

# **EVALUASI PEMBELAJARAN**



**Penulis**  
**Prof. Dr. Ida Ayu Gde Yadnyawati, M.Pd**

**UNHI PRESS**  
**2019**

# **EVALUASI PEMBELAJARAN**

DISUSUN OLEH

**Prof. Dr. Ida Ayu Gde Yadnyawati, M.Pd**

EDITOR

**Prof. Dr. I Ketut Suda, M.Si**

**UNHI PRESS**  
*Publishing*

# EVALUASI PEMBELAJARAN

Penulis : Prof. Dr. Ida Ayu Gde Yadnyawati, M.Pd

ISBN : 978-623-91636-3-1

Editor : Prof. Dr. I Ketut Suda, M.Si

Penyunting : Ida Bagus Putu Eka Suadnyana,SH.H., M.Fil.H

Desain Sampul dan Tata Letak : Ni Luh Putu Trisdyani,S.Sn., M.Sn

Penerbit : UNHI Press

Redaksi :

Jl. Sangalangit,  
Tembau, Penatih,  
Denpasar -Bali Telp.  
(0361) 464700/464800  
Email :  
unhipress@unhi.ac.id

Distributor Tunggal :  
UNHI Press  
Jl. Sangalangit, Tembau Penatih, Denpasar-Bali  
Telp. (0361) 464700/464800  
Email : unhipress@unhi.ac.id

Cetakan pertama, Oktober 2019

Hak cipta dilindungi undang-undang

Dilarang memperbanyak karya tulis ini dalam bentuk dan dengan cara apapun tanpa ijin tertulis dari penerbit.

## KATA PENGANTAR

Dengan memanjatkan puji syukur kehadapan Ida Sang Hyang Widhi Wasa (Tuhan Yang Maha Esa), atas segala rahmatnya pada penulis, akhirnya penulis dapat menyelesaikan sebuah buku yang berjudul: Evaluasi Pembelajaran. Dengan diterbitkannya buku ini, maka akan dapat menjadi refrensi bagi para guru, calon guru, pengelola diklat, mahasiswa, dosen dalam melaksanakan pembelajaran. Buku Evaluasi Pembelajaran ini, akan memberikan tuntunan bagaimana melaksanakan evaluasi dalam pembelajaran dan sekaligus juga agar lebih mudah menyusun alat evaluasi yang valid dan reliable.

Penulis menyadari bahwa buku ini dapat diselesaikan berkat dukungan dan bantuan dari berbagai pihak, oleh karena itu penulis berterima kasih kepada semua pihak yang secara langsung maupun tidak langsung memberikan kontribusi dalam penyelesaian buku ini. Secara khusus pada kesempatan ini penulis mengucapkan terimakasih kepada suami tercinta dr Ida Bgs Putra Paramartha, M.PH dan anak-anak tersayang, dr Ida Bgs Gde Suwibawa Putra.M.Biomed, S.pA, Ida Ayu Gde Suwiprabayanti Putra.S.Kom. M.Kom, Ida Bgs Gde Suwitrigna Putra ST. Menantu Ida Ayu Sri Wedari.SST serta cucu-cucu Ida Bgs Gde Mahabhakta Prabawa, Ida Ayu Gde Sundari Surya Patni. Yang dengan setia dan penuh kesabaran memberikan motivasi untuk menyelesaikan buku ini. Harapan penulis semoga hasil karya ini dapat bermanfaat bagi semua pihak yang berkecimpung dalam bidang pendidikan terutama dalam melaksanakan pembelajaran.

Denpasar, oktober 2019  
Penulis

Prof. Dr. Ida Ayu Gde Yadnyawati.M.Pd

# DAFTAR ISI

	Halaman
<b>KATA PENGANTAR</b> .....	i
<b>DAFTAR ISI</b> .....	ii
<b>BAB I. PENGERTIAN DAN PERANAN EVALUASI PEMBELAJARAN</b>	<b>1</b>
<b>.Pengertian Tes, Pengukuran, Penilaian dan Evaluasi</b>	
A. Pendahuluan .....	1
B. Pengertian Tes, Pengukuran, dan Penilaian.....	2
C. Peranan Pengukuran dan Penilaian dalam Pengajaran .....	6
<b>BAB II. PRINSIP DAN ALAT EVALUASI</b>	<b>11</b>
<b>Prinsip dan Prosedur Penilaian</b>	<b>11</b>
A. Pendahuluan .....	11
B. Hasil Belajar Sebagai Objek Penilaian .....	16
C. Tes sebagai Alat Penilaian Hasil Belajar .....	18
<b>Syarat Tes</b>	<b>19</b>
A. Pendahuluan .....	19
<b>BAB III. PENGADMINISTRASIAN TES HASIL BELAJAR</b>	<b>35</b>
<b>Langkah Penyusunan Tes</b>	<b>35</b>
A.Langkah-langkah Penyiapan Perangkat Tes .....	35
B.Format-Format Perencanaan .....	39
C.Petunjuk Tes .....	42
D.Waktu Tes .....	46
<b>Penyajian Tes</b>	
A. Penyajian Tes .....	48
B. Pemeriksaan Jawaban Tes .....	55

<b>BAB IV. PENGOLAHAN DATA HASIL PENILAIAN</b>	63
<b>Data Hasil Penilaian</b>	63
A. Pendahuluan .....	63
B. Batas Kelulusan .....	64
C. Kecenderungan Memusat dan Keragaman .....	67
<b>Tranformasi Data Penilaian</b>	79
A. Skor baku .....	79
B. Konversi Nilai .....	83
C. Pengolahan Data Hasil Non Test .....	94
<b>BAB V. MENGANALISIS TES</b>	103
<b>Menganalisis Tes</b>	103
A. Pendahuluan .....	103
B. Analisis Tingkat Kesukaran .....	104
C. Analisis Daya Pembeda .....	114
<b>Validitas dan Reliabilitas Tes</b>	
A. Analisis Validitas .....	118
B. Analisis Reliabilitas .....	124
Rangkuman .....	129
<b>BAB VI. PELAPORAN DAN PEMANFAATAN HASIL PENILAIAN</b>	130
<b>Pelaporan Hasil Penilaian</b>	130
A. Pendahuluan .....	130
B. Pelaporan Data Hasil penilaian .....	134
C. Pemanfaatan Data Hasil Penilaian .....	140
D. Manfaat Data Hasil Penilaian PBM .....	147
E. Manfaat Hasil Penilaian Bagi Penelitian Pendidikan .....	149
Rangkuman .....	150
Daftar Pustaka .....	152

# **BAB I**

## **PENGERTIAN DAN PERANAN EVALUASI PEMBELAJARAN**

### **A. Pendahuluan**

Dalam proses pendidikan tes dan pengukuran merupakan faktor sangat perlu diperhatikan karena hasil evaluasi amat diperlukan untuk menentukan berbagai macam tujuan dalam pengambilan keputusan antara lain seleksi, penempatan, prediksi, pengembangan kurikulum, perbaikan proses belajar-mengajar, dan pertanggungjawab pelaksanaan program pendidikan. Berkaitan dengan bidang pendidikan, evaluasi secara khusus bertujuan untuk mengetahui sejauhmana siswa telah menguasai tujuan-tujuan belajar yang telah ditetapkan sebelumnya dan mendiagnosis kesulitan belajar siswa (Gronlund, 1985). Untuk dapat membuat keputusan dengan tepat maka dalam evaluasi dibutuhkan informasi tentang tujuan-tujuan belajar siswa yang telah dicapai dengan akurat, relevan, dan komprehensif. Agar informasi yang diperoleh betul-betul merupakan gambaran kemampuan siswa yang sebenarnya maka diperlukan instrumen pengukuran dan prosedur pelaksanaan pengukuran yang dapat

memperoleh hasil yang berpedoman dengan objektivitas tinggi karena seringkali kita temukan pengukuran dan pengambilan keputusan mengandung subjektivitas disebabkan proses evaluasi merupakan kegiatan yang terdiri dari kegiatan yang kompleks.

