

**DIKTAT MATA KULIAH**  
**PENILAIAN PENCAPAIAN HASIL BELAJAR**



**Disusun Oleh:**

**TRI KUSNAWATI**  
**NUNING CATUR S.W.**

**NIP. 132303689**  
**NIP. 132299492**

**JURUSAN PENDIDIKAN BAHASA PRANCIS**  
**FAKULTAS BAHASA DAN SENI**  
**UNIVERSITAS NEGERI YOGYAKARTA**  
**2010**



Kesalahan terbesar yang bisa dibuat oleh manusia  
di dalam kehidupannya adalah terus-menerus  
mempunyai rasa takut bahwa mereka akan  
membuat kesalahan (Elbert Hubbard)

## KATA PENGANTAR

Puji syukur penulis ucapkan ke hadapan Allah SWT karena berkat rahmat, hidayah, dan karuniaNya penulis dapat menyelesaikan Diktat untuk Mata Kuliah Teknologi Pengajaran Bahasa Prancis ini.

Tujuan penulisan diktat kuliah ini adalah untuk memberikan pemahaman, pengetahuan, serta contoh-contoh yang praktis dan sederhana tentang Penilaian Pencapaian Hasil Belajar. Contoh-contoh disesuaikan dengan lingkungan dan kehidupan sehari-hari para pembaca, khususnya mahasiswa Jurusan Pendidikan Bahasa Prancis. Oleh karena itu, diharapkan setelah membaca, mempelajari, dan membuat beberapa latihan serta tugas, mahasiswa mampu membuat perencanaan dan pelaksanaan penilaian pencapaian hasil belajar Bahasa Prancis.

Ucapan terima kasih penulis sampaikan kepada semua pihak yang telah membantu penulis dalam menyelesaikan penulisan diktat ini. Akhirnya sumbang saran sangat penulis harapkan demi kesempurnaan diktat ini. Semoga diktat ini bermanfaat bagi seluruh pembaca, khususnya mahasiswa dan guru Bahasa Prancis.

Yogyakarta, Desember 2010

Penulis

# **BAB 1**

## **HAKIKAT EVALUASI**

### **1.1 Pendahuluan**

Kegiatan penilaian tidak mungkin dapat dipisahkan dari kegiatan pendidikan dan pengajaran. Semua kegiatan pendidikan dan pengajaran pasti selalu diikuti dengan kegiatan penilaian. Istilah penilaian yang digunakan pada diktat ini sinonim dengan evaluasi. Untuk itu, seorang pengajar (calon pengajar) harus mengetahui hakikat evaluasi. Dalam bab ini kita akan membahas tentang pengukuran dan penilaian, definisi dan komponen penilaian, penilaian sebagai suatu proses, serta tujuan dan manfaat evaluasi.

### **1.2 Pengukuran dan Penilaian**

Penilaian merupakan suatu proses untuk mengetahui (menguji) apakah suatu kegiatan, proses kegiatan, keluaran suatu program telah sesuai dengan tujuan atau criteria yang telah ditentukan (Tuckman dalam Burhan Nurgiyantoro, 2001: 5). Untuk dapat memberikan penilaian secara tepat, misalnya tentang kemampuan siswa dalam membuat surat pribadi dalam bahasa Prancis, pengajar memerlukan data-data tentang kemampuan siswa dalam hal tersebut. Selanjutnya, untuk memperoleh data-data yang diperlukan, pengajar memerlukan alat penilaian yang berupa pengukuran. Dengan demikian, pengukuran merupakan bagian atau alat penilaian saja dan selalu berhubungan dengan data-data kuantitatif yang biasanya berupa skor siswa.

Penilaian selalu berhubungan dengan aspek kualitatif dan aspek kuantitatif, sedangkan pengukuran hanya berhubungan dengan aspek kuantitatif. Aspek kuantitatif pada penilaian diperoleh dengan bantuan pengukuran.

### **1.3 Definisi dan Komponen Penilaian**

Penilaian pada hakikatnya merupakan suatu proses. Cronbach (dalam Burhan Nurgiyantoro, 2001: 7) penilaian adalah proses pengumpulan dan penggunaan informasi yang dipergunakan sebagai dasar pembuatan keputusan tentang program pendidikan.

Scriven dalam Burhan Nurgiyantoro (2001: 7) mengemukakan bahwa proses penilaian terdiri dari tiga komponen, yakni:

- a. Pengumpulan Informasi. Informasi merupakan suatu hal yang sangat esensial karena informasi memberikan data-data yang akan dipergunakan sebagai dasar pembuatan pertimbangan. Informasi ini dapat bersifat kuantitatif atau kualitatif, umum atau khusus, berkaitan dengan orang, materi, prestasi, program, proses, dan sebagainya.
- b. Pembuatan Pertimbangan. Pertimbangan yang dibuat diharapkan tepat jika didukung oleh akuratnya informasi yang diperoleh dan tepatnya penafsiran terhadap informasi tersebut. Pertimbangan

merupakan tafsiran terhadap kondisi yang ada sekarang dan merupakan prediksi penampilan pada masa mendatang.

- c. Pembuatan Keputusan. Pembuatan keputusan merupakan pemilihan di antara sejumlah alternatif. Keputusan yang diambil menuntut diikuti oleh tindakan. Pihak yang terlibat dalam pengambilan keputusan tidak harus orang yang membuat pertimbangan, namun bisa juga orang yang sama.

#### **1.4 Penilaian sebagai Suatu Proses**

Penilaian pada hakikatnya merupakan suatu proses yang harus dilakukan pengajar sebagai bagian dari kegiatan pengajaran. Ten Brink dalam Burhan Nurgiyantoro (2001: 10) mengemukakan tiga komponen proses penilaian sebagai berikut.

- a. Persiapan; berupa kesiapan dan persiapan pihak yang akan melakukan kegiatan penilaian.
- b. Pengumpulan data; berupa kegiatan untuk mendapatkan informasi yang diperlukan.
- c. Evaluasi; berupa kegiatan pembuatan pertimbangan dan pengambilan keputusan.

#### **1.5 Tujuan dan Manfaat Evaluasi**

Tujuan dan manfaat evaluasi antara lain yakni (Burhan Nurgiyantoro, 2001: 15-16):

- a. Untuk mengetahui seberapa jauh tujuan-tujuan pendidikan yang telah ditetapkan dapat dicapai dalam kegiatan belajar mengajar yang dilakukan.
- b. Untuk memberikan objektivitas pengamatan pengajar terhadap tingkah laku hasil belajar siswa. Penilaian merupakan suatu proses yang dilakukan secara berkesinambungan selama proses pembelajaran yang diperoleh melalui pengamatan. Namun pengamatan ini memiliki kelemahan yakni adanya subjektivitas. Untuk meminimalkan subjektivitas ini, maka data-data hasil pengamatan digabungkan dengan hasil pengukuran.
- c. Untuk mengetahui kemampuan siswa dalam bidang-bidang tertentu atau topik-topik tertentu.
- d. Untuk menentukan layak tidaknya seorang siswa dinaikkan ke tingkat di atasnya atau dinyatakan lulus dari tingkat pendidikan yang ditempuhnya.
- e. Untuk memberikan umpan balik bagi kegiatan belajar mengajar yang dilakukan.

#### **1.6 Latihan dan Tugas**

- a. Jelaskan hakikat penilaian dan pengukuran.
- b. Mengapa penilaian dapat memberikan umpan balik bagi kegiatan belajar mengajar yang dilakukan? Jelaskan jawaban Anda.

## **REFERENSI**

Burhan Nurgiyantoro. 2001. *Penilaian dalam Pengajaran Bahasa dan Sastra*. Yogyakarta: BPFE.

## **BAB 2 PENILAIAN**

### **A. PENDAHULUAN**

Pada bab sebelumnya telah dibahas mengenai evaluasi yang di dalamnya memuat perbedaan antara pengukuran dan penilaian. Penilaian dalam bidang pendidikan khususnya dilakukan untuk menaksir tingkat kemampuan seseorang secara tidak langsung melalui respon yang diberikan. Untuk melakukan penilaian diperlukan alat, yang selanjutnya disebut sebagai alat penilaian. Secara umum alat penilaian dibedakan menjadi 2 (dua) macam yaitu tes dan nontes. Uraian berikut ini akan membahas permasalahan tes dan nontes

### **B. TES DAN NONTES**

#### **1. Tes**

Secara sederhana, tes merupakan metode/ cara untuk mengukur kemampuan seseorang/ individu dalam hal kemampuan, pengetahuan, dan keterampilannya dalam bidang tertentu. Jadi tes mempunyai 4 (empat) komponen utama, yaitu metode/ cara, alat ukur, yang diukur (individu), dan bidang yang dites-kan.

Tes memiliki cakupan yang lebih luas dan kompleks daripada nontes. Oleh karena itu, hal tentang tes akan dibahas lebih mendalam di bagian lain.

#### **2. Nontes**

Nontes juga merupakan alat penilaian untuk memperoleh informasi atau data. Informasi atau data yang diperoleh dari nontes ini umumnya merupakan data bidang afeksi dan psikomotor yang secara tidak langsung diperlukan untuk melengkapi data-data bidang kognisi yang diperoleh dari penilaian secara tes. Data nontes akan memberikan penilaian secara kualitatif. Ada 3 (tiga) macam nontes yang dapat dibaca sebagai berikut.

##### **a. Kuesioner**

Kuesioner (*questionnaire*) atau disebut juga angket, yaitu serangkaian pertanyaan yang ditujukan kepada siswa (dalam penelitian disebut sebagai responden). Pertanyaan tersebut diperlukan untuk melengkapi data atas masalah-masalah tertentu yang dibutuhkan. Daftar pertanyaan tersebut dapat bersifat terbuka, artinya siswa diperbolehkan menuliskan pendapat mereka dan bersifat tertutup, siswa hanya memilih alternatif jawaban yang telah disediakan. Daftar jawaban siswa selanjutnya diberi skor dan dimasukkan dalam skala-skala tertentu.

Kuesioner ini merupakan cara yang mudah, murah dan sangat mungkin dilaksanakan dalam suatu penelitian. Langkah-langkah dalam melaksanakan kuesioner adalah sebagai berikut: 1) menentukan tujuan kuesioner, 2) menentukan kelompok sampling (dalam hal ini siswa/ responden), 3) menyusun kuesioner itu sendiri, 4) dokumentasi/ pencatatan kuesioner, dan 5) interpretasi hasil kuesioner.

## **b. Wawancara**

Wawancara (*interview*) merupakan cara untuk memperoleh informasi yang dilakukan oleh pewawancara dengan cara mewawancarai (menanyakan) kepada yang diwawancarai. Dalam hal ini, penanya atau pewawancara adalah guru atau peneliti, sementara yang diwawancarai adalah siswa atau responden. Secara umum, wawancara dibedakan menjadi wawancara terstruktur atau terpin dan wawancara bebas.

## **c. Pengamatan**

Sering juga disebut dengan istilah observasi, yaitu metode atau cara-cara menganalisis dan mengadakan pencatatan secara sistematis mengenai suatu hal (tingkah laku) secara langsung baik terhadap individu maupun kelompok. Pengamatan memerlukan pencatatan karena membutuhkan waktu lebih lama jika dibandingkan dengan tes. Pengamatan secara garis besar dapat dibedakan menjadi pengamatan secara terstruktur dan pengamatan tak berstruktur.

## **C. TEKNIK TES**

Ada beberapa definisi tentang tes dari beberapa ahli. Secara sederhana, tes merupakan tugas (atau tugas-tugas) yang diberikan guru kepada murid. Biasanya orang awam atau bahkan murid akan menyebutnya sebagai soal, soal tes, atau soal ujian. Padahal tugas-tugas tersebut merupakan alat tes atau instrument tes. Kebalikan dari nontes, maka tes ini berurusan dengan data kuantitatif. Setelah tes dilaksanakan, maka data yang diperoleh perlu diolah dan dilambungkan dalam sebuah angka atau skor. Ada beberapa tes yang dapat dilihat sebagai berikut.

### **1) Berdasarkan Jumlah Individu yang Mengikuti Tes**

#### **a) Tes Individu**

Tes yang dikerjakan secara individu (seorang diri)

#### **b) Tes kelompok**

Tes yang dikerjakan secara kelompok

### **2) Berdasarkan Jawaban yang dikehendaki atau tagihan**

#### **a) Tes Perbuatan**

Tes yang menuntut jawaban atau respon dengan cara melakukan sesuatu secara psikomotor. Misalnya ketika guru memberikan tes kepada murid, "*Fermez la porte, s'il vous plait!*". Jawaban yang dikehendaki adalah tindakan seorang siswa menutup pintu.

#### **b) Tes Verbal**

Tes yang menghendaki jawaban dari siswa secara tingkah laku verbal (berbicara atau menulis), yaitu jawaban dalam kalimat. Oleh karena ini dari cara menjawabnya, tes verbal ini dapat dibedakan lagi menjadi tes lisan dan tes tulis.



### 3) **Speed Test dan Power Test**

Keduanya termasuk dalam tes yang dikerjakan secara individu . Ada yang menyatakan kedua tes ini masuk kategori *Aptitude and Ability Test*.

#### a) **Speed Test**

Dari kata *speed* kita sudah dapat menebak, tes ini berhubungan dengan kecepatan. Tes diberikan untuk diselesaikan dalam durasi yang sudah ditentukan, pertanyaan terbatas dalam 1 (satu) hal, dan metode mengerjakan dan menjawab juga jelas. Tes ini menitikberatkan pada banyaknya pertanyaan yang dapat dijawab dengan tepat. Artinya, semakin banyak soal yang diselesaikan dengan benar, maka hasil tes akan semakin baik.

#### b) **Power Test**

Berbeda dengan *Speed test*, *Power test* ini pertanyaan yang diberikan bersifat kompleks, dengan durasi waktu yang sama dengan *speed test*, dan metode penyelesaian jawaban tidak mudah diketahui. Biasanya model tes ini lebih banyak dipakai untuk sekolah Pasca Sarjana atau jenjang menejer dan professional.

Sering dikatakan bahwa hasil *speed test* dapat untuk memprediksi hasil *power test*, karena biasanya seseorang yang *speed testnya* bagus akan bagus pula dalam *power test*.

### 4) **Tes Buatan Guru dan Tes Standar**

#### a) **Tes Buatan Guru**

Tes ini dibuat oleh guru (guru kelas). Tes ini disusun untuk mengetahui atau mengukur tingkat penguasaan siswa terhadap materi/ bidang studi yang telah diajarkan oleh guru tersebut. Soal atau tes buatan guru bisa dikatakan memenuhi syarat kelayakan, kesahihah, dan ketepercayaan bila dalam teknik penyusunan dan pengolahan hasil nilainya mempunyai kriteri kelayakan dan validitas isi (kesahihan isi). Tes buatan guru ini dapat berupa tes lisan, tes tertulis, dan tes perbuatan.

Hasil dari tes buatan guru tidak dapat sebagai pembanding antara sekolah yang satu dengan lainnya. Namun tes ini dapat membedakan hasil belajar siswa yang satu dengan siswa lainnya selama masih dalam satu kelas dan atau satu sekolah.

Bagaimanapun tes buatan guru ini tetap berguna karena gurulah yang menyusun materi pembelajaran, mengujikan, dan menilai hasil belajar siswa.

#### b) **Tes Standar**

Tes standar yang dimaksudkan di sini adalah tes pencapaian atau tes prestasi (*achievement test*). Tes ini mempunyai pengertian bahwa butir-butir soalnya dikerjakan oleh semua

siswa dengan petunjuk yang sama dalam tingkat yang sama pula. Misalnya, ada soal tes standar untuk tingkat Pendidikan Dasar dan Pendidikan Menengah dan bersifat seragam/nasional. Tes standar ini dapat digunakan berkali-kali. Hal ini salah satu yang membedakannya dengan tes buatan guru. Setelah menentukan bahan tes, lalu dilakukan penyusunan butir-butir soal. Selanjutnya butir-butir soal tersebut diujicobakan terhadap sejumlah siswa. Inilah perbedaan selanjutnya jika dibandingkan tes buatan guru. Hasil ujicoba selanjutnya dianalisis untuk mencari koefisien tingkat kesulitan dan daya beda (Tingkat kesulitan dan daya beda akan dibicarakan lebih lanjut dalam bab lain). Butir soal yang terlalu sulit dan terlalu mudah akan direvisi, atau bahkan diganti dengan butir soal baru. Tes standar ini selain berguna untuk mengetahui tingkat hasil belajar siswa juga dapat untuk pembandingan siswa satu sekolah dengan siswa sekolah yang lain bahkan siswa satu wilayah dengan siswa wilayah lain. Satu hal lagi yang membedakan dari tes buatan guru.

## **5) Tes Kemahiran/Keberhasilan (*Proficiency Test*)**

### **a) Tes Penempatan (*Placement Test*)**

Tes penempatan bertujuan untuk setepat-tepatnya menempatkan siswa ke dalam tingkat tertentu atau kelas tertentu dalam suatu sekolah. Jika tes ini dilaksanakan oleh lembaga kursus, maka tujuannya adalah menempatkan peserta kursus di kelas atau tingkat yang sesuai dengan tingkat kemahirannya. Jika satu kelas terdiri dari siswa dengan kemampuan yang seimbang, maka akan memudahkan guru atau pihak sekolah atau pihak lembaga kursus dalam menyiapkan materi dan melaksanakan pengajarannya. Misalnya dalam bahasa asing, sering siswa dikategorikan ke dalam tingkat pembelajar pemula, tingkat menengah, dan tingkat lanjut.

### **b) Tes Diagnostik (*Diagnostic Test*)**

Tes diagnostik disusun untuk tujuan mendiagnosis siswa terhadap aspek khusus bidang yang dipelajari, misalnya tes pelafalan bunyi-bunyi dalam bahasa Prancis. Oleh sebab itu, tes diagnostik dapat dilaksanakan sebelum atau selama pelajaran dimulai.

### **c) Tes Pencapaian/Prestasi (*Achievement Test*)**

Tes ini disusun dengan materi khusus, pada tingkat tertentu, dan tujuan yang telah ditentukan. Tes dilaksanakan setelah masa pengajaran selesai. Kebalikan dari tes penempatan dan tes diagnostik yang dapat dikategorikan sebagai tes kemampuan

awal, maka beberapa ahli menyebut *achievement test* sebagai tes kemampuan akhir.

## 6) Tes Formatif dan Tes Sumatif

### a) Tes Formatif

Tes formatif bertujuan untuk mengetahui tingkat ketercapaian bahan/ materi pengajaran untuk setiap bab, atau unit, atau pokok bahasan tertentu. Oleh karena itu untuk tes ini dapat dilaksanakan beberapa kali bergantung pada banyaknya pokok bahasan yang ada. Ada yang menyebut tes formatif ini sebagai tes atau ulangan harian.

### b) Tes Sumatif

Tes sumatif dilaksanakan setelah suatu kegiatan pengajaran selesai. Misalnya pada akhir semester,akhir *term* atau pada akhir tingkat. Oleh sebab itu tes ini juga sering disebut sebagai ulangan umum atau ulangan/ ujian tingkat.

