pengolahan skor dengan mempergunakan kriteria tertentu (skor yang telah diolah).

Jika kita akan mengubah skor mentah ke nilai jadi, maka masalah pertama yang timbul adalah standar (pendekatan) manakah yang akan digunakan? Standar yang dapat dipergunakan sebenarnya cukup banyak, namun yang sering

disebut dan dipergunakan orang adalah *standar mutlak* (PAP) dan *standar relatif* (PAN). Berikut akan diuraikan kedua standar penilaian tersebut.

## **A. Penilaian Acuan Patokan (PAP)**

Penilaian Acuan Patokan (*criterion referenced evaluation*) yang dikenal pula dengan sebutan standar mutlak, berusaha menafsirkan hasil tes yang diperoleh siswa dengan membandingkannya dengan patokan yang telah ditetapkan. Sebelum hasil tes itu sendiri diperoleh dan bahkan sebelum kegiatan pengajaran dilakukan, patokan yang akan dipergunakan untuk menentukan batas kelulusan itu telah ditetapkan.

Standar “kelulusan” tersebut berisi ketentuan-ketentuan yang dipergunakan untuk menentukan batas kelulusan siswa, atau batas untuk memberikan “nilai” tertentu kepada siswa. Jika skor seorang siswa memenuhi batas minimal untuk lulus, atau untuk tingkatan nilai yang lain, siswa yang bersangkutan dinyatakan lulus karena dianggap telah memenuhi tingkat penguasaan minimal terhadap bahan pelajaran. Demikian pula sebaliknya jika seorang siswa tak memenuhi kriteria batas kelulusan, ia akan dinyatakan tidak lulus karena belum memenuhi tingkat penguasaan minimal yang disyaratkan. Oleh karena batas kelulusan tersebut bersifat baku, tidak boleh ditawar-tawar, standar penilaian ini disebut standar mutlak.

Dalam salah satu harian ibukota terdapat satu informasi mengenai kesempatan bagi pemuda Indonesia untuk menjadi penerbang dengan syarat

antara lain:

- (1) Pria
- (2) Warga negara Indonesia
- (3) Tinggi badan minimal 165 cm
- (4) Panjang kaki minimal 100 cm

Syarat diatas merupakan patokan pada waktu mengajukan lamaran. Bilamana salah satu dari patokan itu tidak terpenuhi lebih baik tidak usah melamar. Khusus mengenai tinggi badan dan panjang kaki persyaratan minimalnya adalah 165 cm dan 100 cm. Mereka yang melebihi ukuran tersebut dapat diterima lamarannya dan lulus, sebaliknya mereka yang kurang dari ukuran tersebut tidak berhasil atau tidak lulus.

Dalam proses penilaian terdapat juga pendekatan yang menggunakan pola diatas yaitu menentukan kelulusan seseorang ditentukan sejumlah patokan. Bilamana seseorang telah memenuhi patokan tersebut ia dinyatakan berhasil atau lulus atau telah menguasai bahan tersebut. Sebaliknya bila seseorang belum memenuhi patokan, ia dikatakan gagal atau belum menguasai bahan tersebut. Patokan dalam proses pembelajaran selalu mengacu pada tujuan instruksional umum dan tujuan instruksional khusus. Keberhasilan seseorang dalam pembelajaran ditentukan oleh tingkat penguasaan tujuan instruksional. Dengan kata lain nilai atau kelulusan seseorang ditentukan oleh penguasaan tujuan instruksional. Jadi berbeda dengan penilaian acuan norma dimana nilai atau kelulusan seseorang ditentukan oleh kelompoknya.

Berhubung standar penilaian ditentukan secara mutlak, banyaknya siswa yang lulus dan memperoleh nilai tinggi akan mencerminkan prestasi siswa sekaligus mencerminkan penguasaannya terhadap bahan pelajaran. Sebagai akibat logis penggunaan standar mutlak ini, dimungkinkan sekali terjadi sebagian besar siswa dalam satu kelompok lulus dengan nilai tinggi, sebagian besar siswa tidak lulus karena nilainya dibawah standar minimal, atau jumlah yang mendapat nilai tinggi dan rendah kurang lebih sama. Hasil tes yang demikian jika digambarkan dalam bentuk kurva akan berwujud kurva juling positif, kurva juling negatif, dan kurva normal.

Penafsiran hasil tes yang mempergunakan pendekatan PAP dilakukan dengan membandingkan antara skor hasil tes yang diperoleh dengan patokan yang telah ditetapkan sebelumnya. Akan tetapi, kriteria yang dipergunakan untuk menetapkan besarnya patokan itu sendiri belum ada kesepakatan yang disetujui oleh semua pihak. Belum ada kesepakatan tentang besarnya angka batas minimal kelulusan, batas minimal penguasaan bahan atau pencapaian tujuan (dalam persentase) dan batas minimal untuk memberikan nilai tertentu.

Seorang guru (dosen) misalnya, menentukan batas minimal kelulusan bagi siswa (mahasiswa)-nya sebesar 60 persen, artinya siswa dinyatakan lulus jika mampu mengerjakan 60 persen butir soal yang tersedia dengan betul. Guru (dosen) yang lain misalnya, diatas atau dibawah 60 persen. Dalam hal

ini tiap sekolah biasanya telah menetapkan patokan yang berlaku di sekolahnya masing-masing.

Penentuan patokan tidak hanya menyangkut penetapan batas minimal kelulusan dan pemberian nilai tertentu, melainkan juga penetapan skala penilaian yang dipergunakan. Kedua hal tersebut memang saling berkaitan. Skala penilaian yang dipergunakan mungkin skala lima (0 – 4 atau E – A), skala 10 (1 – 10), skala sebelas (0 – 10) atau skala 100 (0 – 100). Skala mana yang harus dipergunakan oleh seorang guru adalah tergantung pada ketentuan yang berlaku di sekolahnya.

Penentuan batas minimal kelulusan dan pemberian nilai tertentu dapat dilakukan dengan perhitungan persentase. Artinya seorang siswa dinyatakan lulus jika ia mampu mengerjakan dengan betul “sekian” persen butir soal yang disediakan. Jika butir-butir soal itu telah mencerminkan seluruh bahan pelajaran, kemampuan siswa mengerjakan “sekian” persen tersebut sekaligus mencerminkan tingkat penguasaannya terhadap bahan pelajaran itu.

Perhitungan patokan dengan penghitungan persentase cukup sederhana dan mudah dilakukan, tanpa memerlukan prosedur penghitungan yang rumit. Di bawah ini dicontohkan penentuan patokan dengan penghitungan persentase untuk skala lima dan sepuluh (seratus).

Tabel 1. Contoh Penentuan Patokan dengan Penghitungan Persentase untuk Skala Lima

Interval Persentase Tingkat Penguasaan	Nilai Ubah Skala Lima		Keterangan
	0 - 4	E - A	
85% - 100%	4	A	Baik sekali
75% - 84%	3	B	Baik
60% - 74%	2	C	Cukup
40% - 59%	1	D	Kurang
0% - 39%	0	E	Kurang Sekali

Sebagai contoh perhitungan, misalnya seorang siswa memperoleh skor 62 dari 75 butir soal yang tersedia. Hal itu berarti ia mampu mengerjakan 83 persen (62 dibagi 75 kali 100), dan jika ditransformasi ke skala lima, ia akan memperoleh nilai 3 atau B.

Tabel 2 Contoh Penentuan Patokan dengan Penghitungan Persentase untuk Skala Sepuluh

Interval Persentase Tingkat Penguasaan	Nilai Ubahan Skala Sepuluh	Keterangan
96% - 100%	10	Sempurna
86% - 95%	9	Baik Sekali
76% - 85%	8	Baik
66% - 75%	7	Cukup
65% - 65%	6	Sedang
46% - 55%	5	Hampir Sedang
36% - 45%	4	Kurang
26% - 35%	3	Kurang Sekali
16% - 23%	2	Buruk
0% - 15%	1	Buruk Sekali

Sebagai contoh penghitungan, siswa yang memperoleh skor 62 dari 75 butir soal yang disediakan di atas mampu mengerjakan 83 persen. Tingkat penguasaan siswa tersebut berada dalam interval 76 persen – 85 persen, dan setelah diubah ke dalam skala sepuluh ia memperoleh nilai 8. Jika kita mempergunakan skala seratus, hasil penghitungan persentase tersebut sekaligus menunjukkan nilai siswa.