Dari uraian di atas seorang guru perlu menguasai pengembangan tes, penggunaan tes, prinsip-prinsip dan teknik pengukuran, penilaian hasil belajar, memberikan nilai terhadap informasi yang diperoleh untuk selanjutnya memberikan keputusan yang obyektif. Kegiatan penilain ini hendaknya dilakukan dengan berpedoman pada prinsip bahwa penilaian yang diberikan mengenai hasil belajar siswa dimaksudkan sebagai umpan balik yang bersifat korektif baik bagi siswa maupun bagi guru, dengan hasil yang peroleh dapat meningkatkan minat dan motivasi siswa untuk memperoleh hasil belajar yang lebih baik lagi.

## **B. Pengertian Tes, Pengukuran, dan Penilaian**

Untuk menghindari kesalahan penafsiran terlebih dahulu perlu diuraikan pengertian tes (test), pengukuran (measurement), dan penilaian atau evaluasi (evaluation). Tes adalah merupakan suatu prosedur sistematis yang dipakai untuk mengukur tingkah laku atau karakteristik seseorang (Popham, 1980).

Selanjutnya Gronlund (1985) mendefinisikan tes sebagai suatu instrumen atau prosedur sistematis untuk mengukur suatu sampel dari tingkah laku. Lebih lanjut Cronbach (1984) mengemukakan bahwa: *a test is a systematic prosedur for observing behavior and describing it with the aid of numerical scales or fixed categories*. Secara singkat dapat didefinisikan bahwa tes merupakan proses pengamatan yang sistematis untuk mengetahui tingkah laku atau kemampuan siswa dan meng gambarkannya dengan skala atau kategori-kategori yang pasti. Dari definisi terakhir ini ada tiga unsur pokok yang tercakup dalam tes yaitu: prosedur sistematis, tingkah laku, dan skala atau kategori.

*Prosedur sistematis* berarti memberikan pemahaman bahwa tes direncanakan dan disusun sesuai dengan aturan-aturan penyusunan tes yakni memperhatikan aspek tingkah laku, materi pelajaran, dan tujuan pengukuran. Selain itu, penulisan dan pembuatan tes harus dapat disesuaikan dengan bentuk tes yang diharapkan sehingga dapat diinterpretasikan dengan mudah. Berkaitan dengan penggunaannya, tes dilaksanakan dengan disiplin yang tinggi agar dapat mengukur apa yang seharusnya diukur.

*Tingkah laku (behavior)* memberikan implikasi bahwa tes bertujuan untuk mengukur respon-respon

siswa yang diungkap melalui butir-butir tes. Dengan demikian tes bukan mengukur siswa secara langsung tetapi mengukur karakteristik-karakteristiknya melalui respon-responnya terhadap butir tes. Selain itu tingkah laku yang diamati bukanlah keseluruhan tingkah laku melainkan sebagian atau cuplikan dari keseluruhan butir. Hal ini berarti tingkah laku yang diobservasi merupakan sampel dari kemampuan seseorang secara keseluruhan, dengan perkataan lain dapat dinyatakan soal yang jumlahnya terbatas itu merupakan sampel dari soal yang tak terhingga jumlahnya.

*Skala angka atau kategori* berarti hasil tes dideskripsikan dengan bantuan skala pengukuran. Pada umumnya skala angka diberikan untuk melukiskan taraf atau bobot kemampuan atau tingkat ketercapaian hasil belajar siswa misalnya 0 -10 atau 0 – 100, sedangkan kategori yang diberikan lulus-gagal, baik–buruk, baik-cukup–kurang–buruk, banyak-cukup-kurang-rendah, dan lain-lain.

Menurut Gronlund (1985) evaluasi merupakan proses mendapatkan tingkat deskripsi angka bagi individu dengan karakteristik tertentu. Sedangkan Brown (1976) berpendapat pengukuran adalah deskripsi tingkah laku atau karakteristik seseorang yang bersifat kuantitatif dan diperoleh berdasarkan prosedur-prosedur

tertentu dengan bantuan alat ukur atau instrumen khusus yang dirancang untuk tujuan pengukuran. Hasil pengukuran dijabarkan dengan bantuan skala pengukuran. Dalam pemberian angka ada empat skala pengukuran yaitu nominal, ordinal, interval, dan ratio.

Evaluasi atau penilaian adalah proses sistematis mengumpulkan, menganalisis, dan menginterpretasi informasi dalam menentukan sejauhmana siswa telah mencapai tujuan pengajaran. Hasil penilain ini digunakan untuk mengambil keputusan berdasarkan pertimbangan yang dilakukan. Penilaian tentu saja didasarkan pada data hasil pengukuran (Gronlund, 1985). Dalam penilaian hasil yang diperoleh dapat bersifat kuantitatif dan kualitatif.

Dari uraian di atas tampak bahwa penilaian hasil belajar siswa cukup kompleks dan mungkin saja subyektif, hal ini disebabkan pengukuran berkenaan dengan aspek psikologis dengan data kuantitatif dan kualitatif. Untuk dapat memperoleh hasil belajar siswa secara obyektif diperlukan pemahaman guru yang lebih baik dalam hal penilaian dengan memanfaatkan bermacam-macam pengukuran.

### **C. Peranan Pengukuran dan Penilaian dalam Pengajaran**

Dalam kegiatan pengajaran *pengukuran* dan *penilaian* merupakan kegiatan yang sangat penting. Kedua kegiatan tersebut merupakan salah satu dari empat tugas pokok seorang pengajar. Keempat tugas pokok tersebut adalah *merencanakan, melaksanakan* dan *menilai keberhasilan pengajaran*, serta *memberikan bimbingan*.

Dalam praktek pengajaran, keempat kegiatan pokok ini merupakan sebuah kesatuan yang padu, yang tidak dapat dipisah-pisahkan. Dalam melaksanakan tugas mengajarnya, seorang pengajar berupaya untuk menciptakan situasi belajar yang memungkinkan siswa dapat belajar, memotivasi siswa, menyajikan bahan ajar, serta menggunakan metode dan media yang telah disiapkan. Selain itu ia mengolah dan menafsirkan hasil belajar siswa, serta mengambil keputusan untuk kepentingan peningkatan efektivitas pengajaran yang akan datang. Guna mencapai tujuan pendidikan yang optimal, guru juga memberikan bimbingan kepada siswa dengan berupaya untuk memahami kesulitan belajar yang dialami siswa beserta latar belakangnya dan sekaligus memberikan bantuan untuk mengatasinya sebatas kemampuan dan kewenangannya.

Terhadap seluruh komponen kegiatan belajar mengajar, pengukuran dan penilaian memberikan sumbangan yang sangat berarti. *Pengukuran dan penilaian* berfungsi sebagai pemantau kinerja komponen-komponen tersebut dalam mencapai tujuan akhir proses belajar mengajar. Informasi yang diberikan oleh hasil analisis terhadap hasil pengukuran dan penilaian sangat diperlukan bagi pembuatan kebijakan-kebijakan yang harus dilakukan seorang guru bagi peningkatan mutu kegiatan belajar mengajar di kelasnya.

Untuk membantu anda memahami penjelasan di atas, perhatikan ilustrasi di bawah ini.