## D. BENTUK TES

### 1. Tes Objektif

Tes ini sering juga disebut dengan tes jawaban singkat. Butir tes objektif sangat bermanfaat untuk mengukur hasil belajar kognitif tingkat rendah. Namun sebaliknya tes ini tidak cocok untuk mengukur hasil belajar yang bersifat kompres. Tes hasil belajar ini terdiri dari butir-butir soal yang harus dijawab oleh siswa dengan cara memilih satu jawaban benar, menjodohkan, mengisi uraian pendek, dan lain-lain. Tes ini dikatakan mempunyai kebenaran yang objektif. Dalam tes objektif ini hanya satu jawaban benar (betuk skor 1, salah skor 0). Tes objektif ini dapat dijabarkan dalam 4 (empat) macam, yaitu:

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

Contoh:

*Vrai-Faux Paris est la capital de France*

B - S Paris adalah ibukota Negara Prancis

#### b) Tes Objektif Bentuk Menjodohkan (*Matching Test*)

Contoh:

- |   |                |
|---|----------------|
| 1. Quelqu'un qui habite a Paris                     | a. La Seine    |
| 2. La fleuve qui passé Paris                        | b. Le Parisien |
| 3. Le jour d'indépendant de la République Française | c. Asterix     |
| 4. L'une de BD française très célèbre               | d. 14 Juillet  |

#### c) Tes Objektif Bentuk Melengkapi (*Completion Test*), Isian (*Fill in Test*) atau Menyempurnakan

Contoh:

1. Napoléon Bonaparte est français, mais Marie Antoinette est.....
2. M. Bernard enseigne à l'université. Il est.....

d) Tes Objektif Bentuk Pilihan Ganda (*Multiple Choice Test*)

Contoh:

1. Anna est étudiante. Je suis étudiante aussi.  
Nous.....étudiantes.  
a. sont            b. sommes            c. avons            d. allons
2. J'ai un chat. ....chat s'appelle Toto  
a. Un            b. Le            c. La            d. L'

## 2. Tes Esai

Tes ini terdiri dari pertanyaan (pertanyaan) yang menhendaki jawaban dengan uraian panjang. Bentuk tes ini cocok untuk mengukur hasil belajar yang bersifat kompleks (tingkat kognisi tinggi). Dalam tes ini, siswa diharuskan menyusun sendiri kata- kata/ kalimatnya.

Kedua bentuk tes (tes objektif dan tes esai) di atas masing-masing mempunyai kelebihan dan kelemahan. Guru yang ideal akan mampu meminimalkan kelemahan sehingga, tetap dapat memberikan penilaian secara objektif.

## E. PENILAIAN SEJATI (SEBENARNYA) (*Authentic Assessment*)

Penilaian hasil belajar jenis tes dan nontes yang telah disebutkan pada bagian sebelumnya, biasanya dibuat secara homogen dan bersifat diskret. Siswa dinilai dengan soal yang sama, seragam, tanpa membedakan siswa satu dan lainnya. Sementara dalam belajar bahasa, penilaian sebaiknya tidak mengemsidekan perbedaan antarsiswa. Oleh karena itu, muncullah istilah *authentic assessment*. Beberapa crri dan criteria penilaian sejati ini antara lain adalah sebagai berikut 1) lebih mendorong siswa agar dapat mengembangkan jawaban daripada hanya menghafal jawaban, 2) memungkinkan siswa untuk berpikir ke tingkat yang lebih tinggi sesuai dengan kemampuannya, 3) memberikan penilaian secara holistic, 4) penilaian berdasarkan pada desain pengajaran dan kemajuan kelas, 5) mengakomodasi jawaban yang beragam (*multiple responses*), 6) memanfaatkan portofolio siswa, dan 7) mendorong siswa agar dapat melaksanakan evaluasi terhadap tugasnya masing-masing.

*Authentic Assessment* akan membawa siswa pada "dunia nyata" dalam mengerjakan tugas. Misalnya, membaca 1 (satu) text secara utuh, atau menulis 1 (satu) paragraf. Penilaian tertulis akan meliputi konsep pemilihan topik, proses penulisan, dan tentu saja hasil akhir dari tulisan tersebut. Hal ini berlaku juga dalam penilaian lisan.

### 1. Cloze Test

Tes ini bertujuan untuk mengetahui pemahaman siswa terhadap konten suatu teks. Penyusunan *cloze test* ini dengan cara menghilangkan beberapa kata dalam 1 (satu) teks. Siswa kemudian diminta untuk mengisi tempat yang kosong tersebut dengan kata yang tepat, sesuai dengan

konteksnya. Selama jawaban itu masih bisa berterima dengan konteks, maka dimungkinkan ada lebih dari 1 (satu) jawab benar.

Contoh:

To jump to a scene, click \_\_\_\_\_ Henry's shirt pocket. The note pad \_\_\_\_\_. Then, click on the options tab. \_\_\_\_\_ options page appears. Then, click on \_\_\_\_\_ arrows on the right and left \_\_\_\_\_ of the scene pictured near the \_\_\_\_\_ of the page to \_\_\_\_\_ through the scenes. When you find the scene \_\_\_\_\_ want, click the picture to go \_\_\_\_\_. (Diambil dari tulisan Richard E. Mayer, Cognitive Education, USA).

## 2. Tes Lisan

Tes lisan lebih sering terdengar pada jenjang pendidikan yang lebih tinggi. Tes lisan ini bukan hanya menilai pengetahuan saja, namun juga keterampilan dalam menyampaikan pengetahuan tersebut. Dalam hal ini keterampilan berkomunikasi secara lisan juga dinilai. Tes lisan dapat berlangsung di depan kelas (guru berhadapan langsung dengan seorang siswa), secara panel, atau dalam 1 (satu) kelas.

Dalam tes lisan secara formal, biasanya sudah terdapat daftar pertanyaan pada penguji (guru). Siswa harus menjawab pertanyaan tersebut, menjawab secara tepat, jelas, tidak boleh memberikan informasi berlebihan kecuali jika diminta. Sebaliknya jika secara informal, siswa boleh memberikan informasi yang lebih daripada secara formal.

## 3. Portofolio

Penilaian berdasarkan portofolio bukan hanya sekedar berisi kumpulan daftar observasi dan karya siswa secara acak, namun merupakan kumpulan yang dibuat secara sistematis. Observasi tercatat dan hasil karya siswa harus sesuai dengan tujuan pembelajaran. Misalnya buku perantara (*log book*) yang dipegang oleh siswa akan mampu menampilkan refleksi tingkat perkembangan perilaku dan kebiasaan siswa dalam membaca. Tingkat perkembangan pemahaman siswa dalam membaca, kemudian mengkonstruksinya dalam perkembangan perilaku menulis dapat dijadikan sebagai parameter kemajuan siswa.

Selain itu, portofolio juga mampu mencerminkan catatan kemajuan dan pertumbuhan siswa dalam hal membaca dan menulis (karena catatan dikumpulkan secara teratur dalam masa satu tahun, misalnya). Tentu saja, bukan hanya catatan yang dapat dijadikan bahan portofolio. Bahan lain yang dapat dipakai adalah catatan guru, daftar *checklist* guru, buku refleksi milik siswa, *log book* bacaan siswa, contoh jurnal siswa, rangkuman hasil karya siswa, rekaman audio siswa ketika membaca atau bercerita, rekaman visual suatu proyek kelompok (drama), dll.

## **F. CECRL-DELF**

### **1. CECRL**

CECRL atau *Cadre Europeen Cadre Reference de Langue* adalah kerangka acuan umum Eropa tentang bahasa. CECRL ini dikembangkan melalui proses penelitian dan diskusi panjang yang akhirnya menghasilkan perangkat praktis dan susunan standar yang harus ditempuh secara bertahap dalam mempelajari bahasa. CECRL juga sekaligus memuat cara pengajaran dan penilaiannya secara Internasional. Jadi bukan hanya berlaku untuk bahasa Prancis saja.

Pada tahun 2001 Uni Eropa menyetujui adanya system validasi kompetensi bahasa-bahasa di Eropa, yang diikuti juga dengan penetapan Portofolio bahasa-bahasa di Eropa, dan deklarasi hari bahasa.

CECRL dibagi dalam tingkatan-tingkatan sebagai berikut:

- a) *Niveau/* tingkat A1
- b) *Niveau/* tingkat A2
- c) *Niveau/* tingkat B1
- d) *Niveau/* tingkat B2
- e) *Niveau/* tingkat C1
- f) *Niveau/* tingkat C2

Sampai saat ini ada beberapa Negara di luar Uni Eropa yang juga mengadopsi CECRL ini, misalnya Jepang dan Taiwan.

### **2. DELF**

DELF atau *Diplome D'Etudes en Langue Françaises* adalah satu-satunya sertifikat penguasaan bahasa Prancis sebagai bahasa asing yang dikeluarkan oleh Kementerian Pendidikan Prancis. Sampai saat ini ujian DELF diselenggarakan di 163 negara di dunia.

## **G. TUGAS**

Bekerjalah secara kelompok, 1 (satu) kelompok terdiri dari 2 (dua) atau 3(tiga) orang. Buatlah kunjungan ke SMA/ SMK/ MA untuk mendokumentasikan penilaian dalam 1 (satu kelas) selama 1 (satu) semester. Selanjutnya, analisis tes/ instrument tes yang diberikan sesuai dengan criteria jenis tes dan bentuk tes yang telah kita bahas dalam halaman-halaman sebelumnya. Kemudian, simpulkan kelebihan dan kelemahan masing-masing tes dan atau penilaian dalam kelas yang menjadi objek penelitian Anda.

## **H. REFERENSI**

Brown, Dauglas.2004. *Language Assessment: Principle and Classroom Practices*.New York: Longman

Buba, Egon G & Lincoln, Yvonne S. 1989. *Fourt Generation Evaluation*.Newberry Park: Sabe Publication

[http://www.ciel-strasbourg.org/niveau\\_a1.html](http://www.ciel-strasbourg.org/niveau_a1.html)

Lamme & Yhsmith. 1991. *Portfolio approaches to assessing literacy have been described in a wide variety of publications*

Nurgiyantoro, Burhan.2001. *Penilaian dalam Pengajaran Bahasa dan Sastra*. Yogyakarta: PT. BPFE

Valencia, Sheila, et al.1990. "Assessing reading and writing." In Gerald G. Duffy (Ed.), *Reading in the Middle School* (2nd ed.)

## **BAB 3**

### **KRITERIA INSTRUMEN**

#### **A. PENDAHULUAN**

Instrumen sebagai alat pengukur hasil belajar siswa, diharapkan mampu memberikan informasi yang dapat dipertanggungjawabkan kebenarannya. Artinya, instrument dapat memberikan informasi tentang siswa sesuai dengan keadaan yang mendekati sesungguhnya. Agar instrumen dapat memberikan informasi yang dapat dipertanggungjawabkan, instrumen tersebut harus dapat dipertanggungjawabkan sebagai alat penilaian yang baik. Untuk itu, kita membutuhkan informasi apakah alat tes yang disusun telah memenuhi syarat "baik" yang dimaksud.

#### **B. KRITERIA KELAYAKAN ALAT TES**

Untuk mengetahui apakah tes yang telah disusun memenuhi kriteria kelayakan, maka dapat dengan cara menjawab pertanyaan berikut yang jawabannya dapat menunjukkan kadar kelayakan alat tes tersebut. Pertanyaan-pertanyaan tersebut adalah sebagai berikut.

1. Apakah butir-butir tes itu telah sesuai dengan tujuan ?
  - a. Apakah paling tidak telah ada dua butir soal untuk tiap tujuan?
  - b. Apakah semua butir soal mempunyai tujuan, atau telah dimaksudkan untuk mengukur tujuan tertentu?
  - c. Apakah jumlah butir soal per tujuan telah secara tepat mencerminkan kadar pentingnya tujuan itu?
2. Apakah butir-butir soal telah mencerminkan tingkah laku yang sesuai dengan kata-kata kerja operasional yang terdapat dalam tujuan (kompetensi dasar) ?

#### **C. PENTINGNYA TUJUAN**

Untuk menentukan tingkat kelayakan alat tes, kesesuaian dengan tujuan merupakan kriteria utama. Tes yang dapat mengukur keluaran hasil belajar sesuai dengan yang disarankan oleh tujuan itulah tes yang memenuhi kriteria kelayakan. Tiap butir tes harus secara jelas dapat mengacu pada tujuan tertentu. Sebaliknya, setiap tujuan harus mempunyai alat ukurnya dan harus dapat ditunjuk (butir-butir soal nomor berapa, berapa jumlah, apakah telah sesuai dengan tingkat pentingnya dan cakupan bahan yang ditunjuk).

Kadang-kadang terjadi adanya satu atau beberapa tujuan yang tidak mempunyai butir-butir tes yang dimaksudkan untuk mengukur ketercapaiannya. Atau mungkin sebaliknya, ada sejumlah butir soal yang tidak mempunyai tujuan, tidak jelas yang dimaksudkan untuk mengukur ketercapaian tujuan yang mana. Jika kasus seperti ini terjadi, berarti alat tes tersebut tidak memenuhi kriteria kelayakan, karena itu instrumen tersebut bukanlah alat ukur yang baik.



#### **D. KESESUAIAN DENGAN BAHAN**

Tes yang baik adalah yang sesuai dengan bahan pelajaran yang telah diajarkan. Bahan pelajaran itu sendiri dikembangkan berdasarkan tujuan. Dengan demikian, kaitan antara ketiga komponen tujuan, bahan, dan alat penilaian cukup erat. Penyusunan alat tes pada kenyataannya lebih mendasarkan diri pada bahan pelajaran. Penilaian terhadap ketercapaian tujuan pada hakikatnya tidak lain dari penilaian terhadap penguasaan bahan.

Semua bahan (pokok bahasan) yang diajarkan harus diteskan. Pada kenyataannya, juga sering terjadi terdapat beberapa pokok bahasan yang diajarkan tidak diteskan, atau pengambilan bahan tes tidak bersifat mewakili. Sebaliknya, kadang-kadang terjadi adanya sejumlah butir soal yang dimaksudkan untuk mengukur penguasaan terhadap suatu bahan (pokok bahasan) yang belum atau tidak diajarkan.

Kesesuaian alat tes dengan tujuan dan bahan pelajaran ini merupakan salah satu jenis kesahihan, yaitu kesahihan isi (*content validity*), suatu jenis kesahihan yang penting dalam tes buatan guru. Agar alat tes dapat dipertanggungjawabkan tingkat kelayakan atau kesahihan isinya, penyusun tes itu hendaknya mendasarkan diri pada deskripsi bahan yang diajarkan. Dengan kata lain, kita perlu membuat deskripsi bahan yang diajarkan.

#### **E. TUGAS**

Bekerjalah secara kelompok, 1 (satu) kelompok terdiri dari 2 (dua) atau 3(tiga) orang. Carilah contoh tes sumatif Bahasa Prancis SMA. Analisislah soal-soal pada tersebut berdasarkan tujuan, bahan, dan alat penilaiannya. Apakah sudah sesuai dan seimbang jumlah butir soalnya untuk tiap-tiap tujuan (kompetensi dasar).

#### **F. REFERENSI**

Brown, Douglas.2004. *Language Assessment: Principle and Classroom Practices*.New York: Longman

Nurgiyantoro, Burhan.2001. *Penilaian dalam Pengajaran Bahasa dan Sastra*. Yogyakarta: PT. BPFE

## BAB 4 VALIDITAS DAN RELIABILITAS TES

### A. PENDAHULUAN

Ketika kita ingin mengukur sesuatu, kita harus dapat memilih alat ukur yang sesuai sehingga kita dapat memperoleh hasil pengukuran dengan tepat. Ketepatan hasil pengukuran inilah yang dinamakan validitas. Ada beberapa macam validitas yang sudah dibicarakan dalam bab sebelumnya. Sehubungan dengan penyusunan alat tes, ada istilah bahwa selain valid, tes juga harus reliabel. Apa itu Reliabilitas? Tes harus reliabel, artinya jika tes tersebut digunakan lebih dari satu kali pada kelompok yang sama maka tes tersebut harus dapat memberikan hasil pengukuran yang tetap. Reliabilitas suatu tes dapat ditingkatkan dengan cara menambah jumlah butir soal, dengan catatan bahwa soal yang ditambahkan harus homogen dengan butir soal yang sudah ada.

### B. VALIDITAS

#### 1. Teknik t-test

Teknik t-test ini dapat dipakai untuk menghitung koefisien dengan cara mencari perbedaan nilai rerata hasil tes yang sama pada 2 (dua) kelompok siswa yang berbeda. Misalnya tes teori sastra yang diberikan pada kelompok pertama yang sudah memperoleh mata pelajaran teori sastra, sementara kelompok kedua belum memperoleh teori sastra. Kelompok pertama sebagai ( $X_1$ ) dan kelompok kedua sebagai ( $X_2$ ). Contoh penghitungannya, seperti terdapat dalam buku Nurgiyantoro (2001).

a) Jika subjeknya berbeda, maka rumus yang digunakan adalah sebagai berikut.

$$t = \frac{X_1 - X_2}{\sqrt{\frac{s_1^2}{n_1} - \frac{s_2^2}{n_2}}}$$

#### Keterangan=

t = koefisien yang akan dicari

$X_1$  = nilai rerata kelompok pertama (yang sudah memperoleh teori sastra)

$X_2$  = nilai rerata kelompok kedua (yang belum memperoleh teori sastra)

n = jumlah subjek

$s^2$  = taksiran varian

Dari rumus dan keterangan di atas, ternyata kita masih harus mencari nilai taksiran varian, yaitu  $s^2$ . Dengan rumus berikut di bawah ini kita akan dapat memperoleh taksiran varian.

$$S^2 = \frac{(\sum X_1^2 - \frac{(\sum X_1)^2}{n_1}) + (\sum X_2^2 - \frac{(\sum X_2)^2}{n_2})}{n_1 + n_2 - 2}$$

Penghitungan koefisien dimulai dengan membuat daftar rerata kelompok pertama ( $X_1$ ) dan kelompok kedua ( $X_2$ ) seperti berikut ini (Nurgiyantoro, 2001).