## **B. Penilaian Acuan Norma (PAN)**

Penilaian Acuan Norma (*Norm Referenced Evaluation*) dikenal pula dengan sebutan standar relatif atau norma kelompok. Pendekatan penilaian ini menafsirkan hasil tes siswa yang lain dalam kelompoknya. Alat pembandingan yang dijadikan standar kelulusan dan pemberian nilai tertentu ditentukan berdasarkan skor yang diperoleh siswa dalam satu kelompok. Dengan demikian berbeda halnya dengan PAP, standar kelulusan baru dapat ditentukan setelah diperoleh skor para siswa.

Hal itu berarti standar yang dibuat berdasarkan (dan untuk) satu kelompok tak dapat diterapkan untuk kelompok yang lain. Standar yang dibuat berdasarkan hasil tes sebelumnya pun tidak dapat dipergunakan untuk hasil tes yang sekarang dan mendatang. Setiap kali memperoleh data hasil tes, kita dituntut untuk membuat norma yang baru. Jika kita bandingkan, norma yang satu dengan lain akan berbeda, ada yang “tinggi” tuntutannya, ada yang cukup, dan ada yang rendah. Kesemuanya itu tergantung prestasi siswa kelompok yang bersangkutan. Jika kelompok

tertentu kebetulan para siswanya pandai, norma kelulusan akan menjadi tinggi. Demikian sebaliknya jika suatu kelompok kebetulan rata-rata siswanya tergolong bodoh. Itulah sebabnya pendekatan penilaian ini disebut standar relatif.

Pendekatan PAN berdasar diri pada asumsi distribusi normal. Oleh karena mean dan simpangan baku dihitung berdasarkan skor hasil tes, distribusi skor selalu cenderung normal, walau kadar kenormalan tidak selalu sama untuk tiap kelompok. Dengan demikian, walau sama-sama menghasilkan kurva normal di antara kelompok-kelompok tersebut, mean kurva yang satu berbeda dengan mean kurva yang lain. Sebagai konsekuensinya, seorang siswa yang memperoleh nilai tinggi dalam satu kelompok mungkin akan memperoleh nilai rendah jika dimasukkan ke kelompok lain. Demikian juga sebaliknya, misalnya seorang siswa yang memperoleh skor 62 mendapat nilai 8 dalam kelompoknya, tetapi jika dibandingkan ke kelompok lain mungkin hanya memperoleh nilai 6.

*Contoh:*

1. Satu kelompok anak didik terdiri dari 9 orang mendapat skor (nilai mentah)  
50, 45, 45, 40, 40, 40, 35, 35, 30  
Dari skor ini dapat dibaca bahwa perolehan nilai tertinggi adalah 50 dan perolehan terendah 30. Dengan demikian nilai tertinggi diberikan kepada skor tertinggi, misalnya nilai 10, secara proporsional skor diatas dapat diberi nilai 10, 9, 9,8  
Cara lain dengan menghitung persentase jawaban benar yang dijawab oleh setiap orang. Kemudian kepada persentase tertinggi diberikan

nilai tertinggi. Jika nilai mentah diatas didapat dari 60 butir pertanyaan atau skor maksimumnya 60, maka:

Tabel 3. Pengolahan Nilai Mentah Menjadi Nilai (1-10)

a.	Nilai mentah	50	45	45	40	40	40	35	35	30
b.	Persentase yang betul	83,3	75,0	75,0	66,7	66,7	66,7	58,3	58,3	50,0
c.	Nilai (1 – 10)	10	9	9	8	8	8	7	7	6

2. Sekelompok siswa terdiri dari 40 orang dalam satu ujian memperoleh nilai mentah sebagai berikut:

55	43	39	38	37	35	34	32
52	43	40	37	36	35	34	30
49	43	40	37	36	35	34	28
48	42	40	37	35	34	33	22
46	39	38	37	36	34	32	21

Penyebaran nilai tersebut dapat ditulis sebagai berikut:

Tabel 4 Pengolahan Nilai Mentah Menjadi Nilai (1 – 10)

No	Nilai Mentah	Jumlah Mahasiswa	Jika 55 diberi 10 maka	Jika skor maks 75 maka % yang benar	Persentase diubah menjadi nilai (1 – 10)
1	2	3	4	5	6
1.	55	1	10,0	73,3	10,0
2.	52	1	9,5	69,3	9,5
3.	49	1	8,9	65,3	8,9
4.	48	1	8,7	64,0	8,7
5.	46	1	8,4	61,3	8,4
6.	43	3	7,8	57,3	7,8
7.	42	1	7,6	56,0	7,6
8.	40	3	7,3	53,3	7,3
9.	39	2	7,1	52,0	7,1
10.	38	2	6,9	50,7	6,9
11.	37	5	6,7	49,3	6,7
12.	36	4	6,5	48,0	6,5
13.	35	3	6,4	46,7	6,4
14.	34	4	6,2	45,3	6,2
15.	33	2	6,0	44,0	6,0
16.	32	2	5,8	42,7	5,8
17.	30	1	5,5	40,0	5,5
18.	28	1	5,1	37,3	5,1
19.	22	1	4,0	29,3	4,0
20.	21	1	3,8	28,0	3,8

Jumlah siswa = 40

Jika nilai mentah yang paling tinggi adalah 55 diberi nilai 10, maka untuk

52 adalah  $52/55 \times 10 = 9,5$

49 adalah  $49/55 \times 10 = 9,0$  dan seterusnya.

Sekiranya soal tersebut nilai maksimumnya adalah 75 anda ingin menghitung nilainya dengan persentase, maka cara menghitung adalah:

Skor mentah 55 diubah menjadi persen:

$$55/75 \times 100\% = 73,3\%$$

Skor mentah 52 diubah menjadi:

$$52/75 \times 100\% = 69,3\%$$

dan seterusnya seperti yang tercantum pada Tabel diatas kolom 5.

Bilamana persentase pada kolom 5 dijadikan nilai (1 – 10) dimana 73,3% diberi nilai 10 maka perubahan nilai persen menjadi (1 – 10) untuk setiap nilai mentah tercantum pada kolom 6.

Nilai akhir yang dihitung berdasarkan perubahan nilai mentah menjadi nilai (1 – 10) atau nilai mentah menjadi persentase kemudian menjadi nilai (1 – 10) hasilnya sama. Angka-angka yang tercantum pada kolom 4 sama dengan kolom 6.

Kalau pada Contoh 1) nilai hasil ujian bergerak antara 6 dan 10, pada Contoh 2) nilai hasil ujian bergerak antara 3,8 dan 10,0. Dengan kata lain pada kedua kelompok ini tidak ada nilai yang diharapkan adalah 1 sampai 10.

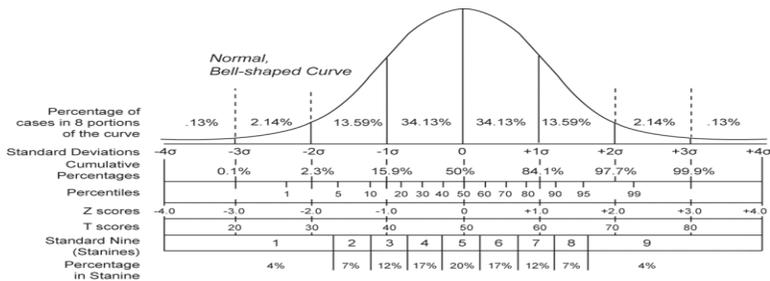
Bilamana jumlah anggota kelompok tidak hanya satu kelas tetapi beberapa kelas sehingga banyaknya siswa atau ratusan jumlahnya, maka untuk memberi

nilai kepada setiap anggota kelompok digunakan statistik sederhana yang menentukan besarnya skor rata-rata kelompok dan simpangan baku kelompok (mean dan standar deviasi).

Jumlah anggota kelompok yang besar, distribusi (penyebaran) kemampuannya mulai dari yang paling pandai, pendek, sedang, kurang dan sangat kurang.

Di bawah ini digambarkan penyebaran kemampuan anggota kelompok menurut kurva normal.