*Dalam kehidupan profesional seorang guru SD, serta pengukuran dan penilaian akan tercermin dalam langkah-langkah utama yang disebutkan diatas (membuat persiapan, melaksanakan kegiatan belajar mengajar, dan mengevaluasi hasil belajar, serta memberikan bimbingan). Dalam pembuatan persiapan (program pengajaran) yang efektif, hasil-hasil pengukuran dan penilaian terhadap program pengajaran sebelumnya bisa dijadikan dasar bagi perbaikan aspek-aspek persiapan program pengajaran yang dikembangkan. Misalnya, jika pada penilaian yang dilakukan pada akhir program sebelumnya diperoleh informasi bahwa hasil belajar yang dicapai para siswa*

*kurang memuaskan, maka pada pengembangan persiapan program pengajaran selanjutnya, guru tersebut dapat mengambil langkah-langkah berikut:*

- a. Jika pencapaian yang kurang memuaskan tersebut terjadi pada sebagian besar bahan uji yang diberikan, maka program pengajaran tersebut harus diperbaiki dan diulangi;*
- b. Jika pencapaian yang kurang memuaskan tersebut hanya terjadi pada bagian-bagian tertentu dari keseluruhan bahan uji yang diberikan, maka guru memasukkan bagian-bagian tersebut ke dalam rencana program pengajaran selanjutnya.;*
- c. Jika pencapaian yang kurang memuaskan tersebut hanya terjadi pada sebagian besar siswa, maka program pengajaran harus diulangi;*
- d. Jika pencapaian yang kurang memuaskan tersebut hanya terjadi pada sebagian kecil siswa, maka guru harus memberikan program remedial kepada siswa-siswa tersebut.*

Begitu pun juga dengan pelaksanaan kegiatan belajar mengajar, hasil – hasil pengukuran dan penilaian hasil belajar siswa akan mempengaruhi langkah-langkah selanjutnya yang harus diambil seorang guru dalam mengembangkan kegiatan belajar mengajar. Apabila

seorang guru gagal menyadari dan menggunakan hasil pengukuran dan penilaian bagi pelaksanaan kegiatan belajar mengajar maka kegiatan-kegiatan belajar mengajar yang tidak efektif akan tetap berlangsung dan hasil program pengajaran tersebut akan tetap kurang efektif. Begitu pun, jika guru gagal menyadari kekurang berhasilan sejumlah siswa dalam menyelesaikan tugas-tugas belajarnya pada sebuah program pengajaran, maka anak tersebut akan mendapatkan kesulitan yang sinambung dan akan sangat fatal bagi kehidupan belajarnya di sekolah dasar bahkan juga bagi keseluruhan sejarah persekolahannya.

Peranan penting lainnya adalah dalam memberikan bimbingan. Data yang diperoleh dari kegiatan pengukuran dan penilaian dapat memberikan informasi mengenai siswa-siswa yang telah berhasil menunaikan tugas-tugas belajarnya dengan baik dan mereka yang belum berhasil. Mereka yang telah berhasil dapat diberikan kegiatan-kegiatan dan bahan-bahan pengayaan, sedangkan mereka yang belum berhasil dapat diberikan kegiatan-kegiatan dan bahan-bahan remedial.

Dengan demikian jelaslah bahwa cakupan peran pengukuran dan penilaian sangatlah luas. Berkaitan dengan luasnya cakupan peran pengukuran dan

penilaian dalam pengajaran secara luas, anda dapat membaca daftar pertanyaan serta jenis pengukuran dan penilaian yang dapat menjawabnya yang dikemukakan Linn dan Grondlund (1995: 4-5), yang disajikan di bawah ini:

- a. *Sejauh mana rancana pengajaran yang dibuat guru realitas bagi kelompok siswa yang berada dalam tanggung jawabnya? (Tes potensi akademik, tes hasil belajar yang telah lalu).*
- b. *Bagaimanakah para siswa dapat dikelompokkan sehingga dapat belajar dengan efektif (Tes buatan guru, tes hasil yang lalu, pengamatan).*
- c. *Sejauh mana para siswa telah siap untuk pelajaran yang selanjutnya? (Pretest, tes hasil belajar yang lalu).*
- d. *Sejauh mana para siswa mencapai tujuan pengajaran ? (Tes buatan guru, proyek kelas, pengujian lisan, pengamatan).*
- e. *Sejauh mana siswa memperoleh kemajuan yang melampui batas ambang? (Tes buatan guru, tes prestasi umum, proyek kelas, pengamatan).*
- f. *Pada bagian apakah sebuah pengulangan bahan ajar akan bermanfaat? (Kuis berjangka, pengujian li*

# **BABII**

## **PRINSIP DAN ALAT EVALUASI**

### **Prinsip dan Prosedur Penilaian**

#### **A. Pendahuluan**

Penilaian pembelajaran merupakan dapat menjadi penentu kulaitas pendidikan, maka upaya merencanakan dan melaksanakan penilaian hendaknya memperhatikan beberapa prinsip dan prosedur penilaian. Prinsip-prinsip tersebut antara lain adalah:

- (a) Dalam menilai hasil belajar hendaknya dirancang sedemikian rupa sehingga jelas abilitas yang harus dinilai, materi penilaian, alat penilaian, dan interpretasi hasil penilaian. Sebagai patokan atau rambu-rambu dalam merancang penilaian hasil belajar adalah kurikulum yang berlaku dan buku pelajaran yang digunakannya. Dalam kurikulum hendaknya dipelajari tujuan-tujuan kurikuler dan tujuan instruksionalnya, pokok bahasan yang diberikan, ruang lingkup dan urutan penyajian, serta pedoman bagaimana pelaksanaannya.

- (b) Penilaian hasil belajar hendaknya menjadi bagian integral dari proses belajar-mengajar. Artinya, penilaian senantiasa dilaksanakan pada setiap saat proses belajar-mengajar sehingga pelaksanaannya berkesinambungan. Tiada proses belajar-mengajar tanpa penilaian, hendaknya dijadikan semboyan bagi setiap guru. Prinsip ini mengisyaratkan pentingnya penilaian formatif sehingga dapat bermanfaat baik bagi siswa maupun bagi guru.
- (c) Agar diperoleh hasil belajar yang objektif dalam pengertian menggambarkan prestasi dan kemampuan siswa sebagaimana adanya, penilaian harus menggunakan berbagai alat penilaian dan sifatnya komprehensif. Dengan sifat komprehensif dimaksudkan segi atau abilitas yang dinilai tidak hanya aspek kognitif, tetapi juga aspek afektif dan psikomotoris. Demikian pula dalam menilai aspek kognitif sebaiknya dicakup semua aspek, yakni pengetahuan, pemahaman, aplikasi, analisis, sintesis, dan evaluasi secara seimbang.
- (d) Penilaian hasil belajar hendaknya diikuti dengan tindak lanjutnya. Data hasil penilaian sangat bermanfaat bagi guru maupun bagi siswa. Oleh

karena itu, perlu dicatat secara teratur dalam catatan khusus mengenai kemajuan siswa. Demikian juga data hasil penilaian harus dapat ditafsirkan sehingga guru dapat memahami para siswajiya terutama prestasi dan kemampuan yang dimilikinya. Bahkan jika mungkin, guru dapat meramalkan prestasi siswa pada masa mendatang. Hasil penilaian juga hendaknya dijadikan bahan untuk menyempurnakaq program pengajaran, memperbaiki kelemahan-kelemahan pengajaran, dan memberikan bimbingan belajar kepada siswa yang memerlukannya. Lebih jauh lagi dapat dijadikan bahan untuk memperbaiki alat penilaian itu sendiri.

Ada beberapa langkah yang dapat dijadikan pegangan dalam melaksanakan proses penilaian hasil belajar, yakni:

- (a) Merumuskan atau mempertegas tujuan-tujuan pengajaran. Mengingat fungsi penilaian hasil belajar adalah mengukur tercapai-tidaknya tujuan pengajaran, maka perlu dilakukan upaya mempertegas tujuan pengajaran sehingga dapat memberikan arah terhadap penyusunan alat-alat penilaian.

- (b) Mengkaji kembali materi pengajaran berdasarkan kurikulum dan silabus mata pelajaran. Hal ini penting mengingat isi tes atau pertanyaan penilaian berkenaan dengan bahan pengajaran yang diberikan. Penguasaan materi pengajaran sesuai dengan tujuan-tujuan pengajaran merupakan isi dan sasaran penilaian hasil belajar.
- (c) Menyusun alat-alat penilaian, baik tes maupun nontes, yang cocok digunakan dalam menilai jenis-jenis tingkah laku yang tergambar dalam tujuan pengajaran. Dalam penyusunan alat penilaian hendaknya diperhatikan kaidah-kaidah penulisan soal
- (d) Menggunakan hasil-hasil penilaian sesuai dengan tujuan penilaian tersebut, yakni untuk kepentingan pendeskripsian kemampuan siswa, kepentingan perbaikan pengajaran, kepentingan bimbingan belajar, maupun kepentingan laporan pertanggungjawaban pendidikan.