Daftar Nilai Rerata Kelompok Pertama ( $X_1$ )

Daftar Nilai Rerata Kelompok Kedua

			$(X_2)$		
No urut	$X_1$	$X_1^2$	No urut	$X_2$	$X_2^2$
1.	7,5	56,25	1.	5,0	25,00
2.	7,0	49,00	2.	4,5	20,25
3.	7,0	49,00	3.	4,0	16,00
4.	6,5	42,25	4.	4,0	16,00
5.	7,5	56,25	5.	3,5	12,25
6.	5,5	30,25	6.	4,5	20,25
7.	8,0	64,00	7.	4,0	16,00
8.	7,0	49,00	8.	3,5	12,25
<b><math>n_1 = 8</math></b>	$\sum X_1 = 56$	$\sum X_1^2 = 396$	<b><math>n_2 = 8</math></b>	$\sum X_2 = 33$	$\sum X_2^2 = 138$

Rerata masing-masing kelompok adalah sebagai berikut.

$$X_1 = \frac{\sum X_1}{n_1} = \frac{56}{8} = 7$$

$$X_2 = \frac{\sum X_2}{n_2} = \frac{33}{8} = 4,125$$

Langkah selanjutnya, tinggal memasukkan nilai rerata masing-masing kelompok untuk mengetahui taksiran varian ( $S^2$ ).

$$\begin{aligned}
S^2 &= \frac{(396 - \frac{56^2}{8}) + (138 - \frac{33^2}{8})}{8 + 8 - 2} \\
&= \frac{396 - 392 + 138 - 136,125}{14} \\
&= \frac{4 + 1,875}{14} = \mathbf{0,42}
\end{aligned}$$

Selanjutnya tinggal memasukkan angka-angka hasil penghitungan di atas ke dalam rumus t-test sebagai berikut.

$$\begin{aligned}
t &= \frac{X_1 - X_2}{\sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}} \\
&= \frac{7 - 4,125}{\sqrt{\frac{0,42}{8} + \frac{0,42}{8}}} \\
&= \frac{2,875}{0,324} = \mathbf{8,873}
\end{aligned}$$

Koefisien nilai **t** sudah diperoleh. Langkah selanjutnya adalah melihat dalam tabel nilai kritis t (terlampir). Derajat Kebebasan (DB) 14, dengan tingkat signifikansi 0,1% hanya membutuhkan koefisien **4,140**. Sementara nilai **t** di atas adalah **8,873**, jadi sangat signifikan bahkan pada batas kesahihah 0,1 %. Maknanya, tes teori sastra tersebut sangat sahih.

a) Jika subjeknya sama, maka rumus yang digunakan adalah sebagai berikut.

$$t = \frac{\Sigma D}{\sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{N - 1}}}$$

### Keterangan

D = Perbedaan skor antara 2 (dua) tes ( $X_1 - X_2$ )

$\sum D$  = Jumlah perbedaan skor antara  $X_1 - X_2$

N = Jumlah subjek

Skor  $X_1$  dapat diperoleh dari hasil *pretest* teori sastra, dan skor  $X_2$  diperoleh dari hasil *posttest*. Selanjutnya, langkah yang harus dilakukan adalah menyusun daftar skor agar lebih mudah dihitung seperti di bawah ini.

Subjek	$X_1$	$X_2$	D	$D^2$
1.	4,0	7,0	-3,0	9,00
2.	3,5	7,3	-3,8	14,44
3.	3,5	6,8	-3,3	10,89
4.	4,3	7,5	-3,2	10,24
5.	4,2	7,5	-3,3	10,89
6.	3,3	6,2	-2,9	8,41
7.	4,2	6,8	-2,6	6,76
8.	4,3	7,5	-3,2	10,24
n= 8	$\sum X_1= 28,3$	$\sum X_2 = 56,5$	$\sum D= -25,3$	$\sum D^2= 80,87$

Tahapan berikutnya, masukkan hasil penghitungan pada tabel di atas ke dalam rumus sebagai berikut.

$$\begin{aligned} t &= \frac{-25,3}{\sqrt{\frac{8 \times 80,87 - (-25,3)^2}{8 - 1}}} \\ &= \frac{-25,3}{\sqrt{\frac{646,96 - 640,09}{7}}} \\ &= \frac{-25,3}{0,991} = -25,53 \end{aligned}$$

Harap diingat bahwa yang diperhatikan adalah angka mutlak, sehingga koefisien (-25,53) dapat diperhitungkan sebagai 25,53. Selanjutnya silakan konsultasikan dengan tabel kritis nilai t. DB 7 pada tingkat kesahihan 0,1 % hanya 5,405, sementara hasil penghitungan di atas

memperlihatkan koefisien yang sangat jauh dari koefisien dalam tabel. Artinya, tes teori sastra yang diujikan pada subjek yang sama tadi mempunyai tingkat kesahihan yang tinggi.

## 2. Korelasi Product Moment

Alat tes yang hendak diukur ke-validannya atau tingkat kesahihannya dengan cara mengorelasikan 2 (dua) hasil tes yang sejenis. Misalnya hasil tes *Comprehension Ecrite* dan *Vocabulaire*. Kedua mata kuliah (atau mata pelajaran) ini bersifat reseptif. Contoh penghitungannya, seperti terdapat dalam buku Nurgiyantoro (2001).

Besaran koefisien korelasi dapat mendasarkan pada ketentuan berikut ini

- Koefisien 0,800 sampai 1,00 = sangat tinggi
- Koefisien 0,600 sampai 0,799 = tinggi
- Koefisien 0,400 sampai 0,599 = cukup
- Koefisien 0,200 sampai 0,399 = rendah
- Koefisien 0,00 sampai 0,199 = sangat rendah

**Catatan**= Ketentuan ini sangat dipengaruhi oleh jumlah subjek (N). Ada yang berpendapat bahwa penghitungan korelasi lebih baik berdasarkan table (lihat lampiran).

$$r_{1.2} = \frac{N\sum X_1X_2 - (\sum X_1)(\sum X_2)}{\sqrt{(N\sum X_1^2 - (\sum X_1)^2)(N\sum X_2^2 - (\sum X_2)^2)}}$$

### Keterangan:

- $r_{1.2}$  = Koefisien korelasi yang dicari
- N = Jumlah skor
- $X_1$  = Skor hasil tes pertama (yaitu *Comprehension Ecrite*)
- $X_2$  = Skor hasil tes kedua (yaitu *Vocabulaire*)

Selanjutnya, perhatikan data nilai (skor) dalam table berikut ini. ( $X_1$ ) merupakan skor hasil tes pertama, yaitu *Comprehension Ecrite*. ( $X_2$ ) merupakan skor hasil tes kedua, yaitu *Vocabulaire*. Sementara N merupakan jumlah skor atau jumlah subjeknya.

Tabel Persiapan Penghitungan Koefisien Korelasi untuk mengukur kesahihan tes *Comprehension Ecrite*. Silakan menggunakan program *excel* untuk menghitung.

No Urut Subjek	X <sub>1</sub>	X <sub>2</sub>	X <sub>1</sub> <sup>2</sup>	X <sub>2</sub> <sup>2</sup>	X <sub>1</sub> X <sub>2</sub>
1.	6,5	7,0	42,25	49,00	45,50
2.	6,5	6,5	42,25	42,25	42,25
3.	6,0	6,5	36,00	42,25	39,00
4.	5,6	6,5	31,36	42,25	36,40
5.	7,0	6,8	49,00	46,24	47,60
6.	7,0	7,5	49,00	56,25	52,50
7.	6,5	7,0	42,25	49,00	45,50
8.	6,0	6,8	36,00	46,24	40,80
N= 8	∑X <sub>1</sub> =51,1	∑ X <sub>2</sub> = 54,6	∑ X <sub>1</sub> <sup>2</sup> = 328,11	∑ X <sub>2</sub> <sup>2</sup> = 373,48	∑ X <sub>1</sub> X <sub>2</sub> = 349,55

Kemudian masukkan data hasil penghitungan awal di atas ke dalam rumus korelasi

$$\begin{aligned}
 r_{1.2} &= \frac{8 \times 349,55 - (51,1)(54,6)}{\sqrt{(8 \times 328,11 - 51,1^2)(8 \times 373,48 - 54,6^2)}} \\
 &= \frac{2.796,4 - 2.790,06}{\sqrt{(13,67)(6,68)}} \\
 &= \frac{6,34}{9,56} \\
 &= 0,663
 \end{aligned}$$

Jika berdasarkan ketentuan besarn di atas, maka  $r = 0,663$  masuk dalam range antara 0,600 sampai 0,799 yang berarti tinggi. Dengan demikian dapat ditafsirkan bahwa dengan jumlah siswa 8 (delapan) orang, tes *Comprehension Ecrite* mempunyai kesahihan sejalan yang **tinggi**. Namun tidak demikian halnya jika penafsiran nilai koefisien korelasi tersebut berdasarkan pada **nilai kritis korelasi**. Nilai  **$r = 0,663$**  tersebut di atas tidak signifikan. Silakan lihat pada tabel yang terlampir.

### C. RELIABILITAS

#### 1. KR20, 21 (Kuder-Richardson 20-21)

Penghitungan koefisien reliabilitas KR20,21 ini hanya dikenakan pada tes yang mempunyai skor dikotomi. Maksudnya, jawaban benar diberi skor 1 (satu) dan jawaban salah diberi skor 0 (nol).

### Rumus KR20

$$r = \frac{n}{n - 1} \left( 1 - \frac{\sum pq}{S^2} \right)$$

#### Keterangan

- r = Koefisien reliabilitas tes
- n = Jumlah butir soal
- p = Proporsi jawaban betul
- q = Proporsi jawaban salah (  $q = 1 - p$  )
- S = Simpangan baku
- S<sup>2</sup> = Varian

### Rumus KR21

$$r = \frac{n}{n - 1} \left( 1 - \frac{X(n - X)}{nS^2} \right)$$

#### Keterangan

- r = Koefisien reliabilitas tes
- n = Jumlah butir soal
- X = Rerata (*mean*)
- S<sup>2</sup> = Varian

Tuckman (1975) menyatakan bahwa besarnya koefisien tingkat reliabilitas di antara 0 - 1,0. Jika dari penghitungan diperoleh nilai koefisien 0 atau bahkan *minus* (negatif) maka dapat dikatakan bahwa tes tersebut sangat rendah tingkat ketepatannya. Sebaliknya semakin tinggi nilai koefisien maka tes dikatakan semakin terpercaya. Tes buatan guru dikatakan terpercaya jika mencapai nilai koefisien minimal 0,60 sedangkan tes yang dipublikasikan minimal harus mempunyai tingkat koefisien 0,85.

Ebel (1979) menambahkan bahwa cara penghitungan KR 20 yang rumit, terutama ketika menghitung **p** dan **q**, maka rumus KR 21 lebih direkomendasikan.

Contoh penghitungan lengkap terdapat pada BAB VIII.



## 2. Koefisien Alpha

Berbeda dengan KR20,21, koefisien Alpha atau lengkapnya Alpha Cronbach menghitung koefisien reliabilitas tes dengan jawaban berjenjang. Maksudnya jawaban antara angka 1-5, seperti yang biasa terdapat dalam pemberian skor angket.

Tes dikatakan mempunyai tingkat kepercayaan jika dari penghitungan nilai koefisien menghasilkan angka antara 0,5 – 1,0. Sebaliknya tingkat koefisien di bawah 0,5 dikatakan tidak tepercaya atau instrumen tes itu buruk. Sementara ada pendapat lain yang menyatakan bahwa baik buruknya tes/ instrument tes dengan membandingkan nilai koefisien r hitung ( $\alpha$ ) dengan koefisien r dalam tabel. Jika nilai dari hitungan lebih besar daripada nilai koefisien r ( $\alpha$ ) tabel, maka dikatakan tes itu mempunyai tingkat kepercayaan yang tinggi.

### Rumus

$$r = \frac{K}{K - 1} \left( 1 - \frac{\sum S_i^2}{S_t^2} \right)$$

Keterangan

K = Jumlah butir soal

$\sum S_i^2$  = Jumlah varian butir-butir soal

$S_t^2$  = Total varian (seluruh butir tes)

## 3. Spearman Brown

Koefisien Spearman Brown merupakan kelanjutan dari “Teknik Belah Dua” atau (*Split-Half Reliability*) yang hanya menghitung koefisien reliabilitas separuh soal. Spearman Brown menyempurnakannya sehingga akan memberikan reliabilitas keseluruhan tes. Nilai koefisien r ini akan sangat dipengaruhi oleh jumlah subjek (n).

Rumus Spearman Brown

$$\text{Reliabilitas seluruh tes} = \frac{2 \times \text{reliabilitas separuh tes}}{1 + \text{reliabilitas separuh tes}}$$

Contoh penghitungan terdapat pada BAB 8.

#### **D. TUGAS.**

Tugas kelompok dengan jumlah anggota 3-5 orang. Silakan mengambil sampel soal uji coba SNMPTN yang diselenggarakan oleh suatu lembaga bimbingan belajar di Yogyakarta. Temukan subjek sebanyak 20 orang agar Anda memperoleh skor sebanyak 20 orang. Selanjutnya analisislah tingkat kesahihan dan kepercayaan soal tes tersebut berdasarkan dari skor 20 subjek dari lembaga bimbingan belajar. Selamat mengerjakan!

#### **E. REFERENSI**

Nurgiyantoro, Burhan.2001. *Penilaian dalam Pengajaran Bahasa dan Sastra*. Yogyakarta: PT. BPFE.

<http://www.azuarjuliandi.com>

<http://www.scribd.com>

## **BAB 5**

### **TABEL SPESIFIKASI**

#### **A. PENDAHULUAN**

Agar instrumen penilaian yang dibuat memiliki keseimbangan dalam jumlah dan bobot, sebelum menyusun instrumen perlu dibuat tabel spesifikasi. Tabel spesifikasi berguna bagi guru atau penyusun soal tes untuk selalu mendasarkan diri pada rambu-rambu yang telah ditetapkan. Hal ini penting karena tak jarang guru hanya memfokuskan diri pada beberapa pokok bahasan dan tingkatan kognitif tertentu, sengaja ataupun tidak. Misalnya seorang guru bahasa Prancis yang kebetulan menyenangi tata bahasa, seperti dalam pengajarannya, ia akan lebih banyak membuat soal tata bahasa saja yang mungkin hampir separuh jumlah seluruh soal. Akibatnya, pokok soal yang lain kurang mendapat porsi yang semestinya.

#### **B. ISI TABEL SPESIFIKASI**

Tabel spesifikasi atau rincian sering juga disebut sebagai kisi-kisi (*blue print*). Sesuai namanya tabel spesifikasi berisi (a) perincian bahan pelajaran atau pokok bahasan yang akan diteskan, (b) tingkah laku atau kemampuan tingkatan-tingkatan kognitif yang dikehendaki, (c) perimbangan dan jumlah soal untuk tiap tingkatan kognitif per pokok bahasan, (d) persentase dan jumlah soal per pokok bahasan, per tingkatan kognitif, dan seluruh soal yang akan disusun, dan bahkan ada yang telah berisi, (e) bentuk tes.

#### **C. PERTIMBANGAN BESARNYA PROPORSI TIAP POKOK BAHASAN**

Untuk keperluan pembuatan tes sumatif, pihak yang berwenang yang dalam hal ini Kantor Diknas telah memberikan tabel spesifikasi tersebut ke sekolah-sekolah. Akan tetapi, guru dapat membuat sendiri sesuai dengan perimbangan proposi yang dikehendaki, baik yang menyangkut proporsi untuk tiap pokok bahasan, tingkah laku tingkatan kognitif porsi untuk tiap bahasan, tingkah laku keluaran belajar yang diukur maupun bentuk tes yang dipilih. Akan tetapi, penentuan besarnya proporsi untuk tiap pokok bahasan per tingkatan kognitif tidak boleh menurut selera sendiri, melainkan harus dengan pertimbangan-pertimbangan tertentu.

Pertimbangan *pertama*, berkaitan dengan aspek kejiwaan, menuntut kita untuk mempertimbangkan tingkat perkembangan kognitif siswa. Dalam kaitan ini, kita akan berurusan apakah siswa itu tingkat SD, SMTP, SMTA, atau bahkan mahasiswa. Pertimbangan *kedua*, berkaitan dengan penting dan luasnya pokok bahasan. Pokok-pokok bahasan yang penting, dalam arti sangat menopang tercapainya tujuan pengajaran secara keseluruhan, dan mempunyai cakupan yang lebih luas, dibanding pokok(-pokok) bahasan yang lain, haruslah diberi proporsi yang lebih besar pula. Sebaliknya pokok bahasan yang tidak begitu luas cakupan bahasanya proporsinya juga lebih kecil.

#### D. PENGISIAN TABEL SPESIFIKASI

Kegiatan pertama pengisian tabel spesifikasi menentukan pokok-pokok bahasan yang akan diteskan, kemudian menentukan perimbangan bobot masing-masing pokok bahasan dan tingkatan aspek kognitif yang akan diungkap. Setelah itu, kita menentukan jumlah seluruh soal yang akan diteskan, baik tes yang berbentuk objektif atau esai, dengan mendasarkan diri pada waktu yang tersedia.

**Tabel**  
**Contoh Tabel Spesifikasi dengan Pertimbangan Bobot**  
**per Pokok Bahasan dan Tingkatan Kognitif dalam Persentase**

Tingkatan Kognitif	Ingatan	Pemahaman	Aplikasi	Analisis	Sintesis	Evaluasi	Jumlah
Pokok Bahasan	(C <sub>1</sub> )	(C <sub>2</sub> )	(C <sub>3</sub> )	(C <sub>4</sub> )	(C <sub>5</sub> )	(C <sub>6</sub> )	
CO	3	2	2	1	-	-	20%
CE	6	4	3				25%
EE							10%
Voc							15%
Gram							20%
Bud							10%
Jumlah	20%	30%	25%	15%	5%	5%	100%

Keterangan: CO (*Compréhension Orale*/Mendengarkan)  
CE (*Compréhension Écrite*/Membaca)  
EE (*Expression Écrite*/Menulis)  
Voc (*Vocabulaire*/Kosakata)  
Gram (*Grammaire*/Tata Bahasa)  
Bud (Pengetahuan Budaya)

#### E. PENENTUAN JUMLAH SOAL

Dengan mendasarkan diri pada besarnya bobot per pokok bahasan per tingkatan kognitif di atas, selanjutnya kita dapat menentukan jumlah soal untuk tiap sel, setelah lebih dulu kita tentukan jumlah keseluruhan soal dan jumlah skor maksimal. Misalnya, tes akan diberikan selama 90 menit, jumlah skor maksimal 75 dengan perimbangan tes bentuk objektif 80 % dan esai 20 %. Jumlah soal dan skor tes objektif 60 (skor tiap soal 1, bentuk tes dapat B-S, pilihan ganda, penjumlahan, atau jawaban singkat), sedang jumlah skor esai 15 ( $20\% \times 75 = 15$ ) dengan soal 3 atau 4. Agar dapat mengukur proses berpikir siswa, tes esai khusus untuk mengukur tingkat C<sub>4</sub> (10%), C<sub>5</sub> dan C<sub>6</sub> (juga 10%). Dengan demikian, tak ada tes objektif untuk C<sub>5</sub> dan C<sub>6</sub>, sedang C<sub>4</sub> tinggal 5 %.

Untuk memudahkan pengisian tiap sel, kita perlu menghitung jumlah skor per pokok bahasan (horisontal) dan per tingkatan kognitif (vertikal) dengan (sementara) mengabaikan bentuk tesnya, objektif atau esai. Perhitungan- perhitungan yang dimaksudkan dicontohkan di bawah ini.

(a) perhitungan jumlah skor untuk tiap pokok bahasan:

- Mendengarkan :  $20\% \times 75 = 15$
- Membaca :  $25\% \times 75 = 18,75$  (19)
- Menulis :  $10\% \times 75 = 7,50$  (7)
- Kosakata :  $15\% \times 75 = 11,25$  (11)
- Tata Bahasa :  $20\% \times 75 = 15$
- Budaya :  $10\% \times 75 = 7,5$  (8) +
- Jumlah = 75

(b) perhitungan jumlah skor untuk tiap pokok kognitif:

- Ingatan :  $20\% \times 75 = 15$
- Pemahaman :  $30\% \times 75 = 22,50$  (22)
- Penerapan :  $25\% \times 75 = 18,75$  (19)
- Analisis :  $15\% \times 75 = 11,25$  (11)
- Sintesis :  $5\% \times 75 = 3,75$  (4)
- Evaluasi :  $5\% \times 75 = 3,75$  (4) +
- Jumlah = 75

Catatan : jika semua soal berbentuk objektif dengan bobot satu, hasil perhitungan di atas sekaligus menunjukkan jumlah butir soal.