GAMBAR KURVA NORMAL



Penyebaran Berdasarkan Kurva Normal

Gambar diatas tidak hanya penyebaran kemampuan tetapi juga fenomena lainnya seperti tinggi badan sekelompok orang dewasa, berat badan anak usia 10 tahun, ukuran sepatu orang dewasa, berat masing-masing biji kacang goreng dan sebagainya.

Menurut kurva distribusi normal, sekelompok siswa yang memiliki skor rata-rata 60 dalam kelompok itu, jumlah peserta adalah:

60 sampai dengan  $(60+1S.B)$  adalah 34,13%  
 $(60+1S.B)$  sampai dengan  $(60+2S.B)$  adalah 13,59%  
 $(60+2S.B)$  sampai dengan  $(60+3S.B)$  adalah 2,14%

Begitu pula dengan jumlah anak didik yang mendapat skor 60 ke bawah, adalah:

60 sampai dengan  $(60-1S.B)$  adalah 34,13%  
 $(60-1S.B)$  sampai dengan  $(60-2S.B)$  adalah 13,59%  
 $(60-2S.B)$  sampai dengan  $(60-3S.B)$  adalah 2,14%

Dengan perkataan lain, jumlah anak didik yang memperoleh skor antara  $(+1S.B$  s.d  $-1S.B)$  adalah 86,26%, yang mendapat skor  $(+2S.B$  s.d  $-2S.B)$  adalah 95,44%.

Selanjutnya akan diuraikan bagaimana mencari skor rata-rata dan besarnya simpangan baku menurut cara yang paling sederhana

(a) Mencari harga rata-rata

Untuk perhitungan skor rata-rata kelas digunakan data pada tabel diatas. Jumlah skor mentah keempatpuluh siswa adalah 1556. Jumlah siswa adalah 40 orang, jadi skor rata-rata adalah:

$$\frac{\text{Jumlah skor seluruh peserta didik}}{\text{Jumlah anak didik}} = \frac{X}{N}$$

$$= \frac{1556}{40} = 38,9$$

(b) Mencari simpangan baku (S.B)

Rumus simpangan baku:

$$\frac{\text{Jumlah skor } 1/6 \text{ peserta Kelompok Tinggi} - \text{Jumlah Skor } 1/6 \text{ peserta Kelompok Rendah}}{1/2 \text{ Jumlah Peserta tes}}$$

- a. Jumlah skor 1/6 dari siswa yang memperoleh skor tinggi yaitu  
 $Jumlah = 55+52+49+48+46+43 = 293$
- b. Jumlah skor 1/6 dari siswa yang memperoleh skor rendah yaitu  
 $Jumlah = 32+32+30+28+22+21 = 165$
- c. Simpangan baku =

Jumlah skor 1/6 peserta Kelompok Tinggi - Jumlah Skor 1/6 peserta Kelompok Rendah

$$\frac{1}{2} \text{ Jumlah Peserta tes}$$

$$= \frac{336 - 198}{\frac{1}{2} \times 40} = \frac{128}{20}$$

$$= 6,4$$

Tabel 5. Konversi Skor Mentah ke dalam Nilai 1 – 10

SKOR MENTAH	Nilai (1 - 10)	CONTOH
Skor rata-rata + 2,25 S.B	10	Dari Tabel 4 siswa dengan skor mentah 49 mendapat nilai: $38,9 + 6,4x = 49$ $x = \frac{49-38,9}{6,4}$ $x = \frac{10,1}{6,4} = 1,58$ Jadi, nilai 8,5
Skor rata-rata + 1,75 S.B	9	
Skor rata-rata + 1,25 S.B	8	
Skor rata-rata + 0,75 S.B	7	
Skor rata-rata + 0,25 S.B	6	
Skor rata-rata - 0,25 S.B	5	
Skor rata-rata - 0,75 S.B	4	
Skor rata-rata - 1,25 S.B	3	
Skor rata-rata - 1,75 S.B	2	
Skor rata-rata - 2,25 S.B	1	

Contoh lain:

Juga dari Tabel 4

Siswa yang mendapat skor mentah 32 dengan menggunakan Tabel 5 nilainya adalah:

Skor mentah = skor rata-rata + .... S.B

(misalkan .... = a)

$$\begin{aligned}
 32 &= 38,9 + a6,4 \\
 32 - 38,9 &= a 6,4 \\
 - 6,9 &= 6,4a
 \end{aligned}$$

$$a = \frac{-6,9}{6,4} = -1,08$$

jadi koefisien ( $a$ ) S.B = -1 atau -1 Simpangan Baku di bawah skor rata-rata.

atau

-1 S.B terletak diantara -0,75 S.B dengan

-1,25 S.B tepat di tengah 0,75 S.B dengan - 1,25 S.B

atau

pada nilai tengah antara 4 dan 3 yaitu 3,5.

Sehingga skor mentah 32 memperoleh nilai 3,5

Pemberian nilai dengan huruf didasarkan pada besarnya skor rata-rata dan simpangan baku. Blood dan Budd (1972: 201) menggunakan rentangan berikut untuk masing-masing nilai:

Tabel 6. Konversi Skor Mentah kedalam Nilai Huruf

RENTANGAN SIMPANGAN BAKU	NILAI	CONTOH
di atas 1,5 S.B	A	- Skor mentah 49 dengan koefisien S.B = 1,46 adalah <b>nilai B</b> - Skor mentah 32 dengan koefisien S.B = -1 adalah <b>nilai E</b>
0,5 S.B - 1,5 S.B	B	
-0,5 S.B - 0,5 S.B	C	
-0,5 S.B - (-0,5) S.B	D	
dibawah -1,5 S.B	E	

### C. Penentuan Nilai Akhir

Pada umumnya nilai akhir yang diberikan kepada siswa, misalnya untuk nilai rapor atau STTB tidak semata-mata berdasarkan nilai tes sumatif saja, melainkan juga nilai tes-tes formatif dan nilai tugas-

tugas yang lain, bahkan mungkin juga jumlah presensi siswa. Seorang guru yang baik pasti mempunyai beberapa nilai formatif atau juga nilai tugas-tugas tertentu seperti tugas rumah, tugas mengarang, tugas membuat laporan, dan lain-lain. Nilai-nilai tersebut sudah tentu harus diperhitungkan dalam pemberian nilai akhir siswa. Perlu ditambahkan di sini bahwa “nilai-nilai” formatif dan tugas tersebut memang sudah merupakan nilai jadi artinya ia telah diolah baik dengan PAP maupun PAN.

Sebelum menghitung nilai akhir, terlebih dahulu kita harus menentukan apakah kedudukan nilai formatif sama dengan nilai tugas yang lain. Jika nilai formatif dianggap sama dengan nilai tugas yang lain, kita dengan mudah dapat mencari nilai rata-ratanya untuk dijadikan satu nilai formatif. Sebaliknya jika nilai formatif dan tugas lain itu dianggap tidak sama, kita perlu mencari nilai rata-rata keduanya dan masing-masing diperhitungkan sendiri dalam menentukan nilai akhir.

#### 1. Penentuan Nilai Akhir Berdasarkan Nilai Formatif dan Sumatif

Untuk menentukan nilai akhir, kita mempunyai dua data: nilai-nilai formatif dan satu nilai sumatif. Nilai formatif di sini dapat berupa hasil tes formatif saja, dapat pula nilai hasil tes formatif dan nilai tugas lain (kalau dianggap keduanya seimbang). Nilai akhir seorang siswa diperoleh dengan nilai rata-rata formatif ditambah dua kali nilai sumatif dan kemudian dibagi tiga atau jika dituliskan dengan rumus:

$$\text{Nilai Akhir} = \frac{\text{form1} + \text{form2} + \text{form}_n + 2 \times \text{Sumatif}}{3}$$

Misalnya seorang siswa mempunyai 3 nilai formatif masing-masing 7, 7 dan 6 serta nilai sumatif 7. Nilai akhir siswa tersebut adalah:  $\frac{7+7+6+(2 \times 7)}{3} = 6,9$

atau dibulatkan menjadi 7. (Catatan untuk pembulatan: pecahan kurang dari 0,5 dibulatkan ke bawah, pecahan lebih dari (tepat) 0,5 dibulatkan ke atas. Walau terasa kurang adil, pecahan 0,5 pada angka 5,5 dibulatkan menjadi 6, artinya nilai 5,5 akan sama dengan 6,4, sedang 6,5 tidak sama dengan 7,49 walau selisihnya sama).