Dalam kaitannya dengan penyusunan alat-alat penilaian ada beberapa langkah yang harus ditempuh, yakni:

- (a) Menelaah kurikulum dan buku pelajaran agar dapat ditentukan lingkup pertanyaan, terutama

materi pelajaran, baik luasnya maupun kedalamannya.

- (b) Merumuskan tujuan instruksional khusus sehingga jelas betul abilitas yang harus dinilainya. Tujuan instruksional khusus harus dirumuskan secara operasional, artinya bisa diukur dengan alat penilaian yang biasa digunakan.
- (c) Membuat kisi-kisi atau *blueprint* alat penilaian. Dalam kisi-kisi harus tampak abilitas yang diukur serta proporsinya, lingkup materi yang diujikan serta proporsinya, tingkat kesulitan soal dan proporsinya, jenis alat penilaian yang digunakan, jumlah soal atau pertanyaan, dan perkiraan waktu yang diperlukan untuk mengerjakan soal tersebut.
- (d) Menyusun atau menulis soal-soal berdasarkan kisi-kisi yang telah dibuat. Dalam menulis soal, perhatikan aturan-aturan yang berlaku.
- (e) Membuat dan menentukan kunci jawaban soal.

## **B. Hasil Belajar Sebagai Objek Penilaian**

Pertanyaan pokok sebelum melakukan penilaian ialah apa yang harus dinilai itu. Terhadap pertanyaan ini kita kembali kepada unsur-unsur yang terdapat dalam proses belajar-mengajar. Ada empat unsur utama proses belajar-mengajar, yakni tujuan-bahan-metode dan alat serta penilaian. Tujuan sebagai arah dari proses belajar-mengajar pada hakikatnya adalah rumusan tingkah laku yang diharapkan dapat dikuasai oleh siswa setelah menerima atau menempuh pengalaman belajarnya. Bahan adalah seperangkat pengetahuan ilmiah yang dijabarkan dari kurikulum untuk disampaikan atau dibahas dalam proses belajar-mengajar agar sampai kepada tujuan yang telah ditetapkan. Metode dan alat adalah cara atau teknik yang digunakan dalam mencapai tujuan. Sedangkan penilaian adalah upaya atau tindakan untuk mengetahui sejauh mana tujuan yang telah ditetapkan itu tercapai atau tidak. Dengan kata lain, penilaian berfungsi sebagai alat untuk mengetahui keberhasilan proses dan hasil belajar siswa.

Proses adalah kegiatan yang dilakukan oleh siswa dalam mencapai tujuan pengajaran, sedangkan hasil belajar adalah kemampuan-kemampuan yang dimiliki siswa setelah menerima pengalaman belajarnya. Horward Kingsley membagi tiga macam hasil belajar,

yakni (a) keterampilan dan kebiasaan, (b) pengetahuan dan pengertian, (c) sikap dan cita-cita. Masing-masing jenis hasil belajar dapat diisi dengan bahan yang telah ditetapkan dalam kurikulum. Sedangkan Gagne membagi lima kategori hasil belajar, yakni (a) informasi verbal, (b) keterampilan intelektual, (c) strategi kognitif, (d) sikap, dan (e) keterampilan motoris. Dalam sistem pendidikan nasional rumusan tujuan pendidikan, baik tujuan kurikuler maupun tujuan instruksional, menggunakan klasifikasi hasil belajar dari Benyamin Bloom yang secara garis besar membaginya menjadi tiga ranah yakni ranah kognitif, ranah afektif, dan ranah psikomotoris.

*Ranah kognitif* berkenaan dengan hasil belajar intelektual yang terdiri dari enam aspek, yakni pengetahuan atau ingatan, pemahaman, aplikasi, analisis, sintesis, dan evaluasi. Kedua aspek pertama disebut kognitif tingkat rendah dan keempat aspek berikutnya termasuk kognitif tingkat tinggi

*Ranah afektif* berkenaan dengan sikap yang terdiri dari lima aspek, yakni penerimaan, jawaban atau reaksi, penilaian, organisasi, dan internalisasi.

*Ranah psikomotoris* berkenaan dengan hasil belajar keterampilan dan kemampuan bertindak. Ada enam aspek ranah psikomotoris, yaitu (a) gerakan refleks, (b) keterampilan gerak dasar, (c) kemampuan

perseptual, (d) keharmonisan atau ketepatan, (e) gerakan keterampilan kompleks, dan (f) gerakan ekspresif dan interpretatif.

### **C. Tes Sebagai Alat Penilaian Hasil Belajar**

Dalam bab ini akan dijelaskan secara khusus alat-alat penilaian hasil belajar, yaitu tes, baik tes uraian (esai) maupun tes objektif. Tes sebagai alat penilaian adalah pertanyaan-pertanyaan yang diberikan kepada siswa untuk mendapat jawaban dari siswa dalam bentuk lisan (tes lisan), dalam bentuk tulisan (tes tulisan), atau dalam bentuk perbuatan (tes tindakan). Tes pada umumnya digunakan untuk menilai dan mengukur hasil belajar siswa, terutama hasil belajar kognitif berkenaan dengan penguasaan bahan pengajaran sesuai dengan tujuan pendidikan dan pengajaran. Sungguhpun demikian, dalam batas tertentu tes dapat pula digunakan untuk mengukur atau menilai hasil belajar bidang afektif dan psikomotoris.

Ada dua jenis tes yang akan dibahas dalam bab ini, yakni tes uraian atau tes esai dan tes objektif. Tes uraian terdiri dari uraian bebas uraian terbatas, dan uraian berstruktur. Sedangkan tes objektif terdiri dari beberapa bentuk, yakni bentuk pilihan benar-salah, pilihan berganda dengan berbagai variasinya,

menjodohkan, dan isian pendek atau melengkapi. Setiap jenis tes dijelaskan konsep- konsepnya sepunya, jenis atau bentuk-bentuknya, cara penyusunannya, kelebihan dan kekurangannya, pemeriksaan dan cara skoringnya, termasuk contoh-contohnya.

## **Syarat-Syarat Tes**

### **A. Pendahuluan**

Syarat-syarat suatu alat ukur yang memenuhi syarat baik sebagai alat ukur adalah sebagai berikut : (1) harus efisien, parsimory (hemat), (2) harus baku (standardize), (3) mempunyai norma, (4) objektif, (5) valid, dan (6) reliabel.

Penjelasan masing-masing syarat yang disebutkan di atas adalah sebagai berikut :

#### **1. Harus efisien, parsimory.**

Syarat ini menentukan bahwa dengan menggunakan waktu yang tidak lama maka dapat terekam data atau hasil yang optimal

2. Harus *standardize*

Syarat *standardize* ini merupakan kontrol bagi alat ukur, sehingga isi butir (content item), standar, dan prosedur menskor item berlaku umum bagi semua orang yang akan menggunakan alat ukur tersebut

3. Mempunyai norma

Alat ukur harus memiliki norma tertentu agar dapat mengukur objek dengan tepat, misalnya untuk kelompok umum tertentu atau ciri-ciri tertentu yang akan diukur sehingga dapat menggambarkan penilaian dengan lebih objektif.

4. Objektif.

Alat ukur harus benar-bener menggambarkan apa yang seharusnya diukur sesuai dengan content yang akan diukur, sehingga perlu dilakukan penelitian kualitatif terhadap item. Dengan demikian item-item alat ukur benar-benar merupakan yang dapat mengidentifikasi apa yang diukur.

5. Valid.

Persyaratan ini menuntut alat ukur harus mampu mengukur apa yang dianggap orang seharusnya diukur alat ukur tersebut.

6. Reliabel.

Suatu alat ukur yang baik sebagai alat ukur harus reliabel yaitu yang mengukur ukuran yang harus terukur, dengan demikian mengacu kepada sejauh mana alat ukur secara ajeg (konsisten ) mengukur apa saja yang diukunya.