Dengan mendasarkan diri pada hasil perhitungan di atas, kita selanjutnya dapat menghitung jumlah butir soal untuk tiap sel. Untuk maksud itu, kita melakukannya melalui dua cara: (a) mendasarkan diri pada hitungan yang pertama, dan (b) mendasarkan diri pada penghitungan kedua. Berikut contoh penghitungan yang dimaksud.

- (a) - Membaca, ingatan :  $20\% \times 15 = 3$   
 - Membaca, pemahaman :  $30\% \times 15 = 4,50$  (5)  
 - Membaca, penerapan :  $25\% \times 15 = 3,75$  (3)\*  
 - Dan seterusnya

- (b) - Ingatan, membaca :  $20\% \times 15 = 3$   
 - Ingatan, kosa kata :  $10\% \times 15 = 1,50$  (2)  
 - Ingatan, struktur :  $25\% \times 15 = 3,75$  (3)  
 - Dan seterusnya

Pengisian tiap sel selain dengan cara di atas, juga dapat dihitung langsung berdasarkan besarnya presentase pokok bahasan, tingkatan kognitif, dan jumlah keseluruhan skor. Berikut dicontohkan dua buah penghitungan yang dimaksud.

- Sel membaca, ingatan :  $20\% \times 20\% \times 75 = 3$

- Sel kosa kata, ingtan :  $20\% \times 20\% \times 75 = 1,5 (2)$

Keterangan :

- \*) pembulatan angka di atas 0,5 tidak harus ke atas dengan pertimbangan (-pertimbangan) tertentu, yaitu (i) karena kita mendasarkan diri pada jumlah butir soal per pokok bahasan atau per tingkatan kognitif, sehingga jika pembulatan selalu ke atas, ada kemungkinan jumlah butir soal yang di dapat akan melebihi target; (ii) pengukuran kemampuan terhadap pokok bahasan yang bersangkutan dirasa lebih tepat jika tingkatan kognitif yang di bawahnya yang ditambah porsinya.

Dengan penghitungan tersebut, seluruh sel akan dapat diisi dengan jumlah soal yang sesuai dengan perimbangan bobot yang ditentukan. Akan tetapi, berbeda dengan cara sebelumnya, jumlah butir soal untuk tiap pokok bahasan dan tingkatan kognitif justru akan ditemukan kemudian setelah seluruh sel terisi.

**Tabel**  
**Contoh Tabel Spesifikasi Penyusunan Soal Tes**  
**Sumatif Bahasa Prancis**

Tingkat an Kognitif	Ingatan	Pemahaman	Aplikasi	Analisis	Sintesis	Evaluasi	Jumlah
Pokok Bahasan	(C <sub>1</sub> )	(C <sub>2</sub> )	(C <sub>3</sub> )	(C <sub>4</sub> )	(C <sub>5</sub> )	(C <sub>6</sub> )	
CO	4	4	5	2	-	-	15
CE	5	5	5	2	1	1	19
EE	-	2	2	1	1	1	7
Voc	4	4	2	1	-	-	11
Gram	-	6	4	3	1	1	15
Bud	2	1	1	2	1	1	8
<b>Jumlah</b>	<b>15</b>	<b>22</b>	<b>19</b>	<b>11</b>	<b>4</b>	<b>4</b>	<b>75</b>

Kisi-kisi yang dicontohkan di atas, sekaligus memuat bentuk tes objektif dan esai. Dalam model yang lain, keduanya dapat dipisahkan dan dibuat dalam dua kisi-kisi. Untuk contoh kasus di atas, pertama dibuat dengan kisi-kisi objektif dengan jumlah keseluruhan 60 butir. Cara pengisian sel-sel sama dengan yang dicontohkan. Kedua, dibuat kisi-kisi untuk tes esai yang bentuknya lebih sederhana. Misalnya, dalam kolom itu berisi: (i) jenis kemampuan yang diukur, (ii) pokok bahasan dan (iii) skor. Yang pertama diisi dengan tingkatan kognitif yang dapat lebih dari satu, demikian juga kolom kedua. Kolom skor diisi kemungkinan skor

tertinggi, karena skor tes esai adalah berskala dan belum tentu sama antara nomor satu dengan yang lain (misalnya: No. 1: 1-4; No. 2:1-5; No. 3:1-6)

Demikian pula halnya dengan penyusunan soal yang menyangkut tingkatan kognitif, mungkin sekali hanya mengungkap tingkatan  $C_1$  dan  $C_2$ , atau sedikit tingkatan di atasnya. Sebab, kedua tingkatan tersebut lebih mudah disusun alat tesnya daripada yang lain.

Dengan selalu melihat kisi-kisi, kita akan mengetahui pokok bahasan dan tingkatan mana saja yang sudah cukup, masih kurang, atau belum terisi. Caranya ialah begitu selesai menyusun suatu butir soal, segera kita tuliskan cacahan (*tally*) pada tiap sel yang dibuat alat tesnya. Melalui penghitungan cacahan itulah kita akan tahu sel mana saja yang sudah memenuhi target dan yang belum. Dengan cara seperti itu dapat diharapkan tersusunnya sebuah alat penilaian yang komprehensif seperti yang direncanakan.

## **F. TINDAK LANJUT SETELAH PENYUSUNAN TABEL SPESIFIKASI**

Setelah menyusun tabel spesifikasi, langkah selanjutnya adalah menentukan bentuk soal dan menuliskan soal-soal tes/instrumen.

### **1. Menentukan Bentuk Soal**

Ada dua hal yang perlu dipertimbangkan dalam menentukan bentuk soal, yaitu:

- a. Waktu yang tersedia
  - Untuk tes formatif dari bahan yang diselesaikan dalam waktu 4-5 kali pertemuan @ 45 menit kira-kira memerlukan 15 – 20 menit, sedangkan untuk pelajaran yang berlangsung selama 1 jam pelajaran memerlukan waktu kira-kira 5 – 10 menit.
  - Waktu yang digunakan untuk menyelesaikan soal bentuk objektif pilihan ganda kira-kira  $\frac{1}{2}$  - 1 menit untuk setiap butir tes (untuk pilihan ganda sederhana barangkali dapat lebih singkat).
  - Waktu yang diperlukan untuk menyelesaikan soal bentuk uraian tergantung dari berapa lama siswa harus berpikir dan menuliskan jawaban.
- b. Sifat materi yang diteskan

Hal-hal yang harus diperhatikan dalam menulis soal tes adalah:

  - Bahasanya harus sederhana dan mudah dipahami.
  - Suatu soal tidak boleh mengandung penafsiran ganda atau membingungkan.
  - Cara memenggal kalimat atau meletakkan/menata kata-kata perlu diperhatikan agar tidak ditafsirkan salah.
  - Petunjuk pengerjaan, petunjuk ini harus dituliskan dengan jelas sehingga siswa tidak bekerja menyimpang dari yang dikehendaki guru.

### **G. TUGAS**

Buatlah tabel spesifikasi tes sumatif bahasa Prancis dengan jumlah soal 50 butir, 60 butir, dan 90 butir.

### **H. REFERENSI**

Arikunto, Suharsimi. 2004. *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.

Nurgiyantoro, Burhan.2001. *Penilaian dalam Pengajaran Bahasa dan Sastra*. Yogyakarta: PT. BPFE.



## **BAB 6**

### **TES KOMPETENSI KEBAHASAAN**

#### **A. PENDAHULUAN**

Kompetensi kebahasaan seseorang berhubungan dengan sistem bahasa, struktur, kosa kata atau seluruh aspek kebahasaan serta hubungan dari tiap-tiap aspek kebahasaan yang telah tersebut di atas. Kompetensi kebahasaan tersebut berpengaruh pada kemampuan seseorang dalam melakukan tindak bahasa. Maksudnya, ada ahli yang mengatakan bahwa keterampilan berbahasa seseorang itu bergantung pada kualitas dan kuantitas kosa kata. Bahkan lebih lanjut dikatakan bahwa penguasaan kosa kata seseorang akan menunjukkan kemampuan berkomunikasi orang tersebut. Sementara di sisi lain, penguasaan bahasa lisan maupun tertulis tidak akan pernah dapat dilakukan tanpa adanya penguasaan struktur bahasa. Bab ini akan menguraikan tes kompetensi kebahasaan

#### **B. TES KEMAMPUAN BERBAHASA**

Kegiatan berbahasa merupakan *performance* yang menunjukkan kompetensi berbahasa seseorang. Kemampuan berbahasa ini dibedakan menjadi kemampuan memahami dan kemampuan menggunakan. Kemampuan memahami terjadi ketika siswa pembelajar melakukan *decoding* terhadap ujaran yang didengar dan atau tulisan yang dibaca. Sementara kemampuan menggunakan merupakan proses *encoding* yang akan muncul dalam keterampilan berbicara dan keterampilan menulis.

##### **1. Kemampuan Reseptif.**

Salah satu kemampuan seseorang dalam tindakan bahasa bersifat reseptif. Reseptif maksudnya menyerap, atau pasif. Misalnya dalam mata kuliah *Compréhension Orale* dan *Compréhension écrite*. Secara umum kegiatan menyimak, mendengarkan dan membaca masuk dalam kelompok ini.

Menyimak merupakan usaha memahami informasi yang disampaikan dalam bentuk ujaran. Oleh karena itu siswa membutuhkan pengetahuan tentang sistem bunyi dan ejaan bahasa yang dipelajari. Tes kemampuan reseptif menyimak menuntut siswa untuk dapat memahami informasi lisan wacana sederhana tentang tema-tema tertentu, sesuai *niveau* atau tingkat pembelajar/siswa.

Sedangkan membaca merupakan usaha memahami informasi yang disampaikan dalam bentuk tulisan. Oleh sebab itu siswa bukan hanya membutuhkan pengetahuan tentang ejaan, namun juga struktur dan kosa kata. Tes kemampuan reseptif membaca menuntut siswa untuk dapat memahami informasi tertulis wacana sederhana tentang tema-tema tertentu, sesuai dengan tingkat an pembelajar.

## 2. Kemampuan Produktif

Kebalikan dari reseptif adalah produktif. Yaitu kemampuan seseorang dalam menghasilkan tindak bahasa. Hasil di sini maksudnya adalah hasil bicara, berpidato, monolog, dan menulis. Contoh dari kemampuan ini misalnya terdapat dalam mata kuliah *Expression Orale* dan *Expression Ecrite*.

Kegiatan berbicara (contoh lain: pidato, monolog) merupakan kegiatan yang menghasilkan bahasa untuk mengkomunikasikan atau menginformasikan ide dan pikiran si pembicara. Agar mempunyai kemampuan berbicara yang prima, siswa bukan hanya membutuhkan kemampuan tentang sistem bunyi, namun juga unsur bahasa dan ide/pikiran. Unsur bahasa terdiri dari struktur dan kosa kata. Penyampaian ide membutuhkan ketepatan pemilihan diksi, alur penyampaian, kejelasan ide yang semuanya dapat dijadikan komponen untuk menyusun rambu-rambu tes kemampuan produktif berbicara.

Kegiatan menulis merupakan kegiatan yang menghasilkan bahasa dan tulisan. Ide dan pikiran dari siswa (atau dalam hal ini disebut penulis) ditulis dalam rangkaian kata menjadi kalimat. Kalimat demi kalimat disusun dalam suatu paragraf. Keterampilan menulis membutuhkan kejelian dalam mengungkapkan ide pokok dalam 1 (satu) paragraf dan kalimat-kalimat yang mendukung ide tersebut. Tes kemampuan produktif menulis menuntut siswa agar mampu memberikan kejelasan ide dan pikiran dengan menggunakan bahasa yang tepat dalam menulis.

## C. TES KOMPETENSI KEBAHASAAN

Tes kompetensi kebahasaan secara umum dikelompokkan ke dalam tes struktur dan kosakata. Hal ini disebabkan karena kedua aspek bahasa tersebut mendominasi dalam tindak bahasa. Seorang siswa dikatakan mempunyai kompetensi kebahasaan jika siswa tersebut mampu menggunakan bahasa untuk berkomunikasi sesuai dengan konteks situasinya.

Walaupun saat ini tes lebih sering muncul dalam bentuk integratif, yang disesuaikan dengan pendekatan pengajaran bahasa yang dipakai, namun demikian masih dibutuhkan tes kompetensi secara diskrit. Mengingat bahasa Prancis diajarkan di tingkat SMA/ SMK/ MA maka hendaknya penyusunan butir tes juga memperhitungkan tingkat aspek kognitif siswa (C1, C2, C3, C4, C5, dan C6). Materi tes juga disesuaikan dengan *niveau DELF A1*.

### 1. Tes Struktur

Contoh tes struktur dapat dilihat sebagai berikut:

*Choisissez la bonne réponse*

1) Nous  étudiants. ( suis, sommes, sont)

2) Nous allons manger  croissants? (un, une, des)

- 3) Tu as un stylo. C'est  stylo. (ton, ta, tes)
- 4) Ils  le musée. (visite, visites, visitent)
- 5) Georges et Yves  alcool. (bois, buvez, bouvent)
- 

## 2. Tes Kosakata

Contoh tes kosakata dapat dilihat sebagai berikut:

Complétez ces dialogues avec les mots proposés.

**train - métro - avion - gare - taxi - aéroport**

- Pour aller à la  de Lyon, s'il vous plaît ?  
- Prenez le , c'est à 2 stations.
- Pour aller à l' de Roissy ?  
- Prenez le , c'est plus rapide !
- Le  de 8 heures pour Nantes part de Montparnasse ?  
- Oui, c'est bien ça.
- Tu as ton  à quelle heure ?  
- A 9 heures, d'Orly Sud.

## D. JENIS TES BAHASA

### 1. Tes Diskrit

Tes diskrit merupakan tes yang hanya menekankan pada 1 (satu) aspek kebahasaan atau 1 (satu) keterampilan berbahasa pada materi soalnya (Oller, 1979:37).

Perhatikan contoh soal *Phonetiques Française* berikut ini.

**Ecoutez les paires de mots suivantes. Dites si elle sont identiques (=) ou différentes (≠)**

- 1) bain, pain (=), (≠)
- 2) Elaine, Hélène (=), (≠)
- 3) eau, ou (=), (≠)
- 4) etc

Silakan dianalisis, apa yang mau diukur dengan tes seperti di atas? Aspek bahasa apa saja yang bisa diukur? Atau kemampuan berbahasa yang mana, yang bisa diukur dengan tes tersebut?

Ya. Tes di atas ingin mengetahui kemampuan siswa dalam menyimak fonem bahasa Prancis dengan cara membedakan 2 (dua) fonem bahasa Prancis. Tidak ada tes struktur, kosakata, atau kemampuan berbicara

dalam soal di atas. Ini yang dinamakan tes diskrit. Jadi tes yang hanya menekankan 1(satu) aspek kebahasaan atau 1 (satu) keterampilan berbahasa saja.

Munculnya teori tes diskrit ini sejalan dengan teori strukturalisme (dalam bidang linguistic) dan teori behaviorisme (dalam bidang psikologi) (Brown, 1980:217). Aliran ini beranggapan bahwa keseluruhan bahasa itu dapat dipahami dan dikuasai dengan cara memecah asek dan kemampuan berbahasa itu menjadi bagian-bagian terpisah(baca= dipisah-pisahkan). Misalnya bidang fonetik terlihat pada soal di atas tadi yang hanya mengukur kemampuan menyimak dengan cara membedakan 2 (dua) buah fonem, atau bidang struktur, kosakata, kemampuan membaca, kemampuan menulis saja, dll. Sementara fungsi bahasa secara keseluruhan kurang diperhatikan (jika dibandingkan dengan tes integratif yang akan dibahas di bagian lain dalam bab ini).

Pemahaman bahwa kemampuan dan keterampilan kebahasaan seseorang tidak dapat dipisah-pisahkan memunculkan ketidaksetujuan/keraguan akan keberhasilan siswa dalam belajar bahasa. Tes diskrit telah dianggap gagal untuk mengukur kemampuan kebahasaan siswa. Munculnya pendekatan baru dalam pengajaran bahasa, yaitu pendekatan komunikatif melahirkan teori tes yang menolak tes diskrit ini.

## 2. Tes Integratif

Tes integrative muncul sejalan dengan munculnya pendekatan komunikatif dalam pengajaran bahasa. Telah dijelaskan pada bagian sebelumnya bahwa tes integrative ini bertolak belakang dengan tes diskrit.

Tes integrative merupakan tes yang mengukur beberapa aspek kebahasaan atau beberapa keterampilan berbahasa secara integral (menjadi satu/ bersama-sama) (Oller, 1979:37). Minimal ada 2 (dua) aspek kebahasaan atau 2 (dua) keterampilan berbahasa yang dapat diukur. Bandingkan dengan tes diskrit yang hanya mengukur 1 (satu) aspek kebahasaan atau 1 (satu) keterampilan berbahasa!

Perhatikan contoh soal berikut ini.

***Ecrivez une carte postale pendant les vacances!***

Dari soal tersebut di atas marilah kita analisis bersama-sama.

1. Perintah “*Ecrivez*” berarti siswa harus mengerjakan soal dengan cara menulis, yaitu menulis kartu pos. Apakah hanya pengetahuan menulis saja, yang diperlukan? Tentu saja tidak. Siswa harus berpikir “Apa saja yang harus saya ketahui untuk bisa menulis kartu pos?”. Pengetahuan lain yang diperlukan adalah sebagai berikut.
2. Bagaimana tipografi penulisan kartu pos? Bagaimana susunannya? Apakah sama dengan menulis surat resmi?
3. Pilihan kata yang dipergunakan dalam menulis. Mengingat bahwa tes diberikan dalam bahasa Prancis, maka siswa harus mengetahui tata cara menulis kartu pos dalam bahasa Prancis. Misalnya, bagaimana

menulis tanggal, bagaimana menulis sapaannya, bagaimana salam pembuka dan penutupnya, dll.

4. Pengalaman apa saja yang akan diceritakan selama liburan ke dalam kartu pos? Jika menulis cerita tentang liburan, maka sudah pasti liburan itu sudah dilaksanakan atau sedang dan masih dilaksanakan. Hal ini akan berpengaruh pada pemilihan *temps* atau kala waktu, yang dalam hal ini berbeda cara pengungkapannya dengan cara dalam bahasa Indonesia.

Dari hanya 1 (satu) soal untuk menulis kartu pos yang lebarnya kira-kira seperempat kertas folio itu sudah diperlukan lebih dari 2 (dua) aspek kebahasaan.

### 3. Tes Pragmatik

Tes pragmatik muncul seiring dengan pendekatan komunikatif dalam pengajaran bahasa. Sebagaimana diketahui, pendekatan komunikatif dalam pengajaran bahasa populer pada tahun 1970-1980. Pendekatan ini menekankan pada pembentukan kompetensi bahasa dan kemampuan berbahasa dalam fungsi komunikasi yang wajar (seperti dalam situasi pemakaian bahasa yang sesungguhnya). Dalam konteks berkomunikasi akan terdapat hubungan sistematis dan timbal balik antara konteks linguistik (konteks yang ada dalam bahasa) dan konteks ekstralinguistik (konteks yang ada di luar bahasa).