## 2. Penentuan Nilai Akhir Berdasarkan Nilai Tugas, Formatif dan Sumatif

Untuk menentukan nilai akhir, kita mempunyai tiga data: nilai-nilai tugas, nilai-nilai formatif dan satu nilai sumatif. Dalam hal ini kita membedakan bobot antara nilai tugas dan nilai formatif. Nilai tugas diberi bobot 2 dan nilai formatif dengan bobot 3 serta nilai sumatif dengan bobot 5. Untuk maksud itu, terlebih dahulu kita harus mencari nilai rata-rata tugas dan formatif, dan nilai rata-rata inilah yang kemudian dikalikan dengan bobotnya masing-masing.

Nilai akhir seorang siswa diperoleh dengan cara: dua kali nilai rata-rata tugas, ditambah tiga kali nilai rata-rata formatif, ditambah lima kali nilai sumatif, kemudian dibagi sepuluh. Atau, jika dituliskan dengan rumus:

$$\text{Nilai Akhir} = \frac{2\bar{X}_t + 2\bar{X}_f + 5S}{10}$$

$2\bar{X}_t$  = nilai rata-rata tugas

$2\bar{X}_f$  = nilai rata-rata formatif

S = nilai sumatif

Sebagai contoh, seorang siswa memperoleh nilai tugas mengarang dan menyusun laporan buku masing-masing 6 dan 7, tiga nilai formatif masing-masing 6, 7, dan 7 dan nilai sumatif 7. Nilai rata-rata tugas adalah 6,5  $((6+7):2=6,5)$ , nilai rata-rata formatif 6,7  $((6+6+7):3=6,7)$ . Nilai akhir siswa tersebut adalah:

$$\frac{(2 \times 6,5) + (3 \times 6,7) + (5 \times 7)}{10} = 6,8 \text{ atau dibulatkan menjadi } 7.$$

### 3. Menentukan Siswa ke dalam Tiga Peringkat

Berdasarkan hasil tes yang diperoleh, mungkin kita bermaksud menentukan siswa (dalam satu kelas) ke dalam kelompok tertentu. Pengelompokan siswa yang akan dikemukakan di bawah ini adalah pengelompokan ke dalam tiga peringkat: *atas*, *sedang* dan *bawah*.

Pengelompokan tersebut berdasarkan pada mean dan simpangan baku hasil tes kelompok yang bersangkutan. Adapun batas untuk ketiga kelompok itu adalah sebagai berikut:

- (1) Kelompok atas : mean + S.B ke atas
- (2) Kelompok sedang : di atas mean - S.B sampai di bawah K.A

(3) Kelompok bawah : mean – S.B ke bawah

Sebagai contoh, kita ambil data tes yang disusun ke dalam tabel 4 diatas. Data tersebut mempunyai jumlah subjek 40, mean 38,9 dan simpangan baku 6,4.

Kelompok atas :  $38,9 + 6,4 = 43,5$ . Jadi semua siswa yang memperoleh skor 43,5 ke atas. Dalam contoh data ada 5 orang siswa.

Kelompok sedang : Diatas ( $38,9 - 6,4 = 32,5$  atau bulatkan 33) sampai di bawah 45,3. Jadi semua siswa yang memperoleh skor 33 sampai 44. Dalam contoh data ada 29 orang siswa

Kelompok bawah :  $38,9 - 6,4 = 32,5$  (bulatkan 33). Jadi semua siswa yang memperoleh skor 32 ke bawah; dalam contoh data ada 6 orang.

#### 4. Menghitung Indeks Prestasi Kelompok

Indeks prestasi kelompok merupakan suatu ukuran prestasi yang dicapai oleh siswa dalam satu kelompok untuk suatu mata tes tertentu. Indeks prestasi tersebut menunjukkan tingkat penguasaan siswa terhadap bahan pelajaran yang diajarkan. Oleh karena itu, indeks tersebut sekaligus menunjukkan tingkat penyerapan bahan pelajaran oleh siswa atau dengan kata lain indeks prestasi kelompok sama

dengan tingkat serap atau tingkat penguasaan bahan.

Dengan mengetahui nilai rata-rata yang dicapai kelompok, kita belum dapat mengetahui secara pasti seberapa besar prestasi yang dicapai kelompok tersebut, khususnya jika kita tak mengetahui berapa besar skor maksimal sebuah tes. Indeks prestasi kelompok diperoleh dengan cara: mean dibagi skor maksimal kali seratus. Atau, jika dituliskan dengan rumus:

$$\text{Indeks prestasi kelompok} = \frac{\bar{X}}{\text{Skor Maks.}} \times 100$$

Sebagai contoh penghitungan, kita ambil lagi contoh data di atas (Tabel 4). Kemudian skor tertinggi (skor maksimal) adalah 55, mean 38,9, besarnya indeks prestasi kelompok 38,9 adalah sebagai berikut:

$$\begin{aligned} \text{Indeks Prestasi Kelompok} &= \frac{38,9}{55} \times 100 \\ &= 70,73 \quad (\text{dibulatkan} \\ &\text{menjadi } 71) \end{aligned}$$

Besarnya indeks prestasi 71 menunjukkan bahwa tingkat penyerapan siswa terhadap bahan pelajaran yang diajarkan sebesar 71 persen. Jadi, boleh dikatakan bahwa kelompok tersebut cukup berprestasi atau tinggi tingkat penyerapannya (memadai). Semakin tinggi indeks prestasi yang dicapai kelas, berarti semakin tinggi tingkat penguasaan bahan siswa. Sesuai dengan prinsip belajar tuntas, tingkat penguasaan kelas dianggap memadai jika indeks prestasi kelompok minimal 65 persen.

## 5. Menghitung Koefisien Variasi Kelas

Koefisien variasi kelas adalah koefisien yang menunjukkan besarnya perbedaan kemampuan siswa dalam satu kelas. Semakin besar jarak penyebaran (*range*) skor siswa berarti ada skor yang sangat rendah atau sangat tinggi dan akan semakin besar pula koefisien variasi kelas yang bersangkutan. Demikian pula sebaliknya.

Berdasarkan besar kecilnya koefisien variasi tersebut, kita akan mengetahui apakah keadaan kelas bersifat heterogen normal atautkah homogen. Kelas bersifat *heterogen* jika perbedaan tingkat kemampuan siswa tinggi (ada yang pandai sekali, tetapi ada pula yang bodoh sekali) yaitu jika koefisien variasi mencapai 40 keatas. Kelas dikatakan bersifat *normal* jika kemampuan siswa mengikuti prinsip distribusi normal, yaitu jika koefisien variasi berkisar antara 30 – 39. Kelas dikatakan *homogen* jika perbedaan tingkat kemampuan siswa relatif kecil, yaitu dengan koefisien variasi 29 ke bawah. Koefisien variasi yang jauh diatas 40 (45 ke atas) atau jauh di bawah (25 ke bawah) menunjukkan bahwa keadaan kelas yang bersangkutan sangat heterogen atau sangat homogen.

Koefisien variasi diperoleh dengan cara simpangan baku dibagi mean kali seratus. Atau jika dituliskan dengan rumus:

$$\text{Koefisien Variasi} = \frac{S}{\bar{X}} \times 100$$

Untuk contoh data di atas (berasal dari data Tabel 4) yaitu mean 38,9 dan simpangan

baku 6,9, besarnya koefisien variasi adalah sebagai berikut:

$$\begin{aligned}\text{Koefisien variasi} &= \frac{38,9}{6,9} \times 100 \\ &= 17,74\end{aligned}$$

Koefisien variasi yang sebesar 17,74 menunjukkan bahwa kelas yang data hasil tesnya terlihat pada Tabel 4 di atas adalah kelas yang *sangat homogen*.

# Bab **Delapan**

## **ANALISIS SOAL**

Dalam uraian ini akan dibahas analisis soal yang mencakup analisis butir soal dan analisis perangkat soal. Dalam analisis butir soal, perhatian terutama ditujukan kepada menilai mutu butir soal, yang berdasarkan pada tiga karakteristik butir soal yaitu: (1) tingkat kesukaran, (2) daya pembeda soal, dan (3) pola jawaban soal. Sedangkan perangkat soal akan dianalisis mutunya dengan menggunakan dua spesifikasi, yaitu: (1) validitas dan (2) reliabilitas. Analisis ini diusahakan dengan menggunakan cara yang paling sederhana, sehingga dapat dilakukan oleh setiap guru.