Untuk mendapatkan instrumen yang baik, ada beberapa tahap sebagai berikut :

1. Menentukan kegunaan pengukuran
2. Menentukan tingkah laku yang mewakili konstruk apa yang diukur
3. Domain sampling
4. Spesifikasi item (item specification)
5. Item construction (item pooling)
6. Item review
7. Item try out
8. Item field-test
9. Analisa butir (item)

10. Validitas dan Reliabilitas

11. Membuat panduan (administrasi, scoring dan interpretasi)

12. Tabel norma.

Uraian masing-masing tahap yang disebutkan di atas adalah sebagai berikut:

1. Menentukan kegunaan pengukuran.

Sebelum mengkonstruksi suatu instrumen terlebih dahulu ditentukan kegunaan alat pengukuran berdasarkan penggunaan tes dan tujuan pengukuran. Berdasarkan penggunaan tes, pada umumnya penggunaan pengukuran dibedakan atas tiga hal yaitu: untuk seleksi, penempatan (*placement*), dan penggunaan diagnostik. Tes seleksi bertujuan untuk menentukan kemampuan atau kecenderungan objek yang diukur sesuai dengan kriteria tertentu melalui prinsip kompetitif dengan memunculkan variabel-variabel tertentu. Tes placement ditujukan untuk mengukur kompetensi minimum objek yang diukur.

Tes diagnostik bertujuan untuk menemukan kelemahan-kelemahan tertentu, dengan demikian

alat ukur harus mengandung item-item yang cukup mudah bagi rata-rata peserta.

2. Menentukan tingkah laku yang diukur.

Proses dari konstruk psikologi hingga menjadi item dilakukan berdasarkan pengalaman karena di sini tidak ada satu carapun diberikan dalam buku-buku. Sebelum sampai ke konstruk terlebih dahulu dilakukan konseptualisasi tingkah laku yaitu mengolah konsep tingkah laku atau operasionalisasi konsep. Setelah konsepnya jelas maka dicoba membuat konstruk, dalam membuat konstruk ini dituntut kejelasan tingkah laku berdasarkan teori dari gejala yang diukur. Kemudian menduga butir-butir (item) apa yang dapat mengukur tingkahlaku tersebut dan memperbaiki butir yang diduga kurang benar, subjektif, atau tidak valid. Untuk mengetahui apakah butir yang dibuat benar maka harus dilakukan:

- a. *Content analysis*, misalnya meneliti tingkah laku tertentu, kemudian menjadikan aspeknya dibuat menjadi butir (item).
- b. *Review riset*, yakni tingkah laku hasil penelitian untuk definisi konstruk butir. Misalnya meneliti

pilot dimasukkan pertanyaan apakah responden yang dites menjadi pilot memiliki jam tangan.

- c. *Critical incidents*, yaitu untuk menemukan sejumlah tingkah laku ekstrim yang dapat menggambarkan suatu kontinum gejala yang diukur. Misalnya menemukan tingkah laku kritikal yang dapat menentukan orang sukses dibandingkan dengan orang gagal.
- d. *Direct observation*, yaitu mencari tingkah laku dari observasi langsung. Direct observation ini masih merupakan face validity.
- e. *Expert judgement*, yaitu mengadakan konsultasi dengan para ahli baik melalui interview maupun kuesioner untuk menetapkan tingkah laku yang menjadi konstruk.
- f. *Instruction objectives*, ditujukan kepada orang yang ahli untuk mencari *specific content* dari tingkah laku. Instruction objectives ini penting dideskripsikan untuk mengukur sesuatu tujuan. Instruction objectives ini lebih terinci dan terprogram dari *expert judgement*.

### 3. Domain sampling.

Melalui domain sampling ini harus menemukan daerah sampling tertentu atau kelompok konstruk gejala tersebut. Hal ini akan mengarahkan tes pada sasaran yang akan diukur, sebagai contoh pada konstruk psikologis sering mengandung sifat-sifat yang banyak dan kompleks, misalnya inteligensi, kreativitas, moral dan lain-lain.

### 4. Item specification.

Pada tahap item specification ini dituntut bagaimana isi-isi item yang dapat diungkap (sumber item content), harus mendeskripsikan situasi problem, stimuli dan dapat berisi karakteristik respons yang benar atau karakteristik respons yang salah. Bila yang menyusun item pooling banyak orangnya maka item specification akan lebih baik, untuk itu perlu banyak pembuat item. Pada item specification ini perlu diperhatikan hal-hal berikut ini yaitu : arti skor dalam criterion referenced; domain tingkah laku harus

mewakili seluruh domain; bentuk-bentuk butir hendaknya mempunyai alternatif, dengan kata lain bahwa jumlah butir dibuat lebih dari yang dibutuhkan. Selain itu, perlu juga diperhatikan banyaknya proporsi butir bagi setiap item specification, materinya dan hubungannya dengan konsep lainnya. Sebagai contoh, sebelum membuat butir-butir soal terlebih dahulu dibuat kisi-kisi (blue print) sebagai jaminan bahwa apa yang akan diujikan memang mengukur tujuan-tujuan pengajaran yang ditentukan sebelumnya serta bobot yang diberikan untuk tiap-tiap tujuan.

##### 5. Item construction (*item pooling*)

Agar dapat mengukur konstruk yang sebenarnya maka dibuat "*pool of item*". Kegiatan-kegiatan pada tahap item construction ini adalah :

- a. memilih format item, apakah pormat pertanyaan terbuka atau tertutup. Untuk kebutuhan inventory ada tiga format yang terkenal, yaitu :

- format dikotomi
  - format Lickert
  - bipolar adjective checklist
- b. Membuktikan bahwa item feasible untuk sampel, misalnya membuat butir sesuai dengan jenjang pendidikan siswa.
  - c. Memilih dan melatih penulis item, yang dimaksudkan dengan penulis disini mungkin saja kita sendiri
  - d. Menulis item. Mula-mula yang dilakukan adalah membuat sesuai dengan content, kemudian skala teknis dan bahasanya.
  - e. Monitor kemajuan dan kualitas guna mendapatkan item yang benar-benar baik.

#### 6. Item review

Pada tahap ini alat ukur diperiksa, diteliti oleh peneliti sendiri dan oleh ahli, dan bila perlu dilakukan revisi. Adapun rancangan yang diteliti oleh ahli meliputi hal-hal sebagai berikut : ketepatan kata-kata, tata bahasa, ambiguitas, teknis, bias, reliabilitas dan relevansi. Hal-hal lain yang perlu diteliti adalah test specification, item specification, measurement, dan test construction.

7. Item try out.

Sebelum melakukan try out yang sebenarnya terlebih dahulu dilakukan pra try out pada sampel yang kecil, hal ini dimaksudkan untuk mendapatkan item yang mendekati kriteria alat ukur yang baik. Bila item kurang/tidak valid dan reliabel dapat diketahui dan diperbaiki sebelum lebih jauh melakukan try out yang sebenarnya. Pada waktu try out dilakukan observasi, wawancara dan membuat komentar tentang item yang berguna untuk mendapatkan item yang valid, reliabel dan objektif. Melalui try out ini ditemukan tingkat kesukaran dan variasi respons secara statistik.

8. Item field test.

Item field test ini dilakukan pada sampel besar, sesuai dengan karakteristik yang dituju.

9. Analisis butir (item)

Analisis butir bertujuan untuk menentukan apakah butir yang dibuat merupakan butir yang baik atau buruk, dengan demikian dapat diketahui apakah alat ukur tersebut baik atau buruk. Untuk menentukan apakah butir merupakan butir yang baik atau buruk maka diperlukan informasi tentang :

- a. Indeks kesukaran butir.
- b. Indeks validitas butir atau daya pembeda

Indeks validitas bertujuan untuk mengetahui daya pembeda item. Daya pembeda : kemampuan suatu butir (soal) membedakan siswa yang pandai dengan siswa yang bodoh/lemah.

$$\text{Daya Pembeda (D)} = \frac{R_u}{T_u} - \frac{R_L}{T_L}$$

*D = Daya pembeda*

*R<sub>u</sub> = Banyaknya kelompok atas yang menjawab soal dengan benar*

*R<sub>L</sub> = Banyaknya kel. bawah yang menjawab soal dengan benar.*

*T<sub>u</sub> = Banyak siswa kel. atas*

*T<sub>L</sub> = Banyak siswa kel. bawah*

Butir Soal	Atas 27% TCN = 100)	Bawah 27% L (N=100)	Indeks Kesukaran
1	100	90	0,10
2	87	30	0,57
3	95	31	0,64
4	85	25	0,60
5	82	13	0,69
6	52	8	0,44

- 1,00
0,00
1,00

Daya pembeda negatif
Daya pembeda rendah
Daya pembeda tinggi positif

c. Pengecoh yang baik : Minimal 2% ada memilih alternatif tersebut

- d. Khusus untuk bentuk item multiple choice diperlukan informasi tentang penyebaran jawaban dari subjek.
- e. Apakah item dalam test homogen atau heterogen.