Tes pragmatik tidak menyajikan tes struktur atau kosakata ataupun unsur-unsur kebahasaan lainnya secara diskrit. Tes pragmatik lebih menekankan pada kemampuan siswa untuk berkomunikasi dengan bahasa dalam situasi tertentu. Mungkin akan terjadi kesalahan atau kesilapan dalam berbahasa/ berkomunikasi, namun senyampang kesalahan tersebut tidak mengganggu kelancaran berkomunikasi, maka hal tersebut (baca= kesalahan/ kesilapan) tidak akan dianggap sebagai masalah. Tes pragmatik sudah tentu bersifat integrative. Namun tes integrative belum tentu pragmatik.

#### Contoh tes yang bersifat Pragmatik

##### a) Dikte

Dikte tidak hanya menyangkut konteks linguistic saja, namun juga melibatkan konteks ekstralinguistik. Tes dengan dikte ini menuntut siswa untuk mampu memahami yang didengar, kemudian menuliskannya dalam waktu tertentu/ terbatas (Oller, 1979). Dikte dapat didefinisikan sebagai tes yang diberikan dengan cara dibacakan atau dibunyikan secara keras-keras dan siswa harus menuliskannya (merupakan hasil penyimakan).

Prosedur pelaksanaan dikte dapat dilakukan seperti dalam pemberian tes dengan teknik *cloze-test* (lihat Bab II). Bedanya adalah bahwa dikte disampaikan secara lisan. Misalnya guru membacakan suatu teks namun ada beberapa bagian (kata atau kelompok kata) dalam teks tersebut yang dihilangkan/ dikosongkan. Siswa kemudian diminta untuk menulis/

melengkapi ruang yang kosong dengan kata/ kelompok kata yang sesuai dengan konteks sehingga teks tersebut menjadi sempurna.

Dikte juga dapat dilaksanakan dengan member semacam “gangguan” suara. Hal ini dilakukan agar menyerupai peristiwa berbahasa sesungguhnya yang kadang-kadang bisa jelas, namun juga bisa terganggu oleh suara lain di sekitarnya. Jika hasil dikte tersebut ditulis, maka disebut dengan dikte-komposisi atau *dictation-composition* atau *dicto-compo*. Sedangkan yang tidak ditulis, maksudnya disampaikan secara lisan disebut produksi lisan imitasi.

### **b) Berbicara**

Tes kemampuan berbicara bertujuan untuk mengukur kemampuan siswa dalam mengekspresikan ide dan gagasannya secara lisan. Tes ini dapat dilakukan dengan cara menyuruh siswa membangun suatu cerita berdasarkan gambar (dan atau gambar berseri) yang diperlihatkan oleh guru. Teknik lain yang dapat dipakai untuk mengadakan tes berbicara adalah dengan melakukan wawancara, baik wawancara terstruktur maupun takterstruktur. Penyusunan tes tetap harus mempertimbangkan tingkat pembelajar/ siswa. 3 (tiga) macam tes berikut ini termasuk dalam tes berbicara (bercerita, pidato, diskusi)

### **c) Bercerita**

Bercerita juga merupakan tes kemampuan berbicara. Tes dengan teknik bercerita ini dapat dilakukan dengan memberikam rangsangan berupa gambar (dan atau gambar berseri) seperti pada tes kemampuan berbicara. Namun dapat juga dengan menggunakan teknik lain, misalnya menceritakan aktivitas sehari-hari, olah raga kegemaran, liburan, pengalaman memasak, naik gunung, dan lain-lain. Untuk siswa SMA/SMA/MA dapat juga memasukkan tema pengalaman lomba cerdas-cermat, lomba pidato, dll. Sementara untuk tingkat perguruan tinggi, teknik bercerita dapat dijadikan juga sebagai tes. Misalnya dengan tes bercerita tentang 1 (satu) bab atau 1 (satu) buku yang harus dibaca.

### **d) Berpidato**

Tes ini hamper sama dengan tes bercerita. Dalam praktinya tes berpidato ini dapat diberikan dengan cara simulasi. Maksudnya, seorang siswa bertindak sebagai rektor UNY untuk menyampaikan pidato dies natalis, atau sebagai presiden RI yang berpidato menyambut kedatangan presiden AS.

Adapaun model penilaian tes berpidato (dan juga dapat dipakai sebagai pedoman penilaian bercerita) dapat dilihat di bawah ini (Jakobovits & Gordon).



Soal dapat disajikan sebagai berikut.

*Où sont-ils?*

a. *Dans un bar*

b. *Dans un restaurant*

c. *Dans un hotel*

d. *Dans un magasin*

#### **h) Teknik Cloze**

*Cloze* atau *close* ini dimaksudkan sebagai proses “menutup” sesuatu yang belum lengkap. Tes ini muncul dari teori psikologi Gestal. Teknik penyusunan tes ini adalah dengan memberikan wacana, namun dalam wacana tersebut ada bagian-bagian yang sengaja dikosongkan. Wacana bisa diberikan secara lisan atau tertulis, pada umumnya diberikan secara tertulis. Penggosongan itupun bisa dilakukan dengan menghilangkan kata ke-3, atau ke-5, atau ke-7. Singkatnya menghilangkan kata ke-n.

Pemilihan wacana harus mendapatkan perhatian yang cermat, jangan sampai tes ini hanya berakhir pada hapalan, atau terjadi pengulangan jawaban berkali-kali.

Contoh *Cloze test* dapat dilihat di bawah ini. Penyusunan tes dilakukan dengan cara menghilangkan kata ke-5.

#### ***Fill in the blanks.***

*This is a program to help you increase your reading speed as well as your reading comprehension. The reading passages are (1) \_\_\_\_\_ in order according to (2) \_\_\_\_\_ length. Each short program (3) \_\_\_\_\_ 15 passages and the 100 (4) \_\_\_\_\_, for example, includes the (5) \_\_\_\_\_ from 100 to 149 (6) \_\_\_\_\_. You time yourself and (7) \_\_\_\_\_ to read as fast (8) \_\_\_\_\_ possible, while understanding the (9) \_\_\_\_\_. You take a quiz (10) \_\_\_\_\_ you finish reading. Keep (11) \_\_\_\_\_ of your results, and (12) \_\_\_\_\_ your progress. You will (13) \_\_\_\_\_ a gradual improvement in (14) \_\_\_\_\_ reading. Enjoy reading in English.*

#### **4. Tes Komunikatif**

Pendekatan komunikatif muncul dalam kurikulum 1994. Konsekuensi dari pendekatan komunikatif tersebut adalah munculnya penilaian yang harus komunikatif pula dalam pengajaran bahasa. Maksudnya penilaian terhadap kemampuan berbahasa siswa tidak dilakukan secara diskrit (terpisah-pisah) namun meliputi 4 (empat) keterampilan berbahasa, yaitu menyimak, berbicara, membaca dan menulis. Kemampuan menyimak dan membaca termasuk ke dalam pemahaman yang bersifat reseptif. Kemampuan berbicara dan menulis merupakan kemampuan produksi atau bersifat produktif.



Tes pragmatik sebenarnya telah bersifat komunikatif (Oller & Valette), namun dalam perkembangannya masih perlu mendapatkan penekanan pada adanya konteks. Jadi tes komunikatif selain harus memuat kompetensi gramatikal, sosiolinguistik juga harus kontekstual. Dengan demikian tes komunikatif akan berupa tes kebahasaan, pemahaman, dan penggunaan bahasa. Keempat aspek kebahasaan yang ditekankan harus kontekstual, artinya sesuai dengan penggunaan bahasa yang sesungguhnya. Bagaimana contoh-contoh tes komunikatif? Perhatikan beberapa contoh di bawah ini.

- a) Mencermati siaran langsung kuliah umum Presiden AS Obama di UI melalui pesawat televisi. Selanjutnya siswa membuat resume tentang pidato Obama tersebut. Resume dapat disampaikan secara lisan, dengan tampil secara individu di depan kelas atau dengan cara tertulis.
- b) Mengamati seorang *guide* yang sedang melayani tamu di Borobudur. Kemudian siswa tersebut mempraktikkan diri sebagai *guide*.

Dari 2 (dua) contoh soal tersebut akan terlihat konteks yang berbeda antara pidato dan *guide*. Ragam bahasa yang dipakai juga akan memperlihatkan perbedaan. Silakan diamati dan dipraktikkan.

#### **D. TES KESASTRAAN**

Pengajaran sastra di tingkat pendidikan dasar, menengah dan atas di Indonesia merupakan bagian dari pengajaran bahasa Indonesia. Pengampu mata pelajaran bahasa Indonesia maka sekaligus juga mengajar tentang sastra Indonesia. Di tingkat pendidikan tinggi, sastra merupakan mata pelajaran/ mata kuliah tersendiri. Penilaian terhadap pengetahuan sastra melahirkan tes kesastraan.

Penilaian dalam bidang sastra bukan semata-mata mengungkapkan pengetahuan sastra pembelajar/ siswa (atau dalam hal ini mahasiswa) namun juga dapat berfungsi untuk mengungkapkan kemampuan apresiasi sastra. Hal ini selanjutnya akan data menunjang tercapainya tujuan pengajaran apresiasi sastra (Nurgiantoro, 2001). Apresiasi mendapat penekanan agar mahasiswa mampu membaca, memahami, memberikan reaksi (respon), menganalisis, member kritik terhadap karya sastra. Jadi tes sastra hendaknya bukan hanya berisi hapalan nama sastrawan, karyanya, periodisasinya. Jika kedua hal ini dipadukan, maka pengajaran dan penilaian sastra akan dapat dikatakan berhasil.

##### **1. Taksonomi Tes Kesastraan**

Berbicara mengenai taksonomi berarti berbicara tentang teori Bloom. Bloom membedakan keluaran hasil belajar ke dalam tiga ranah, yaitu kognitif, afektif, dan psikomotorik.

Ranah kognitif masih dibagi lagi menjadi beberapa tingkat kognisi. Ke-6 (enam) tingkatan tersebut, yaitu sebagai berikut:

- C1 = tingkat ingatan  
*Quel est l'auteur de Madame Bovary ?*
- C2 = tingkat pemahaman  
*Quel est le theme principal de l'existensialism?*
- C3 = tingkat penerapan  
*Paraphrsez la poesie ci-dessous*

**Chanson d'automne**

*Les sanglots longs  
 Des violons  
 De l'automne  
 Blessent mon coeur  
 D'une langueur  
 Monotone.  
 Tout suffocant  
 Et blême, quand  
 Sonne l'heure,  
 Je me souviens  
 Des jours anciens  
 Et je pleure;  
 Et je m'en vais  
 Au vent mauvais  
 Qui m'emporte  
 Deçà, delà  
 Pareil à la  
 Feuille morte.*

**Par: Paul Verlaine (1844-1896)**

- C4 = tingkat analisis  
*Analysez le roman Miserables par Victor Hugo.*
- C5 = tingkat sintesis
- C6 = tingkat evaluasi.

**2. Tes Kesastraan Kategori Moody**

Moody membedakan tes kesastraan menjadi 4 (empat) bagian tingkat. Keempat tingkat tersebut adalah sebagai berikut.

a) Informasi

Pertanyaan dalam tingkat informasi ini dapat dikembangkan dari *Qui?*, *Comment?*, *Ou?*, *Pourquoi?*, *etc.*

b) Konsep

Tes yang dapat dikembangkan pada tingkat ini misalnya "Unsur-unsur apa saja yang membedakan antara karya fiksi dan nonfiksi. Berdasarkan jawaban anda, berikan contoh karya fiksi dan nonfiksi tersebut."

Pertanyaan tersebut di atas selain mengharuskan mahasiswa memahami tentang teori fiksi dan nonfiksi, mahasiswa juga harus dapat

menerapkan teori yang dipelajarinya untuk membedakan karya fiksi dan nonfiksi.

c) Perseptif

Pengembangan dari pertanyaan pada tingkat konsep, namun diikuti dengan penilaian/ justifikasi diri sendiri. Misalnya, “Setelah menyaksikan teater Moliere, berikan pendapat Anda.”

d) Apresiasi

Pada tingkat apresiasi ini unsur linguistic dalam sastra dapat dikembangkan sebagai tes. Reseptif pembaca sastra juga diperlukan dalam tes jenis ini.

## E. TUGAS

### 1. Tugas Kelompok

Lakukan wawancara terhadap guru-guru bahasa Prancis SMA/ SMK/ MA, tentang bagaimana mereka memberikan tes/tugas untuk siswanya. Selanjutnya kategorikan tes/tugas tersebut berdasarkan jenis-jenis tes yang telah dibahas dalam bab ini.

### 2. Tugas Individu

Mencari soal UAS bahasa Prancis di SMA/ SMK/ MA. Analisis jenis-jenis tes yang muncul, kelompokkan ke dalam tingkat Kognisi (C1, C2, C3, dan seterusnya sampai C6 jika ada).

## H. REFERENSI

Nurgiyantoro, Burhan.2001. *Penilaian dalam Pengajaran Bahasa dan Sastra*. Yogyakarta: PT. BPFE

Arikunto S. (1993). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.

<http://www.enpc.fr>

<http://phonetique.free.fr>

<http://www.edu365.cat>

[http://th.alliance\\_francaise.or.th](http://th.alliance_francaise.or.th)

<http://kkitao.e-leraning-server.com>

<http://www.feelingsurf.net>

## **BAB 7**

### **ANALISIS BUTIR SOAL**

#### **A. PENDAHULUAN**

Secara keseluruhan sebuah tes mungkin tidak terpercaya, akan tetapi tentunya tidak semua butir soal yang ada perlu direvisi. Sebab, tentunya terdapat sejumlah butir soal yang telah memenuhi criteria kelayakan dan karenanya dapat dipertahankan. Untuk memilih butir-butir soal yang layak, dan sebaliknya perlu direvisi, dapat ditentukan berdasarkan kerja analisis butir soal. Analisis butir soal akan memberikan jawab terhadap maksud itu secara dapat dipertanggungjawabkan karena ia sanggup memberikan informasi secara rinci tentang “keadaan” masing-masing butir soal, yaitu berdasarkan tingkat kesulitan (*item difficulty*) dan daya pembeda (*item discrimination*). Sebuah butir soal dinyatakan layak jika indeks tingkat kesulitan dan daya pembedanya memenuhi standar yang ditentukan.

#### **B. LANGKAH-LANGKAH ANALISIS BUTIR SOAL**

Langkah-langkah yang dilakukan untuk menganalisis butir soal adalah sebagai berikut.

1. Mengurutkan skor pada lembar jawaban siswa dari skor yang tertinggi berturut-turut sampai yang yang terbawah.
2. mengambil sebanyak 27,5% dari jumlah siswa dari skor yang tertinggi dan 27,5% dari skor yang terendah. Kelompok yang pertama disebut kelompok tinggi (kelompok siswa yang skornya tinggi), sedangkan yang kedua disebut kelompok rendah, dan sisanya sebagai kelompok tengah. Pembagian menjadi ketiga kelompok tersebut terutama disarankan jika jumlah siswa cukup besar, sebaliknya, jika hanya sedikit, cukup dibedakan menjadi kelompok tinggi dan kelompok rendah saja.
3. menganalisis jawaban yang benar atau salah per butir soal per siswa. Analisis ini hanya dilakukan terhadap jawaban siswa kelompok tinggi dan kelompok rendah, sedangkan kelompok tengah ditinggalkan. Berdasarkan analisis atau identifikasi ini akan dapat dihitung indeks tingkat kesulitan dan daya beda masing-masing butir soal.

Tabel

**Analisis Butir Soal Kelompok Tinggi dan Kelompok Rendah untuk  
Persiapan Penghitungan Indeks Tingkat Kesulitan dan Daya Beda  
Kelompok Tinggi**

Nomor urut subjek	Nomor Butir Soal											Jumlah Skor
	1	2	3	4	5	6	7	8	...	...	40	
1	1	1	1	0	1	1	1	1	...	...	1	36
2	1	1	0	1	1	1	1	1	...	...	1	34
3	1	1	1	0	0	1	1	1	...	...	1	30
4	1	0	1	1	0	1	1	0	...	...	1	30
5	0	1	1	1	1	0	1	1	...	...	0	28
6	1	1	1	0	1	0	1	0	...	...	1	28
<b>Jumlah</b>	<b>5</b>	<b>5</b>	<b>5</b>	<b>3</b>	<b>4</b>	<b>4</b>	<b>6</b>	<b>4</b>	...	...	<b>5</b>	

#### Kelompok Rendah

Nomor urut subjek	Nomor Butir Soal											Jumlah Skor
	1	2	3	4	5	6	7	8	...	...	40	
1	1	1	0	1	0	0	1	1	...	...	1	20
2	0	0	1	1	1	0	1	1	...	...	0	19
3	0	1	1	0	0	1	0	1	...	...	1	19
4	1	1	1	0	0	1	1	0	...	...	0	18
5	0	1	0	0	1	1	1	0	...	...	0	16
6	0	0	1	1	1	0	1	1	...	...	1	15
<b>Jumlah</b>	<b>2</b>	<b>4</b>	<b>3</b>	<b>3</b>	<b>3</b>	<b>3</b>	<b>5</b>	<b>4</b>	...	...	<b>3</b>	

### C. TINGKAT KESULITAN BUTIR SOAL

Tingkat kesulitan (*item difficulty*) adalah pernyataan tentang seberapa mudah atau sulit butir soal bagi siswa yang dikenai pengukuran (Oller dalam Burhan Nurgiyantoro, 2001: 138). Butir soal yang baik adalah yang tingkat kesulitannya cukup, tidak terlalu mudah atau terlalu sulit. Butir soal yang terlalu mudah atau sulit sama tidak baiknya karena keduanya tidak dapat membedakan antara siswa kelompok tinggi dan kelompok rendah. Butir soal yang demikian tidak memberikan informasi apa-apa tentang perbedaan prestasi tiap individu (Noll dalam Burhan Nurgiyantoro, 2001: 138).

Tingkat kesulitan suatu butir soal dinyatakan dengan sebuah indeks yang berkisar antara 0,0 – 1,0. Indeks 0,0 berarti soal tersebut sangat sulit karena tak seorang siswa pun yang dapat menjawabnya. Sebaliknya, indeks 1,0 berarti soal tersebut sangat muda karena semua siswa yang dapat menjawab dengan betul. Oller dalam Burhan Nurgiyantoro (2001: 138) mengemukakan bahwa suatu butir soal dinyatakan layak jika indeks tingkat kesulitannya berkisar antara 0,15 –

0,85. Indeks yang di luar itu berarti soal itu terlalu mudah atau terlalu sulit sehingga ia perlu diganti atau direvisi.

Untuk menghitung indeks tingkat kesulitan digunakan rumus:

$$IF = \frac{FH + FL}{N}$$

Keterangan:

IF = (*Item Facility*) indeks tingkat kesulitan

FH = (*Frequency High*) jumlah jawaban betul kelompok tinggi

FL = (*Frequency Low*) jumlah jawaban betul kelompok rendah

N = Jumlah siswa kedua kelompok

Berdasarkan data pada tabel ... kita dapat menghitung indeks tingkat kesulitan tiap butir soal yang ada. Misalnya butir nomor 1.

$$\text{Butir nomor 1} = \frac{5 + 2}{12} = 0,58$$

Hitung pula indeks tingkat kesulitan tiap butir soal yang lain!