### **A. Alasan Perlunya Analisis Butir Soal**

Butir soal buatan guru pada umumnya dikonstruksi secara tergesa-gesa dan tidak dapat diujicoba sebelum diadministrasikan. Akibatnya banyak butir soal yang digunakan dalam ujian tidak dapat menghasilkan informasi yang benar atau akurat tentang tingkat kemampuan siswa. Hal ini

tentu saja dapat berakibat jauh, karena hasil ujian itu seringkali digunakan untuk mengambil keputusan tentang masa depan siswa. Bila keputusan yang diambil didasarkan kepada informasi yang tidak benar atau tidak akurat, yang disebabkan oleh alat yang digunakan untuk mengumpulkan informasi yaitu butir tes yang tidak terkonstruksi secara baik, maka tentu saja keputusan demikian itu adalah keputusan yang tidak dipertanggungjawabkan. Jadi jelaslah bahwa dibutuhkan adanya alat yang dapat dipercaya untuk dapat mengukur apakah alat ukur (butir soal) yang digunakan memang dapat dijadikan dasar untuk menentukan keputusan yang bijaksana. Inilah peran yang harus dimainkan oleh analisis butir soal, yaitu mengukur butir soal yang akan atau telah digunakan. Hasil pengukuran itu akan memberikan keyakinan pada guru akan tepatnya keputusan yang diambilnya terhadap siswa. Butir soal yang ternyata terlalu lemah, akan sukar dipertanggungjawabkan untuk dijadikan sebagai dasar penentuan keputusan, terutama keputusan yang sifatnya mengenai siswa secara individual.

Kelemahan itu tidaklah terletak pada bentuk atau tipe butir soal yang digunakan, tetapi lebih banyak ditentukan oleh apakah butir soal itu dikonstruksi secara baik atau tidak. Butir soal objektif akan sama baiknya untuk mengukur keberhasilan belajar siswa dengan butir soal uraian yang dikonstruksi secara baik. Bahkan dalam beberapa hal butir soal uraian (esei) jauh lebih besar resikonya daripada butir soal objektif, karena mutu butir soal uraian tidak hanya terletak pada kemampuan siswa untuk menjawab

butir soal tersebut, tetapi lebih banyak ditentukan oleh kemampuan dan objektivitas guru dalam menskor hasil tes tersebut.

Selain itu butir soal objektif dapat dianalisis secara lebih akurat dan bertanggung jawab, sehingga dapat diketahui kelemahannya secara lebih tepat. Hampir tidak mungkin atau belum ditemukan teknik analisis kuantitatif yang dapat mendeteksi kelemahannya secara akurat. Yang dapat dilakukan hanyalah analisis kualitatif, yang banyak bergantung pada kemampuan penguasaan orang yang melakukan analisis tersebut pada bidang studi yang bersangkutan. Karena itu maka uraian selanjutnya ini terutama akan dipusatkan pada butir soal objektif.

Butir soal uraian dikonstruksi untuk digunakan hanya satu kali (tak dapat dipakai ulang dalam tes berikutnya). Karena itu butir tes uraian sebenarnya tidak membutuhkan analisis butir soal setelah butir soal tersebut diadministrasikan. Oleh karena itu pula, maka butir tes uraian tidak dapat diujicoba sebelum tes itu diadministrasikan dalam tes yang sesungguhnya. Jadi berbeda dengan tes objektif. Butir soal tes objektif dapat digunakan berulang-ulang, asalkan tidak dalam perangkat tes yang sama. Karena itu terasa benar adanya kegunaan analisis butir soal, dan kemudian direvisi, sehingga butir soal yang kurang baik konstruksinya dapat diperbaiki, dan akhirnya akan diperoleh butir-butir soal yang telah teruji, dan dapat secara akurat mengukur hasil belajar yang ingin diukur.

Ada beberapa alasan mengapa diperlukan analisis butir soal. Alasan tersebut antara lain:

- a. Untuk dapat mengetahui kekuatan dan kelemahan butir tes, sehingga dapat ditentukan butir soal yang baik atau yang harus direvisi.
- b. Untuk menyediakan informasi tentang spesifikasi butir soal secara lengkap, sehingga akan lebih dimudahkan bagi guru dalam menyusun perangkat soal yang akan memenuhi kebutuhan ujian dalam bidang dan tingkat tertentu.
- c. Untuk segera dapat diketahui masalah yang terkandung dalam butir soal, seperti: soal yang terlalu sukar atau terlalu mudah, atau soal yang tidak dapat membedakan antara siswa yang mempersiapkan diri secara baik atau tidak dalam menghadapi tes. Masalah ini, bila dapat diketahui dengan segera, akan memungkinkan guru untuk mengambil keputusan apakah butir soal yang bermasalah itu akan segera digugurkan atau tidak dalam menentukan nilai peserta didik.
- d. Untuk dijadikan alat untuk menilai butir soal yang akan disimpan dalam kumpulan soal atau bank soal. Kegiatan mengumpulkan butir soal yang baik menjadi kumpulan soal atau bank soal merupakan hal yang dianjurkan. Bila seorang guru telah memiliki sejumlah besar butir soal yang baik, maka ia dengan mudah dapat menyusun suatu perangkat soal yang baik untuk digunakan sesuai dengan tujuan. Dengan demikian guru tersebut tidak perlu khawatir akan mutu butir soal yang digunakannya, karena semua butir soal yang dipakai dalam ujian tersebut telah dapat diketahui karakteristik dan spesifikasinya sebelum ujian dilakukan. Untuk memperoleh informasi

tentang butir soal sehingga memungkinkan untuk menyusun beberapa perangkat soal yang paralel. Penyusunan perangkat seperti ini sangat bermanfaat bila akan melakukan ujian ulang untuk mengukur kemampuan beberapa kelompok peserta tes dalam waktu yang berbeda.

## **B. Karakteristik dan Spesifikasi Butir Soal**

Ada dua istilah yang berkenaan dengan analisis butir soal, yaitu karakteristik butir soal dan spesifikasi butir soal. Yang dimaksud dengan karakteristik butir soal ialah parameter kuantitatif butir soal. Sedangkan yang dimaksud dengan spesifikasi butir soal ialah parameter kualitatif butir soal yang ditentukan atas dasar validitas dan realibilitas.

### **1. Karakteristik Butir Soal**

Dalam bidang tes dan pengukuran, dikenal beberapa karakteristik butir soal. Untuk tes hasil belajar umumnya dipertimbangkan 3 (tiga) karakteristik butir soal, yaitu:

- a) tingkat kesukaran
- b) daya beda
- c) pola jawaban soal

Ketiga karakteristik butir soal ini secara bersama-sama akan menentukan mutu butir soal. Bila salah satu dari ketiga karakteristik ini tidak memenuhi persyaratan maka mutu butir soal akan turun. Dan masing-masing karakteristik itu dapat dikalkulasi dengan mudah, tanpa membutuhkan kemampuan statistik yang kompleks atau kemampuan matematik yang tinggi. Dengan demikian maka



Di dalam istilah evaluasi, indeks kesukaran ini diberi simbol **P** (p besar), singkatan dari kata “proporsi”. Dengan demikian maka soal dengan  $P = 0,70$  lebih mudah jika dibandingkan dengan  $P = 0,20$ . Sebaliknya soal dengan  $P = 0,30$  lebih sukar daripada soal dengan  $P = 0,80$ .

Indeks kesukaran ini disebut juga *indeks kemudahan* atau *indeks fasilitas*, karena semakin mudah soal itu, semakin besar pula bilangan indeksnya. Akan tetapi telah disepakati bahwa walaupun semakin tinggi indeksnya menunjukkan soal yang semakin mudah, tetapi tetap disebut indeks kesukaran.