## 10. Validitas

Pada tahap ini diuji validitas alat ukur. Seberapa jauh tes mengukur apa yang hendak diukur disebut validitas. Adapun fungsi validitas ini adalah menerangkan atribut, konstruk, aspek faktor, apa yang diukur dan seberapa jauh hal itu diukur. Alat-alat ukur dibidang pendidikan dan psikologi dirancang untuk menaksir konstruk seperti prestasi belajar, kecerdasan, kreativitas, bakat, sikap, motivasi dan sebagainya. Akan tetapi tidak ada alat yang langsung dapat mengukur konstruk tersebut seperti yang dipakai pada ilmu pengetahuan alam untuk mengukur ciri-ciri seperti panjang, besar dan berat. Para peneliti harus mengembangkan cara-cara tidak langsung untuk mengukur atribut-atribut yang kompleks tersebut.

Validitas suatu instrumen selalu tergantung kepada situasi dan tujuan khusus penggunaan

instrumen tersebut. Suatu tes yang valid untuk suatu situasi mungkin tidak valid untuk situasi yang lain. Pada dasarnya semua prosedur yang digunakan untuk menentukan validitas suatu tes selalu menyangkut hubungan atau hasil tes dengan fakta-fakta lain yang independent dan observable, mengenai ciri-ciri tingkah laku yang akan kita ukur. Syarat kriteria validitas suatu test adalah : independent atau bebas dari test yang diuji, observable/measurable; reliable yakni dengan meminta judgment dari orang yang ahli; relevant; dan bebas dari *bias*. Cara-cara khusus untuk mencari hubungan antara skor tes dengan kriteria dilakukan dengan bermacam-macam cara, yaitu : (a) content validity, (b) criterion-related validity, dan (c) construct validity.

#### 11. Reliabilitas.

Reliabilitas suatu alat ukur adalah derajat keterandalan alat tersebut dapat mengukur apa yang diukurnya. Sifat ini penting dalam segala jenis pengukuran. Suatu tes dapat dipercaya sejauh skor yang diperoleh seorang individu dalam tes itu hampir sama dalam pengukuran yang berulang-ulang.

Reliabilitas mengandung beberapa arti yaitu consistency, constancy, stability, dipendability, pridictability dan accuracy.

Reliabilitas suatu alat ukur dapat diketahui dengan menggunakan teknik atau metode :

- a.. metode test-retest
- b. metode alternate form
- c. metode single trial

Rumus-rumus lain dapat juga digunakan, antara lain

:

- a. rumus Sperman-Broon
- b. rumus Flanagan
- c. rumus split-half dengan tehnik rulon
- d. rumus Koeder-Richardson
- e. rumus analisis varians Hoyt
- f. rumus koefisien alpha.

## 12. Membuat panduan.

Pada tahap ini dibuat panduan penggunaan alat ukur yang mencakup administrasi item, prosedur menskor item, dan interpretasi item sehingga alat ukur lebih terinci dan lebih jelas. Dengan rincian ini akan memudahkan bagi orang yang menggunakan alat ukur dalam menginterpretasikan alat tersebut.

### 13. Tabel Norma.

Tabel norma ini juga berguna untuk memudahkan bagi orang-orang yang akan menggunakan alat ukur, sehingga penggunaan alat ukur tepat ditujukan bagi siswa atau anak sesuai dengan tingkat pendidikan, kelompok umur atau kematangan.

## **Rangkuman**

1. Beberapa prinsip dalam melakukan penilaian antara lain: merencanakan penilaian sesuai dengan tujuan penilaian, penilaian hasil belajar hendaknya menjadi bagian integral dari proses belajar-mengajar, penilaian hasil belajar hendaknya diikuti dengan tindak lanjutnya, dan hasil penilaian juga hendaknya dijadikan bahan untuk menyempurnakan program pengajaran.
2. Beberapa langkah yang dapat dijadikan pegangan dalam melaksanakan proses penilaian hasil belajar, yakni: (a) merumuskan atau mempertegas tujuan-tujuan pengajaran, (b) mengkaji kembali materi pengajaran berdasarkan kurikulum dan silabus mata pelajaran, (c) menyusun alat-alat penilaian, baik tes maupun nontes, yang cocok digunakan dalam menilai jenis-jenis tingkah laku yang tergambar dalam tujuan pengajaran, dan (d) menggunakan hasil-hasil penilaian sesuai dengan tujuan penilaian.

3. Beberapa langkah yang harus ditempuh dalam menyusun alat penilaian, yakni: (a) menelaah kurikulum dan buku pelajaran agar dapat ditentukan lingkup pertanyaan, terutama materi pelajaran, baik luasnya maupun kedalamannya, (b) merumuskan tujuan instruksional khusus, (c) membuat kisi-kisi atau *blueprint* alat penilaian, (d) menyusun atau menulis soal-soal berdasarkan kisi-kisi yang telah dibuat, dan (e) membuat dan menentukan kunci jawaban soal.
4. Hasil belajar sebagai objek penilaian dapat dibedakan ke dalam beberapa kategori, antara lain keterampilan dan kebiasaan, pengetahuan dan pengertian, sikap dan cita-cita. Kategori lain yang banyak digunakan menjadi tiga ranah, yaitu (a) kognitif, (b) afektif, dan (c) psikomotorik. Masing-masing ranah terdiri dari sejumlah aspek yang saling berkaitan.
5. Pada umumnya hasil belajar dinilai melalui tes, baik tes uraian maupun tes objektif. Pelaksanaan penilaian dapat secara lisan, tulisan, dan tindakan atau perbuatan
6. Beberapa syarat yang harus dipenuhi oleh suatu alat ukur yang baik adalah sebagai berikut : (1) harus efisien, parsimony (hemat), (2) harus baku (standardize), (3) mempunyai norma, (4) objektif, (5) valid, dan (6) reliabel.

# **BAB III**

## **PENGADMINISTRASIAN TES HASIL BELAJAR**

### **Langkah Penyusunan Tes**

Dalam melakukan suatu tes perlu dilakukan beberapa tahapan, tahapan pertama adalah penyiapan perangkat tes. Untuk melakukan penyiapan perangkat tes maka langkahnya adalah:

#### **A. Langkah-langkah penyiapan perangkat tes**

##### *1). Menetapkan tujuan tes*

Dalam menetapkan tujuan tes sangat berkaitan dengan maksud penggunaan tes. Tes prestasi belajar dapat dibuat untuk berbagai tujuan. Oleh sebab itu perlu ditetapkan terlebih dahulu penggunaan dari pada tes yang akan dikembangkan. Tujuan penggunaan tes akan memberikan corak dan bentuk terhadap penyusunan butir soal, misalnya bila tes itu dimaksudkan sebagai Ujian Akhir Nasional (UAN) maka butir soal harus disusun mulai dari yang muda sampai kepada yang sukar.

## 2). Analisis Kurikulum

Isi bahan pengajaran yang disajikan di kelas senantiasa mengikuti kurikulum yang berlaku. Pemahaman dan pendalaman akan kurikulum merupakan langkah pertama untuk menyusun dan mengembangkan suatu perangkat tes yang baik. Dengan pemahaman dan pendalaman akan kurikulum tersebut, maka indikator kemampuan atau ketrampilan yang menuntut untuk dikuasai subjek didik akan dapat dirumuskan dengan baik dan akan dapat dituangkan pula dalam bentuk kisi-kisi.