#### D. DAYA PEMBEDA BUTIR SOAL

Daya pembeda (*item discrimination*) maksudnya adalah seberapa besar suatu butir soal dapat membedakan antara siswa kelompok tinggi dan siswa kelompok rendah. Butir soal tersebut adalah dapat membedakan kedua kelompok secara layak. Hal tersebut berdasarkan logika bahwa siswa dari kelompok tinggi dapat menjawab dengan betul yang lebih banyak daripada kelompok rendah (Oller, 1979: 251; Ebel, 1979: 258).

Daya pembeda dihitung berdasarkan perbedaan jumlah jawaban betul untuk tiap butir soal antara kelompok tinggi dan kelompok rendah. Jika terjadi kelompok rendah menjawab betul lebih banyak daripada kelompok tinggi, butir soal tersebut kurang baik karena menyalahi logika. Lebih jauh, hal itu berarti bahwa butir soal tersebut tidak terpercayai karena tidak memiliki konsistensi internal (*internal consistency*, Olley, 1979: 284).

Besar kecilnya daya pembeda suatu butir soal dinyatakan dengan suatu indeks yang berkisar antara -1,00 sampai dengan 1,00. Indeks yang semakin besar atau mendekati 1,00, butir soal yang bersangkutan semakin baik sebab semakin nyata perbedaan antara kelompok rendah dan kelompok tinggi. Indeks negatif berarti siswa kelompok rendah justru menjawab dengan betul lebih banyak daripada kelompok tinggi.

Untuk mencari indeks daya beda suatu butir soal dilakukan dengan cara sebagai berikut: jumlah jawaban betul kelompok tinggi dikurangi jumlah jawaban betul kelompok rendah kemudian dibagi dengan jumlah subjek kelompok tinggi atau kelompok rendah (27,5 persen). Jika ditulis

dengan rumus indeks daya beda (IDB, atau ID) tersebut adalah sebagai berikut.

$$ID = \frac{FH - FL}{n}$$

ID = (Item discrimination) indeks daya pembeda yang dicari

FH = Jumlah jawaban betul kelompok tinggi

FL = Jumlah jawaban betul kelompok rendah

n = Jumlah subjek kelompok tinggi atau rendah, atau 27,5 persen subjek.

Berdasarkan data-data yang ada pada Tabel 4.8 di atas kita dapat menghitung indeks daya pembeda tiap butir soal, misalnya dicontohkan di bawah ini.

- Butir nomor 1  $= \frac{5-2}{6} = 0,50$
- Butir nomor 2  $= \frac{5-4}{6} = 0,17$
- Butir nomor 3  $= \frac{5-3}{6} = 0,33$
- Butir nomor 4  $= \frac{3-3}{6} = 0,00$
- Butir nomor 7  $= \frac{6-5}{6} = 0,17$
- Butir nomor 40  $= \frac{5-3}{6} = 0,33$

Butir soal yang baik indeks daya pembeda paling tidak harus mencapai 0,25 atau bahkan 0,35 (Oller, 1979: 252). Butir soal yang indeks daya pembedanya kurang dari 0,25 dianggap tidak layak, dan karenanya perlu revisi atau diganti. Indeks yang kurang dari 0,25 berarti butir soal yang bersangkutan kurang mampu membedakan antara kelompok tinggi dan rendah. Indeks daya pembeda butir 2, 4, dan 7 di atas kurang 0,25, sedang butir nomor 1, 3, dan 40 di atas 0,25.

Sebuah butir soal dinyatakan layak indeks tingkat kesulitan maupun daya pembeda dapat memenuhi standar yang telah ditetapkan. Indeks tingkat kesulitan mungkin telah memenuhi persyaratan, tetapi jika indeks daya pembedanya rendah, di bawah 0,25, butir soal yang bersangkutan tetap dinyatakan kurang layak. Demikian pula sebaliknya. Hal ini biasanya cukup berat dipenuhi terutama yang berkaitan dengan tuntutan indeks daya pembeda. Akan tetapi, jika butir-butir soal tes memenuhi persyaratan tersebut, tingkat kepercayaan tes akan menjadi lebih tinggi.

Penghitungan indeks tingkat dan indeks daya beda dapat dilakukan dengan cara lain, yaitu dengan mempergunakan tabel analisis butir soal (Fan, 1952. *Item analysis table*). Untuk maksud ini, terlebih dahulu kita harus menghitung proporsi jawaban betul kelompok tinggi (PH) dan

proporsi jawaban betul kelompok rendah (PL). Proporsi diperoleh dengan: jumlah jawaban betul (FH atau FL) dibagi jumlah subjek (n, 27,5 persen). Setelah besarnya proporsi masing-masing kelompok ditemukan, langkah selanjutnya adalah konsultasi tabel.

Berdasarkan angka-angka pada tabel 4.10 di atas, misalnya, kita dapat menghitung besarnya proporsi yang dimaksud. Butir soal nomor 1:  $PH = 5: 6 = 0,83$  sedang  $PL = 2: 6 = 0,33$  (n, 2,75 persen = 6); butir soal nomor 2:  $PH = 5:6 = 0,83$  sedang  $PL = 4: 6 = 0,67$ . Demikian seterusnya dengan nomor-nomor butir soal yang lain.

Walau terjadi perbedaan kecil indeks yang dihasilkan, penghitungan seperti di atas menghasilkan kesimpulan yang sama dengan yang dilakukan seperti pada tabel 4.10 di atas. Artinya, indeks-indeks butir soal yang dinyatakan layak atau tidak layak pada tabel 4.10 juga dinyatakan secara sama pada tabel 4.11.

Kerja analisis butir soal, selain dapat dilakukan dengan cara-cara seperti di atas, juga dapat dilakukan dengan mempergunakan jasa komputer. Misalnya, kita dapat mempergunakan program SPSS (*Statistical Package for Social Sciences*), Lisrel, dan Iteman (*Item Analysis*)

## E. ANALISIS DISTRAKTOR

Penentuan revisi terhadap suatu butir soal tidak semata-mata berdasarkan besarnya indeks tingkat kesulitan dan daya beda saja, melainkan juga bagaimana sebaran distribusi frekuensi jawaban pada alternatif yang disediakan. Dengan kata lain, kita perlu juga menganalisis efektivitas butir-butir pengecoh (*distractors*) untuk tiap butir soal.

Dasar pemikiran analisis efektivitas distraktor tersebut tidak berbeda halnya dengan daya beda suatu butir soal: harus ada perbedaan frekuensi jawaban antara siswa kelompok tinggi dan kelompok rendah. Untuk setiap alternatif jawaban yang betul, kelompok tinggi harus memilih secara lebih banyak karena besarnya selisih jawaban betul inilah yang akan menentukan besar kecilnya indeks daya beda. Sebaliknya, alternatif-alternatif jawaban yang merupakan distraktor, kelompok rendah harus memilih secara lebih banyak. Di samping itu, semua alternatif jawaban yang disediakan harus ada siswa yang memilihnya. Namun, jika hanya ada satu orang yang memilih distraktor, ia harus dari kelompok rendah. Jika yang memilih distraktor itu dari kelompok tinggi dan rendah dengan jumlah (misalnya sama-sama dua orang), distraktor tersebut kiranya masih dapat dipandang layak. Jika terjadi penyimpangan, terhadap hal-hal tersebut, suatu butir soal disarankan untuk direvisi. Kegiatan butir soal tidak harus mencakup butir soal dengan seluruh alternatif jawabannya, melainkan cukup pada distraktor yang mengalami penyimpangan saja.

Untuk mengetahui efektivitas tiap alternatif jawaban, atau sebaliknya, ada penyimpangan, perlu dilakukan kegiatan analisis



distraktor, karena dari kegiatan itulah akan diketahui sebaran (distribusi) frekuensi jawaban. Langkah pertama yang dilakukan, yaitu setelah kita memisahkan lembar-lembar jawaban untuk siswa kelompok tinggi dan kelompok rendah, adalah meneliti pilihan terhadap alternatif-alternatif jawaban semua butir soal untuk seluruh siswa.

Setelah kita mendapatkan data-data sebaran sebaran distribusi frekuensi jawaban siswa baik dari kelompok tinggi maupun kelompok rendah seperti di atas, kita langsung dapat melakukan analisis distraktor tiap butir soal untuk mengetahui keefektifitasannya. Akan tetapi, kita juga dapat mengalihkan ke dalam tabel lain yang sekaligus untuk menganalisis butir soal untuk mencari indeks tingkat kesulitan dan daya beda. Berikut dicontohkan model yang kedua, dengan pertimbangan bahwa model tersebut dapat menghemat tempat. Di samping itu, dengan mempergunakan model itu sekaligus dapat diketahui ketiga hal yang dianalisis (indeks tingkat kesulitan, indeks daya beda, dan efektivitas distraktor), sehingga hal-hal yang perlu direvisi dapat dilihat atau ditentukan secara menyeluruh. Bentuk tabel yang demikian itulah yang disarankan untuk dilakukan (dan ditampilkan) dalam kegiatan analisis butir soal.

## F. ANALISIS BUTIR SOAL ESAI

Untuk tes yang berbentuk esai, penghitungan indeks tingkat kesulitan dan indeks daya beda dipergunakan rumus (Noll dkk, 1979: 214-215) berikut

$$\text{Indeks Tingkat Kesulitan} = \frac{Sh + S1 - (2N \times \text{Skor min})}{2N \times (\text{Skor maks} - \text{Skor min})}$$

$$\text{Indeks Daya Beda} = \frac{Sh - S1}{N (\text{Skor maks} - \text{Skor min})}$$

Catatan:

- Sh = Jumlah skor betul kelompok tinggi
- S1 = Jumlah skor betul kelompok rendah
- Skor<sub>maks</sub> = Skor maksimal suatu butir soal
- Skor<sub>min</sub> = Skor minimal suatu butir soal
- N = Jumlah subjek kelompok tinggi atau rendah (27,5 persen).

Langkah- langkah yang ditempuh sama dengan pada analisis butir soal objektif di atas, tetapi pada langkah identifikasi jawaban benar dan

salah berbeda. Pada tes esai, jawaban benar biasanya berskala 1 sampai dengan 5. Artinya, skor minimal (jawaban salah) 1 dan skor maksimal (jawaban tepat) 5.

Berdasarkan data-data di atas berikut dicontohkan penghitungan indeks tingkat kesulitan (IF) dan daya beda (ID).

$$\begin{aligned} \text{Butir nomor 1: IF} &= \frac{22+13-(2 \times 6 \times 1)}{2 \times 6 \times (5-1)} \\ &= 0,48 \end{aligned}$$

$$\text{ID} = \frac{22-13}{6 \times (5-1)} = 0,38$$

$$\begin{aligned} \text{Butir nomor 2: IF} &= \frac{20+10-(2 \times 6 \times 1)}{2 \times 6 \times (5-1)} \\ &= 0,38 \end{aligned}$$

$$\text{ID} = \frac{20-10}{6 \times (5-1)} = 0,42$$

$$\begin{aligned} \text{Butir nomor 3: IF} &= \frac{23+11-(2 \times 6 \times 1)}{2 \times 6 \times (5-1)} \\ &= 0,46 \end{aligned}$$

$$\text{ID} = \frac{23-11}{6 \times (5-1)} = 0,50$$

$$\begin{aligned} \text{Butir nomor 4: IF} &= \frac{25+21-(2 \times 6 \times 1)}{2 \times 6 \times (5-1)} \\ &= 0,71 \end{aligned}$$

$$\text{ID} = \frac{25-21}{6 \times (5-1)} = 0,17$$

Dengan mempergunakan kriteria kelayakan butir soal seperti di atas, indeks tingkat kesulitan 0,15 sampai 0,85 dan indeks skor daya beda minimal 0,25, dari keempat butir soal di atas hanya nomor 4 yang kurang layak karena indeks daya bedanya yang terlalu kecil. Hasil analisis butir soal untuk tes bentuk esai pun perlu disajikan dalam tabel.

## G. TUGAS

Matrik jawaban siswa terhadap butir soal tertera seperti di bawah ini. Carilah indeks kesukaran dan Daya Beda butir soal tersebut.

No Siswa	Nomor Item										Skor yang diperoleh
	1	2	3	4	5	6	7	8	9	10	
1	1	1	0	0	1	0	0	1	0	0	4
2	0	1	0	0	1	1	0	1	0	1	5
3	0	1	0	1	1	1	0	1	1	1	7
4	1	1	1	1	1	1	1	1	1	1	10
5	0	1	0	0	0	0	0	0	0	0	1
6	1	0	1	1	1	1	0	1	1	1	8
7	0	1	1	1	0	0	1	1	0	0	5
8	1	1	1	0	1	1	0	1	1	1	8

## H. REFERENSI

Arikunto, Suharsimi. 2006. *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.

Nurgiyantoro, Burhan.2001. *Penilaian dalam Pengajaran Bahasa dan Sastra*. Yogyakarta: PT. BPFE.

## BAB 8 UJI COBA INSTRUMEN

### A. PENDAHULUAN

Bab ini akan membahas tentang uji coba penghitungan ketepatan alat tes yang jajarannya dan rumus-rumusnya terdapat dalam BAB IV. Selain itu dalam bab ini juga akan membahas tentang analisis distraktor dan revisi butir soal.

### B. PENGHITUNGAN PENGUJIAN ALAT TES

#### 1. Reliabilitas Spearman Brown

##### a. Pertama

Menganalisis lembar jawaban siswa. Dalam hal ini yang akan dicontohkan adalah menghitung koefisien soal tes objektif yang dikotomis. Artinya, jawaban benar skor 1 (satu) dan jawaban salah skor 0 (nol). Kemudian masukan tabulasi dengan menggunakan teknik belah dua yang dalam penghitungan ini menggunakan nomor soal ganjil dan nomor soal genap. Belah dua juga dapat dilakukan dengan pemilihan nomor awal dan nomor akhir.

Tabel akan terlihat seperti berikut ini. (Nurgiyantoro, 2001)

Nomor Urut Subjek	Nomor Butir Soal										Skor Total	Skor Ganjil	Skor Genap
	1	2	3	4	5	6	7	8	9	10			
1.	1	1	0	1	0	1	1	1	1	1	8	3	5
2.	1	0	1	1	1	1	0	0	0	1	6	3	3
3.	0	1	1	1	0	1	1	0	1	1	7	3	4
4.	0	1	1	1	1	0	1	1	1	1	8	4	4
5.	1	1	0	0	1	1	0	1	0	0	5	2	3
6.	1	1	1	1	0	1	1	1	1	1	9	4	5
7.	1	0	1	0	0	1	1	1	0	1	6	3	3
8.	0	0	1	1	1	0	0	1	0	0	4	2	2

##### b. Kedua

Penghitungan langkah awal di atas dilanjutkan dengan penghitungan di bawah ini. Skor ganjil sebagai  $X_1$  dan skor genap sebagai  $X_2$ .

Nomor Urut Subjek	$X_1$	$X_1^2$	$X_2$	$X_2^2$	$X_1X_2$
1.	3	9	5	25	15
2.	3	9	3	9	9
3.	3	9	4	16	12
4.	4	16	4	16	16
5.	2	4	3	9	6
6.	4	16	5	15	20
7.	3	9	3	9	9
8.	2	4	2	4	4
<b>n= 8</b>	<b><math>\sum X_1= 24</math></b>	<b><math>\sum X_1^2=76</math></b>	<b><math>\sum X_2= 29</math></b>	<b><math>\sum X_2^2=113</math></b>	<b><math>\sum X_1X_2= 91</math></b>

c. Ketiga

Mulai hitung koefisien nilai r dengan menggunakan rumus *product moment* angka kasar, seperti terlihat di bawah ini.

$$\begin{aligned}r_{1.2} &= \frac{(8 \times 91) - (24 \times 29)}{\sqrt{(8 \times 76 - 24^2)(8 \times 113 - 29^2)}} \\ &= \frac{718 - 686}{\sqrt{(32 \times 63)}} \\ &= \frac{32}{\sqrt{2.016}} \\ &= \frac{32}{44,9} \\ &= \mathbf{0,713}\end{aligned}$$

d. Keempat

Nilai koefisien **0,713** dalam penghitungan di langkah ketiga di atas baru menghitung koefisien separuh soal. Untuk mengetahui reliabilitas keseluruhan soal, maka dihitung dengan menggunakan rumus Spearman Brown sebagai berikut.

$$\begin{aligned}\text{Reliabilitas seluruh tes} &= \frac{2 \times \text{reliabilitas separuh tes}}{1 + \text{reliabilitas separuh tes}} \\ &= \frac{2 \times 0,713}{1 + 0,713} \\ &= \frac{1,426}{1,713} \\ &= \mathbf{0,832}\end{aligned}$$

e. Kelima

Memberi makna nilai 0,832. Nilai koefisien 0,832 ini termasuk dalam tingkat keterpercayaan yang **tinggi**. Biasanya nilai koefisien **r** adalah 1,0 (pada umumnya di bawah 1,00). Jadi nilai 0,832 itu mendekati

angka 1,00. Selain itu menentukan tingkat reliabilitas dapat menggunakan tabel nilai-nilai korelasi.

2. Rumus KR20

a. Pertama

Sama seperti langkah pertama dalam menganalisis dengan rumus *Spearman Brown*. Buat tabulasi jawaban benar (skor 1) dan jawaban salah (skor 0) (Nurgiyantoro, 2001).

Nomor Urut Subjek	Nomor Butir Soal										Jumlah Skor
	1	2	3	4	5	6	7	8	9	10	
1.	1	1	0	0	1	0	1	1	1	0	6
2.	1	0	0	1	1	0	0	0	1	1	5
3.	1	1	1	0	1	1	1	1	0	1	8
4.	0	1	0	0	0	1	1	1	0	0	4
5.	0	0	0	0	1	0	1	0	1	0	3
6.	1	1	1	1	1	1	1	0	1	1	9
7.	1	1	0	1	0	1	0	0	1	1	6
8.	1	1	1	0	0	1	1	0	1	1	7
Jumlah	6	6	3	3	5	5	6	3	6	5	48
P	.75	.75	.375	.375	.625	.625	.75	.375	.75	.625	
Q	.25	.25	.625	.625	.375	.375	.25	.625	.25	.375	
Pq	.19	.19	.234	.234	.234	.234	.19	.234	.19	.234	2.164

b. Kedua

Dengan menggunakan Excel atau kalkulator, maka akan diketahui

$$n = 10$$

$$S = 1,87$$

$$X \text{ (rerata)} = 6$$

$$\sum pq = 2,164$$

c. Ketiga

Data-data pada langkah kedua masukkan ke dalam rumus KR20

$$r = \frac{n}{n-1} \left( 1 - \frac{\sum pq}{S^2} \right)$$

$$10 \quad 2,164$$

$$\begin{aligned}
&= \frac{10}{10-1} \left( 1 - \frac{2,164}{1,87^2} \right) \\
&= \frac{10}{9} \left( 1 - \frac{2,164}{3,5} \right) \\
&= 1,11 (1-0,619) \\
&= 1,11 \times 0,381 \\
&= \mathbf{0,423}
\end{aligned}$$

d. Memaknai

Besarnya koefisien korelasi antara 0 samapi dengan 1,0. Jika nilai koefisien 1,00 maka soal/ instrument tes tersebut dinyatakan sempurna atau sangat tepercaya. Berdasarkan penghitungan di atas nilai koefisien 0,423 bahkan lebih kecil dari nilai yang dipersyaratkan bagi tes buatan guru (minimal tes buatan guru mempunyai koefisien 0,60).