Rumus mencari P adalah:

$$P = \frac{B}{N}$$

Keterangan:

P = indeks kesukaran

B = banyaknya siswa yang menjawab soal itu dengan betul

N = jumlah seluruh siswa peserta tes

Dari rumus itu kita ketahui bahwa tingkat kesukaran butir soal sangat dipengaruhi oleh tingkat kemampuan anggota kelompok peserta tes. Bila satu butir soal diadministrasikan kepada dua kelompok peserta tes yang berbeda tingkat kemampuannya maka hasilnya dapat diperkirakan akan berbeda pula. Dengan demikian maka dapat dikatakan bahwa tingkat kesukaran butir soal tidak sepenuhnya merupakan ukuran karakteristik butir soal saja, tetapi lebih merupakan kemampuan rata-

rata kelompok peserta tes. Karena itu bila kita jumpai suatu butir tes yang mempunyai tingkat kesukaran 0,45, maka interpretasinya ialah bahwa butir soal itu mempunyai tingkat kesukaran 0,45 untuk kelompok peserta tes tersebut.

Tingkat kesukaran butir soal tidaklah menunjukkan bahwa butir soal tertentu itu baik atau tidak baik. Tingkat kesukaran butir soal hanya menunjukkan bahwa butir soal itu sukar atau mudah untuk kelompok peserta tes tertentu. Butir soal hasil belajar yang terlalu sukar atau terlalu mudah tidak banyak memberikan informasi tentang butir soal atau peserta tes. Untuk tes hasil belajar, tingkat kesukaran yang dianggap baik adalah bila berkisar sekitar 0,50. Atau dengan kata lain, makin dekat tingkat kesukaran suatu butir soal hasil belajar ke 0,50, makin baik butir soal tersebut bagi kelompok tertentu. Sebaliknya makin jauh tingkat kesukarannya dari 0,50 maka makin kurang informasi yang kita peroleh tentang butir soal dan kelompok peserta tes.

Untuk dapat mengkalkulasi tingkat kesukaran butir soal maka perlu dibuat tabel skor hasil tes sebagai berikut:

Tabel 7. Skor Hasil Tes Mata Pelajaran Bahasa Indonesia Siswa Kelas III SMU Negeri 1 Maros

Nama Siswa	Nomor Butir Soal										Skor
	1	2	3	4	5	6	7	8	9	10	
Adam	1	1	1	0	0	1	1	1	0	0	6
Buhari	1	1	0	0	1	1	0	1	1	1	7
Carik	1	1	1	1	1	1	1	1	1	0	9
Daliman	1	0	1	1	1	1	1	1	0	1	8
Elia	1	1	0	0	1	0	0	0	1	0	4
Fatonah	1	0	1	1	1	0	1	1	1	0	7
Gani	1	1	1	1	1	1	1	1	1	1	10
Hamid	1	1	1	0	1	1	0	1	1	0	7
Inem	1	1	0	0	0	0	1	0	1	0	4
Joni	1	1	1	0	0	0	1	1	0	0	5
<b>Jumlah</b>	10	8	7	4	7	6	7	8	7	3	-
<b>P</b>	1,0	0,8	0,7	0,4	0,7	0,6	0,7	0,8	0,7	0,3	-

*Keterangan:* a. 1 = jawaban benar  
 b. 0 = jawaban salah  
 c. dalam contoh ini butir soal hanya 10 buah  
 d. N = 10

Contoh di atas memperlihatkan bahwa tingkat kesukaran soal nomor 1 adalah  $10:10 = 1,0$ , sedangkan butir soal nomor 10 tingkat kesukarannya adalah  $3:10 = 0,3$ . Jadi soal nomor 1 sangat mudah bagi kelompok peserta tes ini, sedangkan butir soal nomor 10 dapat dikategorikan sebagai soal yang sukar untuk kelompok peserta tes tersebut. Jika jumlah siswa besar (50 orang atau lebih) maka perlu dibuat pembagian tiga kelompok, yaitu kelompok atas,

tengah dan bawah untuk memudahkan analisis. Kelompok tengah tidak diikutsertakan dalam analisis butir soal.

Tingkat kesukaran perangkat soal (naskah ujian) dapat ditentukan dengan menjumlah tingkat kesukaran semua butir soal, kemudian dibagi dengan jumlah butir soal. Secara singkat tingkat kesukaran perangkat soal dapat dirumuskan sebagai berikut:

Dalam contoh di atas tingkat kesukaran perangkat ujian adalah:

$$P_{\text{naskah ujian}} = \frac{1,0+0,8+0,7+0,6+0,7+0,6+0,7+0,8+0,7+0,3}{10} = 0,67$$

Untuk sederhananya, tingkat kesukaran butir dan perangkat soal dapat dibagi menjadi tiga kelompok, yaitu mudah, sedang dan sukar. Sebagai patokan dapat digunakan tabel sebagai berikut:

Tabel 8. Tingkat kesukaran

Tingkat Kesukaran	Nilai P
Sukar	0,00 – 0,25
Sedang	0,26 – 0,75
Mudah	0,76 – 1,00

Untuk menyusun suatu naskah ujian sebaiknya digunakan butir soal yang tingkat kesukarannya berimbang sebagai berikut:

$$\begin{aligned} \text{Sukar} &= 25\% \\ \text{Sedang} &= 50\% \\ \text{Mudah} &= 25\% \end{aligned}$$

#### b) Daya Beda

Daya beda butir soal adalah indeks yang membedakan tingkat kemampuan butir soal

membedakan kelompok yang berprestasi tinggi (kelompok atas) dari kelompok yang berprestasi rendah (kelompok bawah) di antara peserta tes. Karena daya beda dihitung dari hasil tes kelompok peserta ujian tertentu, maka dalam penafsiran daya beda pun haruslah selalu dikaitkan dengan kelompok peserta tes (kelompok sampel) tertentu itu. Daya beda suatu butir soal yang didasarkan pada hasil tes suatu kelompok belum tentu akan berlaku pada kelompok yang lain, apalagi bila tingkat kemampuan masing-masing kelompok peserta tes itu berbeda. Misalnya, suatu butir soal yang diujikan kepada siswa jurusan IPS akan sangat berbeda hasil dan interpretasinya bila butir soal tersebut diujikan kepada siswa jurusan IPA. Daya beda butir soal biasa disimbolkan dengan  $D$  (huruf kapital). Langkah-langkah untuk mengkalkulasi daya beda adalah sebagai berikut:

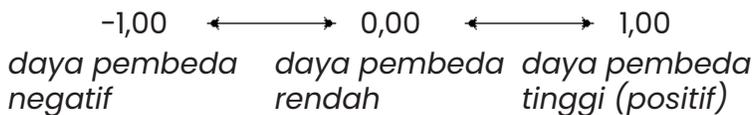
- a. Susunlah urutan peserta tes berdasarkan skor yang diperolehnya, mulai dari skor tinggi sampai ke skor terendah
- b. Bagilah peserta tes tersebut menjadi dua kelompok yang sama jumlahnya. Jika jumlah peserta tes ganjil, maka peserta yang di tengah-tengah tidak usah dimasukkan ke dalam salah satu kelompok. Kelompok peserta dinamakan kelompok prestasi tinggi (kelompok atas) dan kelompok kedua dinamakan kelompok prestasi rendah (kelompok bawah). Bila jumlah peserta cukup besar (lebih dari 50) maka diambil 27% dari kelompok atas dan 27% dari kelompok bawah.
- c. Hitunglah jumlah kelompok atas yang menjawab benar terhadap butir soal yang akan dikalkulasi

daya bedanya. Demikian pula untuk kelompok bawah.

d. Kalkulasikan proporsi peserta yang menjawab benar terhadap butir soal tersebut untuk masing-masing kelompok.

e. Kurangilah proporsi kelompok atas dari kelompok bawah, dan diperoleh indeks daya beda butir soal tersebut.

Dengan demikian ada tiga titik pada daya pembeda yaitu:



Bagi suatu soal yang dapat dijawab benar oleh siswa pandai maupun siswa bodoh, maka soal itu tidak baik karena tidak mempunyai daya pembeda. Demikian pula jika semua siswa baik pandai maupun bodoh tidak dapat menjawab dengan benar. Soal tersebut tidak baik juga karena tidak mempunyai daya pembeda. Soal yang baik adalah soal yang dapat dijawab dengan benar oleh siswa-siswa yang pandai saja.

Jika seluruh kelompok atas dapat menjawab soal tersebut dengan benar, sedang seluruh kelompok bawah menjawab salah, maka soal tersebut mempunyai D paling besar, yaitu salah, yaitu 1,00. Sebaliknya, jika semua kelompok atas menjawab salah tetapi semua kelompok bawah menjawab betul, maka nilai D-nya adalah -1,00. Tetapi jika siswa

kelompok atas dan siswa kelompok bawah sama-sama menjawab benar atau sama-sama menjawab salah, maka soal tersebut mempunyai nilai  $D=0,00$ . Karena tidak mempunyai daya pembeda sama sekali.