Tes prestasi belajar adalah mengenai sejauh mana subjek didik menyerap atau menguasai isi pelajaran, karena itu materi tes harus didasarkan pada kurikulum sebagai patokan dalam kegiatan pembelajaran. Sudah barang tentu tidak semua isi GBPP diambil sebagai dasar penulisan butir soal atau bobot tiap pokok dan sub pokok bahasan dalam GBPP tidak sama dalam penulisan butir soal. Sehingga analisis ini dilakukan guna dapat menyusun seperangkat tes yang benar-benar representatif. Cara yang dapat digunakan untuk memilih, dan menetapkan pokok dan sub pokok bahasan sebagai materi penulisan butir soal adalah dengan menelusuri seluruh isi GBPP sambil menilai esensial tidaknya suatu pokok atau sub pokok

bahasan. Sesuai dengan tingkatan essential pokok atau sub pokok bahasan itulah ditetapkan pembobotan materi butir soal.

### 3). *Analisis Buku Pelajaran*

Analisis buku pelajaran disebut juga timbangan buku, yang berarti bahwa penyusunan soal dapat juga menggunakan buku sumber (literatures), disamping buku paket yang sudah ada, selama buku-buku tersebut sesuai dengan kurikulum yang sedang berlangsung dan buku-buku tersebut memang digunakan bersama oleh pendidik dan anak didik secara keseluruhannya.

### 4). *Menentukan Kisi-kisi*

Istilah lain untuk kisi-kisi ialah blue print, table of test specification, lay out, plan dan frame work. Kisi-kisi adalah suatu daftar berbentuk matriks yang memuat komponen-komponen sebagai berikut :

#### a). Pokok bahasan

Pokok bahasan dan sub pokok bahasa merupakan ruang lingkup butir soal yang susun. Pokok bahasan dan sub pokok bahasan perlu ditetapkan pembobotannya sebagaimana diuraikan pada langkah analisis kurikulum tersebut diatas. Pembobotan ini biasanya ditetapkan dengan kategori; penting, sedang

dan kurang penting. Penetapan bobot dari masing-masing pokok atau sub pokok bahasan disarankan atas tujuan yang akan dicapai dan urgensi materi cakupan pokok atau sub pokok bahasan disarankan atas tujuan yang akan dicapai dan urgensi materi cakupan pokok atau sub pokok bahasan tersebut.

b) Aspek intelektual

Aspek intelektual berupa perilaku yang telah dimiliki subjek didik sebagai hasil belajar, diperinci ke dalam jenjang kemampuan intelektual menurut taksonomi Bloom yang akan diukur dengan tes yang dibuat, adalah mencakup: (1) Ingatan, (2) Pemahaman, (3) Penerapan, (4) Analisis, (5) Sintesis dan (6) Evaluasi

c) Bentuk soal

Bentuk soal pada umumnya dibedakan atas dua macam yaitu : bentuk uraian dan bentuk objektif.

d) Tingkat kesukaran soal

Untuk mengetahui perbandingan yang tepat antara kelompok soal yang dikategorikan mudah dan sukar, maka perlu dicantumkan dalam kisi-kisi tingkat kesukaran butir soal. Penentuan tingkat kesukaran butir soal ditetapkan berdasarkan pendapat penulis butir soal. Perbandingan jumlah butir soal sesuai dengan tingkat kesukaran butir soal biasanya adalah

27% mudah dan 27% sukar, selebihnya adalah sedang.

e) Jumlah dan proporsi butir soal

Jumlah dan proporsi butir soal ditentukan oleh waktu yang disediakan untuk mengerjakan tes tersebut. Oleh sebab itu perlu ditentukan terlebih dahulu waktu yang diperlukan mengerjakan tes tersebut, sesudah itulah baru ditentukan jumlah butir soalnya. Namun demikian perlu juga dipertimbangkan daya konsentrasi subjek didik peserta tes. Disamping waktu yang dibutuhkan untuk mengerjakan tes, perlu juga dipertimbangkan tingkat kesukaran butir soal. Untuk mengerjakan butir soal yang sukar dibutuhkan waktu yang lebih lama dibandingkan dengan mengerjakan butir soal yang tergolong mudah. Dalam waktu yang sama dapat berbeda jumlah butir soal yang dapat dikerjakan, oleh karena tingkat kesukaran butir soal yang berbeda.

## **B. Format-format Perencanaan**

### *1). Format spesifikasi butir soal*

Format ini menggambarkan alokasi butir soal sehubungan dengan tujuan tes (aspek yang akan di tes) dan alokasi butir soal sehubungan dengan isi (materi yang akan dites). Penentuan besarnya alokasi soal berdasarkan tujuan dan materi isi, disesuaikan dengan

penekanan yang terdapat dalam tujuan pengajaran dan sifatnya materi/isi bahan pengajaran. Pada pemahaman/pengertian jadi buka pada aspek pengetahuan siap. Hal ini menunjukkan pada prinsip pengajaran bermakna buka pada prinsip verbalisme.

Dari format spesifikasi butir soal/item dilanjutkan dengan perencanaan alokasi butir soal sehubungan dengan isi/pokok bahasan dan tujuan soal dan isi/ pokok bahasan yang dimaksud. Dengan kata lain, format perencanaan itu menggambarkan alokasi butir soal pada setiap tujuan tes dari setiap isi pokok bahasan. Penentuan dari prestasi yang disebut pada perencanaan tes ini didasarkan dari spesifikasi item. Untuk mendapatkan 15% (alokasi isi/ pokok bahasan), demikian seterusnya. Setelah perencanaan tes ini selesai dilanjutkan dengan lay out tes, yang menggambarkan jumlah butir soal untuk setiap bentuk soal dari setiap persentasi yang terdapat dalam perencanaan tes yang telah dibuat.

## 2). *Format Kisi-kisi Tes*

Format kisi-kisi adalah format yang dibuat terakhir dari seluruh langkah kegiatan perencanaan tes. Selain penentuan yang lebih terperinci dari lay out, pada kisi-kisi

tes diperkirakan tingkat kesulitan butir soal yang akan ditulis.

Tingkat kesukaran soal bergantung dari tujuan tes dan sulitnya pokok isi materi yang akan diuji. Kisi-kisi tes yang baik mempunyai ciri-ciri sebagai berikut: setiap item dapat dilihat apakah tujuan item, isi pokok bahasan yang dites, bentuk tes dan tingkat kesukarannya.

Tabel 3.1. Kisi-kisi Tes Hasil belajar

Bidang Studi: .....

Materi/isi pokok bahasan	Tingkat kesukaran	Tujuan tes					
		Pengetahuan		Pemahaman		Aplikasi	
		B-S	P-G	B-S	P-G	B-S	P-G
..... .....	Sukar						
	Sedang						
	Mudah						
..... .....	Sukar						
	Sedang						
	Mudah						

Keterangan: B-S = benar salah

P-G = pilihan ganda

Sebaiknya penulisan soal dimulai dari item yang mudah untuk setiap bentuk tes menuju kepada item yang sukar.

Bentuk lain dari kisi-kisi tes adalah sebagai berikut:

Tabel 3.2. Format Kisi-kisi Tes Hasil belajar

Mata Pelajaran: .....

No urut	Pokok bahasan/ Sub Pokok Bahasan	Aspek Perilaku	Tipe Soal		
			Pilihan Ganda	Uraian	Jumlah

Setelah kisi-kisi tes ini dilakukan penulisan soal sebagai berikut.

Tabel 3.3. Format Penulisan Soal

Bidang Studi: .....

Pokok bahasan/ Sub Pokok Bahasan	TIK	Aspek Perilaku	Bentuk Soal	Rumusan Soal

Catatan: TIK : Tujuan Instruksional Khusus

Aspek Perilaku : Pemahaman, Pengetahuan, Aplikasi, Sintesa, Analisa, dan Evaluasi

### C. Petunjuk Tes

Petunjuk tes dimaksudkan untuk membimbing peserta tes bagaimana mereka melaksanakan/mengerjakan/menjawab tes (pedoman mengerjakan tes). Petunjuk ini terdiri dari dua macam, yaitu : Petunjuk Umum dan Petunjuk khusus.