3. Rumus KR21

Dengan data yang sama dengan penghitungan memakai rumus KR20, maka kita tinggal memasukkan data yang dibutuhkan oleh Rumus KR21 sebagai berikut ini.

$$\begin{aligned}
r &= \frac{n}{n-1} \left( 1 - \frac{X(n-X)}{nS^2} \right) \\
&= \frac{10}{10-1} \left( 1 - \frac{6(10-6)}{10 \times 1,87^2} \right) \\
&= \frac{10}{10-1} \left( 1 - \frac{6(10-6)}{10 \times 1,87^2} \right) \\
&= \frac{10}{9} \left( 1 - \frac{6 \times 4}{10 \times 3,5} \right) \\
&= 1,11 \left( 1 - \frac{24}{34,97} \right)
\end{aligned}$$

$$= 1,11 (1-0,686)$$

$$= 1,11 \times 0,314$$
$$= \mathbf{0,348}$$

Perlu diingat bahwa Rumus KR20, KR21 ini menghindari sebaran yang merata disisi kanan dan sisi kiri (akan terlihat jika dibuat dalam kurva). Semakin besar varian (koefisien S-nya) maka instrumen tes akan semakin bagus/ terpercaya.

### C. ANALISIS DISTRAKTOR

Dalam mempersiapkan alat tes yang sah dan tepercaya diperlukan langkah-langkah yang tidak sederhana. Pada bab-bab sebelumnya telah dibahas tentang analisis butir soal (anabut), tingkat kesulitan dan daya beda (DB). Revisi terhadap suatu butir soal juga ditentukan oleh efektivitas distraktor.

Distraktor dari kata *distract* atau *distraire* yang artinya mengecoh. Jadi dalam membuat soal (dalam hal ini pilihan ganda) tim penyusun juga harus mempersiapkan pengecoh. Bagaimana menyusun soal pilihan ganda dengan distraktor? Simaklah beberapa kiat berikut ini.

1. Dalam tes pilihan ganda yang disebut dengan distraktor adalah pilihan jawaban yang bukan jawaban sebenarnya.
2. Distraktor harus mampu mengecoh, jadi harus ada yang memilih jawaban tersebut, misalnya siswa yang mempunyai kompetensi rendah. Perlu diingat bahwa di kelas terdapat kelompok atas dan kelompok bawah.
3. Sebagai akibat dari ketentuan di atas, jika ada distraktor yang tidak dijlilih, maka distraktor tersebut harus direvisi dengan alternatif jawaban yang lain.
4. Distraktor memang harus dapat mengecoh atau “menipu” namun bukan “menjebak” siswa dalam memilih jawaban yang tepat.

Nurgiyantoro (2001) memberikan contoh analisis butir soal sehingga kita dapat mengetahui soal soal-soal ataupun jawaban yang harus direvisi. Silakan simak tabel berikut ini.



### Analisis Butir Soal

Nomor Butir Soal dan Opsi	KT	KR	PH	PL	P	R	Keterangan
1. (a) b c d	15 - 2 3*)	10 4 6 -)	0,75	0,50	0,63	0,27	layak dengan revisi
2. a b c (d)	5 - - 15	7 2 4 7	0,75	0,35	0,55	0,41	layak
3. a (b) c d	5 10 3*) 2	8 7 2*) 3	0,50	0,35	0,42	0,16*)	tak layak
4. a (b) c d	6 8 4*) 2*)	7 8 3*) 2*)	0,40	0,40	0,40	0,00*)	tak layak
dst...							

#### Keterangan

KT = Kelompok Tinggi

KR = Kelompok Rendah

() = Alternatif jawaban benar

\*) = soal atau pilihan tidak layak, sehingga harus direvisi

PH = Proporsi jawaban benar kelompok tinggi

PL = Proporsi jawaban benar kelompok rendah

P = Indeks Tingkat Kesulitan

r = Daya Beda

#### D. REVISI BUTIR SOAL

Berdasarkan pada tabel di atas, marilah kita analisis soal-soal yang memerlukan revisi dan alasan melakukan revisi.

##### 1. Butir nomor 1 (satu);

- indeks tingkat kesulitan memenuhi syarat,
- daya beda memenuhi syarat,
- alternatif pilihan jawaban d harus direvisi karena tidak dapat membedakan antara kelompok tinggi dan kelompok rendah. Justru kelompok tinggi ada yang terkecoh (3 orang), sementara kelompok rendah tidak ada yang terkecoh (0 orang).

Kesimpulannya soal sudah layak dengan syarat distraktor d dalam alternatif jawaban direvisi.

##### 2. Butir nomor 2 (dua)= soal sangat layak, tidak perlu mengalami revisi.

##### 3. Butir nomor 3 (tiga)

- indeks tingkat kesulitan memenuhi syarat,
- syarat daya beda TIDAK terpenuhi, karena kurang atau tidak dapat membedakan siswa kelompok tinggi dan siswa kelompok rendah,
- distraktor pilihan jawaban c harus direvisi karena tidak dapat membedakan antara kelompok tinggi dan kelompok rendah. Justru kelompok atas ada yang terkecoh (3 orang), sementara kelompok rendah yang terkecoh 2 orang.

Kesimpulannya soal HARUS direvisi atau DIGANTI

4. Butir nomor 4 (empat)

- indeks tingkat kesulitan memenuhi syarat,
- syarat daya beda TIDAK terpenuhi, karena kurang atau tidak dapat membedakan siswa kelompok tinggi dan siswa kelompok rendah,
- distraktor pilihan jawaban c dan d harus direvisi karena tidak dapat membedakan antara kelompok tinggi dan kelompok rendah. Justru lebih banyak kelompok tinggi yang terkecoh daripada kelompok rendah.

Kesimpulannya soal HARUS direvisi atau DIGANTI.

#### D. TUGAS.

Tugas kelompok dengan jumlah anggota 3-5 orang. Silakan mengambil sampel soal uji coba SNMPTN yang diselenggarakan oleh suatu lembaga bimbingan belajar di Yogyakarta. Temukan subjek sebanyak 20 orang. Buatlah analisis terhadap 20 butir soal. Susun hitungan dalam tabel, kemudian maknai masing-masing nilai koefisien yang ada. Tentukan soal-soal yang perlu direvisi atau diganti. Selamat mengerjakan!

#### E. REFERENSI

Nurgiyantoro, Burhan.2001. *Penilaian dalam Pengajaran Bahasa dan Sastra*. Yogyakarta: PT. BPFE

[jurnal](#) ramainas (on line)

<http://www.scribd.com>

## BAB 9 ANALISIS DATA HASIL TES

Setelah pelaksanaan tes selesai dilakukan, kegiatan guru berikutnya adalah mengoreksi pekerjaan siswa, yang sebenarnya tidak lain dari menentukan dan menghitung jumlah jawaban betul siswa. Kegiatan ini disebut juga menghitung jumlah skor. Jadi, jumlah skor adalah jumlah jawaban betul siswa terhadap butir-butir tes yang dikerjakan.

Skor yang diperoleh dengan cara di atas adalah skor mentah (*raw scores*), dan masih harus diolah lebih lanjut untuk dijadikan **nilai jadi**. Angka-angka skor tersebut belum dapat menafsirkan hasil tes, menentukan kedudukan siswa di antara kelompoknya, menentukan prestasi kelas, dan lain-lain, sebelum diolah dengan teknik-teknik tertentu. Analisis data skor biasanya dilakukan dengan mempergunakan teknik statistik walau dalam bentuk yang sederhana. Beberapa teknik yang dimaksud berikut akan dibicarakan dengan contoh-contoh seperlunya.

### A. Penyusunan Distribusi Frekuensi

Untuk memudahkan analisis selanjutnya, data-data skor tersebut perlu disusun menurut teknik tertentu, misalnya dengan membuat tabulasi dan kemudian menghitung frekuensi masing-masing skor. Jika kita memilih cara itu, berarti kita akan menyusun distribusi frekuensi yang akan memudahkan kita untuk mengerjakan penghitungan selanjutnya, misalnya menghitung nilai rata-rata, simpangan baku, tingkat presentil, dan sebagainya. Distribusi frekuensi dibedakan menjadi dua macam yaitu dengan penjelasan sebagai berikut.

#### 1. Distribusi Tunggal

Dalam distribusi tunggal, tiap skor ditulis dan kemudian dihitung atau ditali (Inggris: *tally*) sendiri-sendiri. Cara ini lebih menguntungkan jika jarak sebaran skor yang ada relatif kecil (maksimal 15). Jarak sebaran skor (*range*) dihitung dengan cara: skor tertinggi dikurangi skor terendah ditambah satu. Misalnya: skor tertinggi 60 dan terendah 25, jarak sebaran =  $(60-25) + 1 = 36$ .

Sebagai contoh distribusi tunggal, berikut disajikan skor hasil bahasa dari 40 orang siswa dengan kemungkinan skor tertinggi 75. Skor yang dimaksud sebagai berikut:

55 60 57 60 55 53 62 60 55 58  
54 62 52 60 54 60 61 61 55 64  
58 58 60 63 62 59 59 61 56 54  
53 53 55 56 60 57 58 62 63 59

Dengan membandingkan skor siswa yang belum disusun dengan yang telah dibuat menjadi distribusi frekuensi di atas, kita

dapat melihat perbedaannya. Data yang pertama belum dapat memberikan gambaran secara jelas, misalnya tentang skor tertinggi, terendah, jumlah frekuensi masing-masing skor, skor yang tertinggi frekuensinya atau sebaliknya yang terendah. Padahal, hal-hal tersebut akan dengan mudah ditemukan jika data-data itu telah dibuat menjadi distribusi frekuensi seperti di atas. Hal itu akan menjadi lebih sulit jika jumlah siswa relatif besar, apalagi jika skor-skoranya terdiri dari angka yang tidak bulat.

## 2. Distribusi Bergolong

Distribusi bergolong dibuat berdasarkan data-data yang telah dikelompokkan ke dalam kelas-kelas interval tertentu. Tiap kelas mempunyai angka batas kelas bawah dan batas kelas atas yang keduanya berjarak interval tertentu. Oleh karena itu, sebelum membuat distribusi frekuensi bergolong, terlebih dahulu kita perlu menentukan besarnya kelas dan interval.

**Penentuan Besarnya Kelas.** Tidak ada ketentuan yang harus diikuti secara mutlak untuk menentukan jumlah kelas. Sebagai bahan pertimbangan perlu dikemukakan bahwa jumlah kelas sebaiknya tidak terlalu kecil, atau sebaliknya terlalu besar. Kelas yang terlalu kecil akan mengaburkan keterangan-keterangan yang diperlukan. Sebaliknya, jumlah kelas yang terlampaui besar, akan tidak praktis dan tidak efisien. Sebagai ancar-ancar di sini dikemukakan bahwa jumlah kelas berkisar antara 7 sampai dengan 15.

**Penentuan Besarnya Interval.** Meskipun tidak ada keharusan mutlak yang harus diikuti, sebaiknya interval merupakan bilangan ganjil agar kita dengan mudah dapat menentukan titik tengah skor-skor pada tiap kelas. Setelah kita menentukan jumlah kelas, besarnya interval dapat ditentukan berdasarkan rumus berikut.

$$i(\text{nterval}) = \frac{\text{jarak sebaran (range)}}{k(\text{elas})}$$

Misalnya, data hasil tes menunjukkan skor tertinggi 69 dan skor terendah 15, sedang kita telah menentukan besarnya kelas adalah 11 buah, besarnya interval dapat dihitung sebagai berikut.

$$\begin{aligned} i &= \frac{(69 - 15) + 1}{11} \\ &= 5 \end{aligned}$$

Secara lengkap hasil tes yang dimaksud adalah sebagai berikut. Catat juga bahwa jumlah siswa adalah 50 orang dengan kemungkinan nilai tertinggi 75. Perhatikan juga bahwa susunan skor yang tak teratur belum dapat memberikan informasi-informasi yang diperlukan sebelum disusun menjadi distribusi frekuensi.

48 56 44 60 34 35 45 64 53 37  
 58 63 28 52 69 34 38 44 49 60  
 23 25 18 54 46 15 22 38 33 57  
 24 40 44 38 47 44 42 34 30 32  
 26 36 42 58 65 40 29 58 40 28

Dengan jumlah kelas (11) dan besarnya interval (5) seperti yang telah ditentukan di atas, kemudian kita susun distribusi bergolong yang dimaksud.

Perlu ditambahkan bahwa besarnya interval yang umum digunakan adalah 1, 2, 3, 5, 10, dan 20 (Guilford, 1956, lewat Nurkancana, 1983: 144), dan Nurkancana sendiri kemudian menambahkan angka 15.

## B. Penghitungan Nilai Rata-Rata, Median, Modus, dan Simpangan Baku

Nilai rata-rata dan simpangan baku adalah dua hal yang banyak dimanfaatkan dalam menganalisis dan atau mengolah skor mentah menjadi nilai jadi. Median dan modus juga sering ingin diketahui. Untuk itu, berikut juga akan dijelaskan bagaimana cara menghitungnya.

1. Nilai rata-rata (*mean*) disimbolkan dengan tanda:  $\bar{X}$  (banyak juga yang mempergunakan M sebagai kependekan mean). Nilai rata-rata dapat dihitung dengan berbagai cara tergantung keadaan data: (a) data yang belum disusun, (b) data yang disusun dengan distribusi tunggal, dan (c) data yang disusun dengan distribusi bergolong.

### (a) Menghitung Mean dari Data yang Belum Disusun

Mean dihitung dengan menjumlah semua skor dibagi dengan jumlah subyek. Rumusnya adalah:

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_n}{N}$$

Sebagai contoh misalnya, kita menghitung mean skor hasil yang belum disusun di atas. Kita ambil contoh yang pertama.

$$\begin{aligned} \bar{X} &= \frac{55 + 60 + 57 + 60 + \dots + 59}{40} \\ &= \frac{2324}{40} \\ &= 58,1 \end{aligned}$$

(b) Menghitung Mean dari Data Distribusi Tunggal

Pada hakikatnya, penghitungan mean data yang telah disusun dalam distribusi tunggal sama dengan cara di atas, yaitu sama-sama menjumlah seluruh skor kemudian dibagi jumlah subyek. Rumus yang digunakan adalah:

$$\bar{X} = \frac{\sum X}{N}$$

Sebagai contoh kita ambil data dalam table 10.1 di atas. Dalam penghitungan table 103 berikut sekaligus akan dihitung  $\sum X^2$  dan  $\sum X^2$  sebagai persiapan menghitung simpangan baku.

(c) Menghitung Mean dari Data Distribusi Bergolong

Penghitungan ini agak berbeda dengan kedua cara sebelumnya. Dalam hal ini kita perlu lebih dulu menentukan adanya angka yang diperlukan sebagai mean, yaitu yang disebut mean duga ( $X_d$ ) atau mean tebakan. Rumus yang digunakan adalah:

$$\bar{X} = \bar{X}_d + i \left( \frac{\sum fd}{N} \right)$$

= mean duga

$i$  = interval

$d$  = derivation

**Cara menentukan mean duga.** Angka yang diperkirakan sebagai mean adalah titik tengah pada kelas interval yang tempatnya di tengah jumlah kelas.

**Cara menentukan derivation.**  $d$  (erivation) menunjukkan besarnya penyimpangan dari nilai rata-rata ( $X_d$ ). Oleh karena itu, kelas interval yang ditempati mean duga,  $d$  adalah 0. Kelas interval di atasnya berturut-turut adalah: +1, +2, +3 dan seterusnya, sedang kelas di bawahnya berturut-turut -1, -2, -3, dan seterusnya. Berikut contoh penghitungan konkret yang diambil dari table di atas.

Berdasarkan data-data di atas kita dapat menghitung mean dengan memasukkan ke dalam rumus berikut.

$$\bar{X} = \bar{X}_d + i \left( \frac{\sum fd}{N} \right)$$

$$\bar{X} = 42 + 5 \left( \frac{-3}{50} \right)$$

$$\begin{aligned} &= 42 + 5 (-0,06) \\ &= 42 + 0,3 \\ &= 41,7 \end{aligned}$$

Penghitungan nilai rata-rata dengan cara di atas sering menghasilkan angka yang berbeda disbanding dengan cara yang pertama yaitu dengan menjumlah semua skor dibagi jumlah subyek. Perbedaan angka itu memang tak terlalu besar. Namun, hal itu dapat dipandang sebagai kelemahan cara tersebut. Sebagai contoh misalnya, data-data di atas jika dihitung dengan teknik pertama akan menghasilkan mean sebagai berikut.

$$\bar{X} = \frac{48 + 56 + 44 + 60 + \dots + 28}{50}$$

$$= \frac{2.099}{50}$$

$$= 41,98$$

Jadi, ada perbedaan angka sebesar 0,28. Selisih perbedaan tersebut disebut kesalahan pengelompokan (*grouping error*), akibat pengelompokan data kasar ke dalam distribusi frekuensi.

## 2. Median

Median ( $M_d$ ) adalah angka sentral atau titik tengah dari sejumlah skor. Secara teoretis median membagi skor menjadi dua bagian yang sama, yaitu skor-skor yang berada di atas dan di bawah median. Itulah sebabnya median disebut juga sebagai rata-rata posisi. Misalnya kita memiliki 15 skor yang telah diurutkan dari yang tertinggi: 65, 64, 62, 60, 60, 60, 59, 58, 56, 55, 55, 54, 53, 52, 50, mediannya adalah skor ke-8, yaitu 58.

Median yang dicari berdasarkan cara tersebut dinamakan median sesungguhnya (*true median*) karena ditentukan berdasarkan posisi skor yang telah diurutkan. Dalam penafsiran

nilai rata-rata populasi melalui teknik sampel, mean lebih banyak dipergunakan daripada median.

### 3. Modus

Modus (*mode*, simbol:  $M_o$ ) adalah skor yang memiliki frekuensi paling tinggi. Mencari modus data skor yang belum dikelompokkan ke dalam distribusi bergolong tidak sukar. Kita tinggal menunjuk saja skor tertentu yang berfrekuensi tertinggi, tanpa melalui penghitungan.