Rumus untuk menentukan indeks diskriminasi adalah:

$$D = \frac{Ba}{Na} - \frac{Bb}{Nb}$$

*Keterangan:*

D = daya beda

Na = banyaknya peserta kelompok atas

Nb = banyaknya peserta kelompok bawah

Ba = banyaknya peserta kelompok atas yang menjawab soal itu dengan benar

Bb = banyaknya peserta kelompok bawah yang menjawab soal itu dengan benar

*Klasifikasi daya pembeda*

D : 0,00 – 0,20 = jelek

D : 0,20 – 0,40 = cukup

D : 0,40 – 0,70 = baik

D : 0,70 – 1,00 = baik sekali

D : negatif, semuanya tidak baik, jadi semua butir soal yang mempunyai D negatif sebaiknya dibuang saja

Bila data pada Tabel 7 di atas diterapkan untuk menghitung daya beda butir soal nomor 5, maka akan diperoleh daftar sebagai berikut:

*Kelompok atas*

Gani        10

Carik        9

Budiman	8
Buhari	7
Fatonah	7

*Kelompok bawah*

Hamid	7
Adam	6
Joni	5
Elia	4
Inem	4

Apabila jumlah siswa besar (lebih dari 50 orang), maka perlu dilihat pembagian 3 kelompok: atas, tengah, dan bawah untuk memudahkan analisis.

Kelompok atas yang menjawab benar soal nomor 5 adalah 5 orang sedangkan kelompok bawah hanya 2 orang. Dengan demikian proporsi kelompok atas yang menjawab benar adalah 1,0 dan kelompok bawah adalah 0,4. Jadi daya beda butir soal nomor 5 adalah  $1,0 - 0,4 = 0,5$ . Jadi daya pembeda soal nomor 5 tergolong *baik*.

c) Pola Jawaban Soal

Yang dimaksud pola jawaban di sini adalah distribusi peserta tes dalam hal menentukan pilihan jawaban pada soal bentuk pilihan ganda. Pola jawaban soal diperoleh dengan menghitung banyaknya peserta tes yang memilih pilihan jawaban a, b, c, atau d atau yang tidak memilih pilihan manapun (dalam istilah evaluasi disebut omit, disingkat O).

Dari pola jawaban soal dapat ditentukan apakah pengecoh (*distractor*) berfungsi sebagai pengecoh dengan baik atau tidak. Pengecoh yang tidak dipilih sama sekali oleh peserta tes berarti bahwa

pengecoh itu jelek. Sebaliknya sebuah distraktor (pengecoh) dapat dikatakan berfungsi dengan baik apabila distraktor tersebut mempunyai daya tarik yang besar bagi pengikut-pengikut yang kurang memahami konsep atau kurang menguasai bahan. Semua distraktor dapat berfungsi dengan baik apabila sudah dipilih oleh lebih dari 5% pengikut tes, kemudian pada segi omit tidak dipilih lebih dari 10% pengikut tes.

Sesuatu distraktor dapat diperlakukan dengan 3 cara: (1) diterima, karena sudah baik, (2) ditolak, karena tidak baik, (3) ditulis kembali karena kurang baik. Kekurangannya mungkin hanya terletak pada rumusan kalimatnya sehingga hanya perlu ditulis kembali, dengan perubahan seperlunya. Perhatikan contoh distribusi jawaban berikut ini:

Contoh 1 distribusi jawaban

Pilihan	A	B*	C	D	O
<b>Kelompok</b>					
<b>Kelompok atas</b>	0	3	1	0	0
<b>Kelompok bawah</b>	1	2	1	1	1
<b>Jumlah</b>	1	5	2	1	1

Jawaban yang benar adalah B (diberi tanda bintang), kebanyakan peserta pada (kedua kelompok ini) memilih B. semua distraktornya sudah berfungsi dengan baik karena sudah dipilih lebih dari 5% pengikut tes. Pengecoh A, C, dan D ada yang memilih terutama mereka yang masuk kelompok bawah. Dengan demikian dapat disimpulkan bahwa pengecoh berfungsi sebagai jawaban yang salah.

Dilihat dari segi omit (kolom pilihan paling kanan) adalah baik karena tidak lebih dari 10% pengikut tes tidak memilih. Jadi butir soal diatas adalah baik.

Contoh 2 distribusi jawaban

Pilihan Kelompok	A*	B	C	D	O
Kelompok atas	1	2	2	0	0
Kelompok bawah	3	0	1	1	0
Jumlah	4	2	3	1	0

Jawaban yang benar adalah A (diberi tanda bintang), peserta paling banyak memilih A terutama Kelompok Bawah. Pilihan pengecoh B, C, dan D berfungsi. Tetapi kelompok atas justru cenderung pilihannya pada B dan C. Dari segi omit adalah baik karena tidak ada peserta tes yang tidak menjawab. Jadi butir soal di atas harus diperbaiki atau diganti. Sebaran jawaban seperti ini kemungkinan disebabkan oleh pokok soal (stem) yang kurang baik atau pilihan B dan C cukup menarik sebagai jawaban yang benar atau pilihan A yang perlu diperbaiki.

## 2. Spesifikasi Butir Soal

Ada dua spesifikasi butir soal yang harus dipertimbangkan dalam analisis butir soal, yaitu (a) validitas dan (b) reliabilitas

### a. Validitas Soal

Kata validitas berasal dari kata valid, namun kedua istilah ini mempunyai konteks makna yang berbeda. "Validitas" merupakan sebuah kata benda, sedangkan "valid" merupakan kata sifat. Dari

pengalaman sehari-hari tidak sedikit orang (siswa atau guru) yang mengatakan “Tes ini baik karena sudah validitas”, jelas kalimat tersebut tidak tepat. Yang benar adalah “Tes ini sudah baik karena sudah valid” atau “Tes ini baik karena memiliki validitas tinggi”.

Sebuah tes disebut valid apabila tes itu dapat tepat mengukur apa yang hendak diukur. Istilah “valid” biasa diartikan *sahih*, sehingga validitas dapat diganti menjadi *kesahihan*.

Ada empat macam validitas pada suatu tes adalah sebagai berikut:

1) Validitas Isi

Sebuah tes yang dikatakan memiliki validitas isi apabila mengukur tujuan khusus tertentu yang sejajar dengan materi atau isi pelajaran yang diberikan. Oleh karena materi yang diajarkan tertera dalam kurikulum, maka validitas isi ini sering juga disebut validitas kurikuler. Validitas isi dapat diusahakan tercapainya sejak saat penyusunan dengan cara memerinci materi kurikulum atau buku pelajaran

2) Validitas Konstruksi

Konstruksi berkaitan dengan konsep bidang ilmu yang akan diuji kesahihan tesnya. Validitas konstruksi menunjuk pada pengertian *apakah tes yang disusun itu telah sesuai dengan konsep ilmu yang diteskan itu*. Jika kita menyusun tes kemampuan apresiasi sastra, pertanyaan mendasar yang perlu dijawab adalah: *Apa pengertian apresiasi sastra itu?*

Sebuah tes dikatakan memiliki validitas konstruksi apabila butir-butir soal yang membangun

tes tersebut mengukur setiap aspek berpikir (ingatan, pemahaman, aplikasi, analisis, sintesis, dan evaluasi) seperti yang disebutkan dalam Tujuan Instruksional Khusus. Dengan kata lain jika butir-butir soal mengukur aspek berpikir tersebut sudah sesuai dengan aspek berpikir yang menjadi tujuan instruksional.

Sebagai contoh jika rumusan Tujuan Instruksional Khusus (TIK): "Siswa dapat membandingkan antara efek biologis dan efek psikologis", maka butir soal pada tes merupakan perintah agar siswa membedakan antara dua efek tersebut.

### 3) Validitas Ukuran

Validitas ukuran (norma, standar, kriteris) menunjuk pada pengertian *seberapa jauh siswa yang sudah diajar dalam bidang tertentu menunjukkan kemampuan yang lebih tinggi daripada yang belum diajar*. Apakah siswa yang telah diajar dalam bidang tertentu, misalnya penguasaan kosakata, mempunyai kemampuan penguasaan kosakata yang lebih tinggi daripada siswa yang belum (pernah) diajar dalam hal penguasaan kosakata. Jika kemampuan siswa yang telah diajar penguasaan kosakata lebih tinggi secara signifikan dengan siswa yang belum atau tidak diajar, tes itu dikatakan mempunyai jenis kesahihan ukuran.

Untuk menguji tingkat kesahihan ukuran tes, kita dapat menguji dua kelompok siswa dengan tes yang sama. Kelompok pertama telah diajar dalam bidang yang diteskan itu, sedang kelompok kedua belum diajar. Kesahihan ukuran sebuah tes dapat juga diujikan pada subjek yang sama, yaitu dengan memberikan dua kali tes dengan alat

tes yang sama. Tes pertama dilakukan sebelum siswa diberi pelajaran yang diteskan itu, dan yang kedua setelah siswa selesai diajar. Dalam hal ini, orang menyebutnya sebagai pretes dan postes.

#### 4) Validitas Ramalan

Validitas suatu alat tes diukur dengan validitas ramalan, baru dapat dilakukan pada masa mendatang setelah jangka waktu tertentu. Validitas ramalan menunjuk pada pengertian *apakah sebuah alat tes mempunyai kemampuan untuk meramalkan prestasi yang akan dicapai kemudian*. Seorang siswa yang diuji dengan alat tes tersebut menunjukkan prestasi yang menonjol, apakah ia juga akan berprestasi secara serupa pada mata-mata pelajaran yang diteskan berikutnya. Jika prestasi siswa tersebut tidak berubah, alat tes tersebut berarti mempunyai kesahihan ramalan yang cukup tinggi.

Sebagai contoh, sebuah tes kemampuan apresiasi sastra diteskan kepada sejumlah siswa di satu kelas. Hasilnya menunjukkan bahwa separuh jumlah siswa mendapat skor di atas 7. Pada semester berikutnya siswa di kelas tersebut dites lagi dengan tes kemampuan apresiasi sastra bahan berikutnya, atau bahkan dapat juga dalam mata-mata pelajaran yang lain. Jika prestasi yang dicapai oleh siswa di kelas tersebut masih seperti komposisi prestasi sebelumnya, artinya separuh jumlah siswa yang di atas mendapat skor yang lebih tinggi pula, tes kemampuan apresiasi sastra yang diujikan pertama di atas mempunyai kesahihan ramalan yang tinggi. Sebaliknya jika prestasi yang dicapai siswa kemudian berbeda dengan komposisi sebelumnya, beberapa siswa

yang semula berprestasi baik kini menurun dan beberapa siswa yang semula prestasinya rendah kini justru baik, alat tes kemampuan apresiasi sastra tersebut rendah validitas ramalannya.

Untuk mengetahui tinggi rendahnya kadar kesahihan ramalan, bisanya dilakukan dengan mencari koefisien korelasi antara hasil tes yang pertama dengan hasil tes atau prestasi yang dicapai kemudian. Tinggi rendahnya koefisien korelasi yang diperoleh menunjukkan tinggi rendahnya kadar kesahihan ramalan atas tes yang diuji kesahihannya itu.

*b. Reliabilias Soal*

Kata *reliabilitas* dalam bahasa Indonesia diambil dari kata *reliability* dalam bahasa Inggris, berasal dari kata asal *reliable* yang artinya *dapat dipercaya*. Seseorang dikatakan dapat dipercaya jika orang tersebut selalu bicara ajeg (konsisten), tidak berubah-ubah pembicaraannya dari waktu ke waktu. Demikian pula halnya sebuah tes. Tes tersebut dikatakan dapat dipercaya jika memberikan hasil yang tetap apabila diteskan berkali-kali. Sebuah tes dikatakan reliabel apabila hasil-hasil tes tersebut menunjukkan ketetapan. Dengan perkataan lain, jika kepada para siswa diberikan tes yang sama pada waktu yang berlainan, maka setiap siswa akan tetap berada dalam urutan (rangking) yang sama dalam kelompoknya.

*Contoh:*

<b>Waktu Tes</b> <b>Nama Siswa</b>	<b>Pengetesan Pertama</b>	<b>Pengetesan Kedua</b>
Amin	7	7
Bayu	6,6	6,6
Cahyani	8	9
Didi	5	6
Elvi	6	7
Parida	7	8

Walaupun nampaknya hasil tes pada pengetesan kedua lebih baik, akan tetapi karena kenaikannya dialami oleh semua siswa, maka tes yang digunakan dapat dikatakan memiliki reliabilitas yang tinggi. Kenaikan hasil tes kedua barangkali disebabkan oleh adanya “pengalaman” yang diperoleh pada waktu mengerjakan tes pertama. Dalam keadaan seperti ini dikatakan bahwa adanya akibat yang dibawa karena siswa telah mengalami suatu kegiatan.

## DAFTAR PUSTAKA

- Arikunto, Suharsimi. 1993. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Buchori, M. 1983. *Teknik-Teknin Evaluasi dalam Pendidikan*. Bandung: Jemmars.
- Mulyasa, E. 2003. *Kurikulum Berbasis Kompetensi: Konsep, Karakteristik, dan Implementasi*. Bandung: Remaja Rosdakarya
- Nurkencana,Wayan. 1986. *Evaluasi Pendidikan*. Surabaya: Usaha Nasional
- Stamboel, Conny Semiawan. 1979. *Prinsip dan Teknik Pengukuran dan Penilaian di Dalam Dunia Pendidikan*. Jakarta: Mutiara
- Thoha, M. Chabib. 1996. *Teknik Evaluasi Pendidikan*. Jakarta: Raja Grafindo Persada
- Zainul, Asmawi dan Noehi Nasution. 2001. *Penilaian Hasil Belajar*. Jakarta: Dikti Depdiknas

## RIWAYAT HIDUP PENULIS



**Dr. Syahrudin, S.Pd., M.Pd.** Lahir di Bone, Tanggal 30 Juni 1974. Menyelesaikan pendidikan SD sampai SMA di Kabupaten Bone. Pendidikan S1 di IKIP Ujung Pandang tahun 1997, S2 di Universitas Negeri Makassar tahun 2000, S3 di Universitas Hasanuddin tahun 2012, tugas belajar pada Program *Sandwich/Visitingschoolar* di Northern Illinois University Amerika Serikat tahun 2010–2011. Pekerjaan sebagai PNS Dosen LLDIKTI Wilayah IX Sultan Batara dpt Universitas Muhammadiyah Makassar. Riwayat Pekerjaan: Guru SMPN 18 Makassar tahun 1998 – 2001, Dosen dpt STKIP Yapim Maros tahun 2001 – 2018, Dosen dpt Universitas Muhammadiyah Makassar tahun 2018 sampai sekarang, Ketua Program Studi Pendidikan Bahasa dan Sastra Indonesia STKIP Yapim Maros tahun 2001 – 2009, Wakil Ketua I STKIP Yapim Maros tahun 2009 – 2016, Dekan FKIP Universitas Pancasakti Makassar tahun 2011 – 2017, Instruktur COTIM (*Consortium of Teaching Indonesian and Malay*) Pengajaran Bahasa Indonesia dan Bahasa Melayu untuk Orang Asing di Universitas Sam Ratulangi Manado tahun 2001, Konsultan Pendidikan Nonformal Dinas Pendidikan Kabupaten Maros tahun 2007 – 2009, Penilai dan Reviewer DIKTI Kementerian Pendidikan Nasional RI pada Bahan Ajar PGSD tahun 2008 – 2011, Instruktur Program Bermutu LPMP Provinsi Sulawesi Selatan tahun 2010 – 2012, Dewan Pendidikan Ka-

bupaten Maros tahun 2018 sampai sekarang, Dosen Luar Biasa di Universitas Pancasakti Makassar, Universitas Terbuka, Politeknik Nasional Makassar dan Institut Teknologi dan Kesehatan Permata Ilmu Maros. Aktif dalam kegiatan penelitian dan publikasi karya ilmiah: Pemenang Hibah Penelitian Doktor tahun 2011, Hibah Penelitian Bersaing tahun 2013, tahun 2014, penelitian hibah kompetensi tahun 2017, 2018, 2019, penelitian unggulan perguruan tinggi tahun 2021.