Menghitung modus dari data yang telah dikelompokkan harus melalui penghitungan. Hasil perhitungan itu merupakan modus yang diperkirakan, bukan modus yang sebenarnya. Angka yang diperoleh biasanya tidak sama dengan modus yang sesungguhnya. Rumus yang digunakan untuk mencari modus adalah:

$M_o$  = modus yang dicari

$B$  = Batas kelas bawah dari kelas modus

$F_o$  = Frekuensi kelas modus

$F_1$  = Frekuensi kelas sebelum kelas modus

$F_{-1}$  = Frekuensi kelas sesudah kelas modus

$i$  = interval

Berikut contoh penghitungan modus berdasarkan data yang disajikan pada table 10.5 di atas. Pertama, kita tentukan kelas modus, yaitu kelas yang berfrekuensi paling tinggi yang dalam table adalah: 40-44 dengan frekuensi ( $F_o$ : 9). Dari kelas itu kita dapat mengetahui frekuensi kelas sebelum dan sesudahnya, serta batas kelas bawah dan atasnya. Data-data itu kemudian dimasukkan ke dalam rumus di atas:

### 4. Simpangan Baku

Simpangan baku (*standard deviation*, dilambangkan dengan huruf Yunani:  $\sigma$ ) adalah ukuran penyebaran (variabilitas) skor yang diperoleh para siswa yang didasarkan pada kuadrat penyimpangan tiap skor dari nilai rata-rata (Tuckman, 1975: 486). Skor yang diperoleh seorang siswa (disebut skor mentah) biasanya mempunyai penyimpangan atau perbedaan dengan nilai rata-rata yang dicapai seluruh siswa. Rumusnya adalah:

$$x = X - \bar{X}$$

$x$  = ( $x$  kecil) besarnya penyimpangan individual

$X$  = ( $X$  besar) skor mentah



$\bar{X}$  = ( $\bar{X}$  bar) nilai rata-rata

Misalnya, seorang siswa mendapat skor mentah 58, sedang nilai rata-rata kelompok adalah 55, maka besarnya penyimpangan individual siswa tersebut adalah 3 ( $58 - 55 = 3$ )

Jika penyimpangan itu positif, berarti skor siswa yang bersangkutan berada di atas nilai rata-rata kelompok, sebaliknya jika negative, berarti di bawah nilai rata-rata. Penyimpangan skor dari nilai rata-rata tersebut terjadi pada seluruh siswa. Besar kecilnya penyimpangan itu akan memberikan petunjuk secara jelas tentang penyebaran (variabilitas) skor pada siswa. Besar kecilnya penyimpangan itu biasanya dinyatakan dengan indeks penyebaran (*index of variability*), dan indeks inilah yang kemudian lebih dikenal dengan sebutan simpangan baku (disingkat: S), atau ada yang mengadaptasi dari bahasa aslinya: standar deviasi (disingkat: SD), adapula yang menyingkatnya SB.

Besar kecilnya (indeks) simpangan baku menunjukkan besar kecilnya jarak penyebaran (*range*) yang dapat dilihat dari skor terendah dan tertinggi. Hal itu berarti semakin besar simpangan baku, semakin besar pula jarak penyebaran skor, itu berarti skor bersifat heterogen. Sebaliknya, semakin kecil simpangan baku, semakin kecil pula jarak penyebaran, dan menunjukkan bahwa skor bersifat homogeny. Hal itu juga berarti, jika semua skor yang diperoleh siswa itu sama, maka indeks simpangan baku adalah nol.

Penghitungan simpangan baku dapat dilakukan berdasarkan (a) penghitungan penyimpangan tiap skor individual, (b) data yang disusun dalam distribusi tunggal, (c) data dalam distribusi bergolong, dan langsung dari kalkulator.

(a) Menghitung Simpangan Baku dari Penyimpangan Skor Individual

Rumus yang digunakan:

$$S = \sqrt{\frac{\sum x^2}{N}}$$

S = Simpangan baku yang dicari

x = Penyimpangan skor individual dari mean

N = Jumlah subyek

Untuk memperjelas, berikut disajikan tiga data sebagai contoh penghitungan.

Data pertama		Data kedua		Data ketiga	
Skor (X)	Mean (X)	Skor (X)	Mean (X)	Skor (X)	Mean (X)
60	60	50	52	40	44
60	60	54	55	45	48
60	60	55	56	50	54
60	60	58	60	55	60
	60		55		49,5

Data pertama menunjukkan bahwa ke-8 orang subyek semuanya mendapat skor 60, berarti meannya juga 60. Dengan demikian tidak terdapat penyimpangan tiap skor individual dari mean. Hal itu berarti tidak ada variabilitas skor, dan sebagai akibatnya, simpangan bakunya pun nol.

Hal di atas tidak berlaku untuk data kedua dan ketiga karena keduanya mempunyai penyimpangan individual dari mean. Untuk menghitung simpangan baku, dan untuk lebih memperjelas, data-data di atas kita sajikan lagi ke dalam table 10.6 berikut.

Data tersebut selanjutnya dimasukkan ke dalam rumus:

$$s = \sqrt{\frac{70}{8}}$$

$$= 2,958$$

Dengan teknik penghitungan yang sama, data ketiga di atas mempunyai simpangan baku sebesar: 6,164. Masalahnya kemudian, apa arti perbedaan besarnya simpangan baku tersebut.

Simpangan baku data kedua lebih kecil daripada data ketiga. Hal itu terbukti bahwa data kedua memang mempunyai jarak penyebaran lebih kecil daripada data ketiga. Jarak penyebaran data kedua adalah 11, dengan skor tertinggi 60 dan terendah 50, sedangkan jarak penyebaran data ketiga

adalah 21, dengan skor tertinggi 60 dan terendah 40. Hal itu juga menunjukkan bahwa data kedua lebih bersifat homogeny daripada data ketiga, atau, data ketiga lebih heterogen daripada yang kedua.

- (b) Menghitung Simpangan Baku dari Data Distribusi Tunggal  
Rumus yang digunakan:

$$\sqrt{\frac{\sum x^2}{N} - \left(\frac{\sum X}{N}\right)^2} \quad \text{atau} \quad \sqrt{\frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N}}$$

Penghitungan dengan rumus di atas dilakukan langsung berdasarkan data kasar. Oleh karena itu, skor penilaian yang belum disusun ke dalam bentuk distribusi pun dapat dihitung dengan mempergunakan rumus di atas. Disbanding dengan rumus pertama, rumus langsung dari angka kasar ini lebih menguntungkan karena kerja penghitungan lebih sedikit. Jadi, lebih bersifat menghemat tenaga walau angka-angka yang dihadapi lebih besar. Sebab, hal itu bukan masalah jika mempergunakan kalkulator.

Sebagai contoh penghitungan, kita ambi data-data seperti yang terdapat dalam table 10.3. Dari table itu, kita telah mempunyai data-data  $N = 40$ ,  $\sum X = 2.324$ , dan  $\sum X^2 = 135.446$ . Data-data tersebut selanjutnya dimasukkan ke dalam rumus sebagai berikut.

$$\begin{aligned} s &= \sqrt{\frac{135.446}{40} - \left(\frac{2.324}{40}\right)^2} \\ &= \sqrt{3.386,15 - 3.375,61} \\ &= \sqrt{10,54} \\ &= 3,246 \end{aligned}$$

- (c) Menghitung Simpangan Baku dari Data Distribusi Bergolong  
Rumus yang digunakan:

### C. Kurve Normal

Penyebaran skor hasil tes biasanya mempunyai kecenderungan-kecenderungan tertentu, yaitu sebagian besar skor akan berada di sekitar mean, di atas atau di bawahnya. Skor yang semakin besar penyimpangannya dari mean akan semakin kecil frekuensinya, baik

penyimpangan ke arah yang lebih rendah. Dengan kata lain, skor yang berada jauh di atas dan di bawah mean frekuensinya akan jauh lebih kecil dibanding yang berada di sekitar mean. Penyebaran skor yang demikian adalah penyebaran yang memenuhi cirri-ciri *distribusi normal* – sebuah konsep yang sangat penting untuk menganalisis dan menaksir skor secara lebih lanjut.

Skor yang berada di titik tengah kurve adalah mean. Penyimpangan skor dari mean baik ke kanan (atas mean) maupun ke kiri (bawah mean) bersifat simetris, sama besarnya. Kurve normal memang disusun berdasarkan distribusi teoretis dengan persamaan matematis, bukan fakta empiris. Akan tetapi pada umumnya skor yang diperoleh dari hasil tes cenderung mengikuti distribusi normal, walau kadar kenormalannya tidak benar-benar simetris seperti dalam kurve normal.

#### 1. Daerah Kurve Normal

Besarnya penyimpangan skor dari mean dinyatakan dalam satuan simpangan baku ( $S$ ). Simpangan ke kanan (di atas mean) dengan:  $+1S$ ,  $+2S$ ,  $+3S$ , sedang simpangan ke kiri (di bawah mean) dengan:  $-1S$ ,  $-2S$ ,  $-3S$  (tanda plus dan minus selalu diikutsertakan). Daerah dinyatakan dalam presentase (ada: 100) atau proporsi (ada: 1).

Sebagai contoh, kita ambil data yang kemudian disusun dalam table 10.2, dari data tersebut diketahui:  $N = 50$ ,  $X = 41,98$  dan  $S = 13,205$  (juga diketahui skor tertinggi 69, terendah 15). Berdasarkan asumsi kurve normal berate siswa yang memiliki skor dari 28,775 sampai 55,185 (dibulatkan: 29 – 55), yaitu  $\pm 13,205$  dari 41,98 adalah  $68,26\% \times 50 = 34,13$  (dibulatkan 34) orang; yang memiliki skor  $\pm 2 \times 13,205$  dari 41,98: 15,57 sampai 68,39 (16 – 68) adalah  $95,44\% \times 50 = 47,72$  (dibulatkan 48) orang). Dan sisanya yang 2 orang adalah yang mempunyai skor lebih dari  $\pm 2 \times 13,205$  dari mean, dan dalam data memang ada skor itu, yaitu 15 dan 69. Walaupun terdapat perbedaan dengan keadaan data yang sesungguhnya, hal itu tidak akan menyimpang jauh.

Perlu ditegaskan bahwa konsep dan asumsi distribusi normal sangat penting. Penghitungan-penghitungan statistik mendasarkan diri pada asumsi distribusi normal. Jika kita bermaksud menguji normalitas data skor yang diperoleh, teknik statistik juga menyediakan cara itu.

## 2. Kurve Juling

Jika hasil tes ditafsirkan berdasarkan skor yang diperoleh siswa dalam sebuah kelompok, pada umumnya skor tersebut cenderung mengikuti asumsi distribusi normal. Pengolahan skor yang demikian adalah pengolahan skor yang berdasar norma kelompok (*norm reference test*). Sebaliknya, jika skor hasil tes ditafsirkan dengan criteria (patokan, standar) yang telah ditetapkan sebelumnya – pengolahan skor yang berdasarkan criteria tertentu (*criterion referenced test*) – skor hasil tes belum tentu mengikuti asumsi distribusi normal.

Jika terjadi penyebaran skor tidak berdistribusi secara normal, maka ada dua kemungkinan penyebaran skor hasil tes tersebut: skor tes cenderung menyebar ke kiri, atau sebaliknya, menyebar ke kanan. Jika sebagian besar skor (cenderung) menyebar ke kiri, berarti sebagian besar siswa memperoleh skor rendah (dalam kurve normal penyimpangan ke kiri di bawah mean). Keadaan skor yang menyebar ke kiri, jika digambarkan dalam bentuk kurve, akan berwujud kurve juling, juling ke kanan atau juling negatif.

Penyebaran skor yang cenderung menyebar ke kiri menunjukkan bahwa rata-rata siswa belum menguasai bahan pelajaran yang diteskan. Butir-butir soal yang diteskan masih terlampaui sulit bagi siswa yang bersangkutan.

Apabila skor cenderung menyebar ke kanan, berarti sebagian besar siswa memperoleh skor tinggi jika skor hasil tes mereka ditafsirkan dengan criteria yang telah ditetapkan. Penyebaran skor yang demikian, jika digambarkan dalam bentuk kurve, akan berwujud kurve juling, juling kiri atau juling positif.

Gambar kurve juling kiri menunjukkan bahwa sebagian besar siswa mampu mengerjakan butir-butir soal dengan betul. Butir-butir soal tersebut dirasakan mudah. Atau, hal itu dapat juga diartikan bahwa rata-rata siswa menguasai bahan pelajaran yang diteskan. Keadaan skor yang cenderung menyebar ke kanan atau ke kiri dapat ditemui jika penafsiran mempergunakan standar tertentu, tetapi hal itu tidak berlaku jika penafsiran mempergunakan standar relatif (norma kelompok).

#### D. TUGAS

Seorang guru bahasa memperoleh data skor hasil tes yang diberikan kepada 44 orang siswanya sebagai berikut.

65 55 58 47 45 54 60 50 52 45 42  
46 50 63 54 68 30 34 38 52 66 40  
33 37 26 46 36 28 54 45 50 43 44  
42 39 35 46 58 65 44 43 50 36 62

1. Buatlah distribusi frekuensi data skor di atas ke dalam 12 kelas. Jika jarak penyebaran besar, distribusi yang dibuat adalah distribusi frekuensi bergolong.
2. Hitunglah mean, median, modus, dan simpangan baku data skor di atas. Penghitungan mean dan simpangan baku dilakukan dua kali berdasarkan data kasar dan data yang telah disusun ke dalam bentuk distribusi frekuensi.
3. Gambarlah sebuah kurve normal berdasarkan mean dan simpangan baku yang diperoleh melalui penghitungan sebelumnya. Tentukan berapa orang yang memiliki skor di atas  $+1S$ ,  $+2S$ , mean ke atas,  $-1S$  sampai  $+1S$ ,  $-1S$ ,  $-2S$ , dan samakah jumlah siswa yang mendapat skor  $-1S$  ke bawah dengan  $+1S$  ke atas.

#### E. REFERENSI

Nurgiyantoro, Burhan.2001. *Penilaian dalam Pengajaran Bahasa dan Sastra*. Yogyakarta: PT. BPFE  
<http://supriyanto1985.wordpress.com>

## BAB 10 LAPORAN HASIL BELAJAR SISWA

### A. PENDAHULUAN

Nilai akhir pada umumnya diberikan dalam bentuk laporan (raport) atau tanda tamat belajar. Nilai akhir tersebut bukan semata-mata terdiri dari nilai sumatif saja, namun juga nilai formatif, tugas-tugas, hasil pengamatan guru, bahkan kehadiran. Sebelum suatu semester atau term dimulai seorang guru sebaiknya menjelaskan penentuan nilai akhir (terdiri dari komponen apa saja, bagaimana bobotnya, dll). Murid berhak mengetahui cara guru mengevaluasi hasil belajar siswa. Berdasarkan pada Kurikulum Tingkat Satuan Pendidikan (KTSP), banyak komponen yang dibutuhkan untuk mengisi laporan kemajuan siswa tersebut. Raport disebut sebagai Laporan Hasil Belajar Siswa (LHBS). LHBS mempunyai keunikan karena mencantumkan deskripsi ketercapaian masing-masing kompetensi dasar (KD) dari setiap mata pelajaran.

### B. KOMPONEN LHBS

#### 1. Kriteria Ketuntasan Minimal (KKM)

KKM ditetapkan pada setiap awal tahun pelajaran. Besarnya nilai KKM ditentukan oleh forum Musyawarah Guru Mata Pelajaran (MGMP). Nilai KKM ditetapkan mulai rentang 0 sampai dengan 100, dalam bilangan bulat. Nilai ketuntasan belajar maksimal adalah 100. Pihak sekolah dapat menetapkan nilai KKM di bawah nilai ketuntasan belajar maksimal. Nilai KKM harus dicantumkan dalam LHBS.

#### Penetapan Nilai KKM

Kompetensi Dasar (KD) dan Indikator	Kriteria Penetapan Ketuntasan			Nilai KKM
	Kompleksitas (Kesulitan & Kerumitan)	Daya Dukung	<i>Intake</i> Siswa	

Penulisan KKM Pada LHBS

Lembar Hasil Belajar Siswa  
Sekolah Menengah Atas

Nama Siswa: \_\_\_\_\_ Nama \_\_\_\_\_  
 Sekolah: \_\_\_\_\_  
 No. Induk: \_\_\_\_\_  
 Kelas/ Semester: \_\_\_\_\_ Tahun \_\_\_\_\_  
 Pelajaran: \_\_\_\_\_

No	Mata Pelajaran	KKM	Nilai Hasil Belajar				Sikap/ Afektif
			Pengetahuan dan Pemahaman Konsep		Praktik		Predikat
			Angka	Huruf	Angka	Huruf	
1.	Pendidikan Agama	75	80	Delapan puluh	-	-	B
2.	Bahasa Indonesia	70	70	Tujuh puluh	75	Tujuh puluh lima	B
3.	Bahasa Inggris	65	70	Tujuh puluh	65	Enam puluh lima	B

Mengetahui tingkat ketercapaian KKM yang telah ditetapkan ini sangat penting sebagai evaluasi dan sebagai dasar untuk menetapkan KKM tahun berikutnya.

2. Nilai Pengetahuan/ Kognitif

Nilai pengetahuan dan pemahaman konsep (kognitif) yang muncul dalam LHBS itu diolah dengan mencari rerata antara KD1, KD2, dan Ujian Akhir Sekolah (UAS). Contoh dapat dilihat berikut ini.

Nilai Pengetahuan			
KD <sub>1</sub>	KD <sub>2</sub>	UAS	Rata-Rata
82	70	67	<b>73</b>

3. Deskripsi

Terdapat kolom Keterangan Ketercapaian Kompetensi (DESKRIPSI). Misal” Telah mencapai KKM.....dan.... (hal ini harus mnegacu pada KKM yang telah ditentukan)

4. Nilai Psikomotor

Nilai ini sifatnya berjenjang dari angka 1 - 4

5. Nilai Praktik



## 6. Nilai Sikap

Nilai Sikap berdasarkan pada rentang nilai sikap sebagai berikut.

Amat Baik	= 85 – 100
Baik	= 69 - 84
Cukup	= 53 – 68
Kurang	= 20 – 52

Pengolahan nilai dapat dilakukan dengan menggunakan Office excel 2003/ 2007.

## C. LAPORAN

LHBS akan sangat berguna bagi:

### 1. Siswa yang bersangkutan

LHBS akan membuat siswa memahami kekurangan dan kelebihan, sehingga dia dapat menentukan langkah dan strategi belajar dalam semester berikutnya.

### 2. Sekolah yang bersangkutan

Pihak sekolah juga membutuhkan LHBS untuk laporan, evaluasi, dan keperluan lainnya.

### 3. Orang tua/ Wali murid

Orang tua/ wali murid sangat membutuhkan LHBS. Orang tua juga bertanggung jawab atas kemajuan belajar anaknya.

### 4. Stakeholders

Masyarakat, khususnya para pemakai lulusan sekolah termasuk kelompok yang turut serta membutuhkan informasi LHBS. Bahkan saat ini untuk mendaftarkan ke Universitas atau untuk mengikuti seleksi beasiswa, beberapa universitas dan atau lembaga memasukkan nilai LHBS sebagai salah satu criteria pertimbangan.

## D. TUGAS

Tugas Kelompok. Kelompok terdiri atas 2-3 orang. Kumpulkan contoh LHBS mulai dari Tingkat Pendidikan Dasar, Menengah, dan Atas. Analisis LHBS tersebut, kemudian bandingkan dengan contoh Raport (Laporan hasil Belajar) yang pernah dipakai tahun-tahun sebelumnya. Bagaimana pendapat Anda?

## E. REFERENSI

Nurgiyantoro, Burhan.2001. *Penilaian dalam Pengajaran Bahasa dan Sastra*. Yogyakarta: PT. BPFE

<http://supriyanto1985.wordpress.com>

<http://akhmadsudrajat.wordpress.com>