#### ***Petunjuk Umum, berisi;***

##### *a. Pengisian identitas peserta*

Pengisian identitas peserta adalah penulisan nama atau nomor tertentu (nomor ujian atau nomor induk), kelas atau mungkin jurusan untuk sebaiknya telah tertulis pada lembar jawaban.

*b. Tempat memberi jawaban*

Tempat memberi jawaban maksudnya adalah apakah ada lembar jawaban tersendiri atau langsung pada kertas soal atau pada kertas tersendiri. Bila bentuk tes adalah tes objektif biasanya mempunyai lembaran tersendiri agar mudah memeriksanya.

*c. Penjelasan tentang sistem penilaian*

Penjelasan tentang sistem penilaian adalah petunjuk bagaimana cara menilai. Dalam bentuk tes objektif hendaknya diberitahukan apakah mempergunakan rumus pengurangan skor bagi jawaban yang salah. Hal ini dijelaskan untuk mengurangi penerkaan jawaban dan agar peserta berhati-hati dalam memberi jawaban. Pada bentuk tes esai penjelasan berisi jumlah skor untuk masing-masing butir soal. Hal ini dimaksudkan agar mahasiswa terarah dalam memilih jawaban. Sehubungan dengan ini maka biasanya pada petunjuk disebutkan apakah ada keterkaitan untuk menjawab/ mengerjakan soal secara berurutan atau boleh memilih jawaban yang paling diketahui peserta.

***Petunjuk khusus, berisi:***

***a. tanda yang dipergunakan dalam memberi jawaban***

Tanda jawaban, dalam hubungan ini biasanya dijelaskan dalam tes objektif. Apakah dengan menyilang huruf jawaban atau melingkari atau membuat garis dibawah huruf jawaban.

***b. jawaban pilihan yang diminta,***

Jawaban pilihan yang diminta, maksudnya adalah pemilihan atas alternatif jawaban. *Pada bentuk tes benar-salah* sudah barang tentu menyilang/melingkari (B) bila pernyataan benar dan (S) bila pernyataan salah. *Pada tes pilihan ganda* yang dipilih kemungkinan jawaban yang benar atau salah atau paling benar atau paling sedikit mengandung kebenaran. *Pada tes pilihan ganda kompleks* biasa-nya pilihan adalah sebagai berikut:

Lingkari A. Jika (1), (2) dan (3) betul

B. Jika (1) dan (3) betul

C. Jika (2) dan (4) betul

D. jika hanya (4) betul

E. jika semua betul

Pada tes pilihan ganda sebab akibat, biasanya pilihan adalah sebagai berikut:

Lingkari:

- A. Jika pernyataan betul, alasan betul dan keduanya mempunyai hubungan sebab-akibat
- B. Jika pernyataan betul, alasan betul tetapi keduanya tidak menunjukkan hubungan sebab-akibat
- C. Jika pernyataan betul, alasan salah
- D. Jika pernyataan salah, alasan betul
- E. Jika pernyataan maupun alasan salah

Penulisan petunjuk tes biasanya ditempatkan pada permulaan perangkat tes. Bila perangkat tes itu terdiri dari beberapa bentuk item hendaknya petunjuk khusus ditempatkan mendahului setiap bentuk item bukan pada permulaan perangkat tes.

Petunjuk hendaknya singkat dan jelas sehingga peserta tidak terlalu sukar untuk mengingat petunjuk khusus dari tes itu.

Pada tes esai bentuk memberi jawaban itu terdiri dari :  
Jawablah secara singkat tetapi jelas !, Atau : Dahulukan mengerjakan pertanyaan yang paling mudah !,

Disamping petunjuk di atas, sesuai dengan kebutuhan, petunjuk tes objektif sering ditambahkan dengan:

- i. jangan mencoret perangkat soal
- ii. jangan menerka jawaban dan lain sebagainya.

#### **D. Waktu Tes**

Waktu tes dimaksudkan dalam hal ini adalah lamanya penyediaan waktu untuk mengerjakan tes, biasanya dituliskan dalam jumlah menit. Perhitungan jumlah waktu disediakan dan ditentukan oleh jumlah butir soal dan tingkat kesukaran soal. Semakin banyak butir soal dan semakin tinggi tingkat kesukaran soal, semakin banyak waktu pelaksanaan tes tersebut. Bagi tes objektif perkiraan jumlah waktu yang disediakan relatif lebih mudah. Untuk tes benar-salah secara umum jumlah waktu yang disediakan sekitar 30 detik untuk setiap soal dan untuk tes pilihan ganda jumlah waktu yang disediakan sekitar 60 detik, untuk setiap butir soal. Waktu yang disediakan harus ditambah untuk identitas dan memahami petunjuk tes.

Untuk tes esai jumlah waktu yang disediakan mempertimbangkan banyaknya butir soal yang sukar dan yang mudah. Ada beberapa hal yang perlu dipertimbangkan dalam penentuan waktu yang disediakan, yaitu :

- 1) Jumlah waktu yang cukup lama misalnya sampai dengan 3,5 jam atau lebih mempunyai beberapa kelemahan antara lain:
  - Menimbulkan rasa bosan/jemu.
  - Kelelahan yang tinggi sehingga peserta cenderung asal memberi jawaban
  - Memberi kesempatan untuk saling kerja sama.
- 2) Sesuaikanlah jumlah waktu dengan daya tahan fisik dan mental peserta. Untuk itu tentunya semakin tinggi tingkatan sekolah jumlah waktu dapat dilakukan semakin lama, sebaliknya semakin rendah tingkatan sekolah semakin sedikit waktu yang mungkin dapat disediakan.
- 3) Untuk mendapatkan jumlah waktu yang paling cocok, sebaiknya tes, itu dicobakan. Kepada murid/peserta (satu atau dua orang) yang sejajar kelas dengan peserta yang akan diuji.
- 4) Sering terjadi seorang guru menambah atau mengurangi waktu pelaksanaan tes dari yang direncanakan sebelumnya. Hal tersebut sangat tidak baik apabila murid/peserta telah mengetahui sebelum pelaksanaan tes dimulai. Walaupun misalnya sebagian besar peserta /murid telah selesai namun waktu masih lama, tidak baik apabila guru mengakhiri ujian sebelum waktu yang ditentukan, selama masih ada peserta yang masih bekerja dan menambahkan

waktu ujian apabila peserta belum ada yang selesai walaupun waktu telah berakhir. Namun akan lebih baik jika mempertimbangkan keadaan tersebut dalam membuat skor.

## **Penyajian Tes**

### **A. Penyajian Tes**

Penyajian tes memerlukan ketentuan yang perlu diikuti, walaupun sementara orang berpendapat bahwa penyajian tes adalah hal yang mudah dan dapat dilakukan oleh sembarang orang. Hal ini sering terlihat di sekolah-sekolah (pada pelaksanaan tes sumatif), pegawai sering diturutsertakan sebagai pengawas. Perlu diingat bahwa ketetapan penilaian turut ditentukan oleh baik tidaknya penyajian tes, terutama dalam tes objektif.

Beberapa hal ini yang berhubungan dengan penyajian tes adalah masalah fisik, lingkungan, disiplin dan pengawasan.

Penyajian tes sehubungan dengan perangkat soal menyangkut hal-hal berikut:

#### **a. Penulisan butir soal**

Bagi tes objektif terutama tes pilihan ganda, sebaiknya alternatif jawaban tersusun ke bawah

walaupun mungkin bisa dibuat dua atau lebih alternatif jawaban dalam satu-baris.

- b. Penyajian perangkat soal sebaiknya dilakukan dengan memperbanyak perangkat soal sesuai dengan jumlah peserta tes. Hal ini terutama bagi tes objektif.
- c. Kondisi tulisan perangkat soal harus jelas dan terang. Untuk itu karena kemungkinan adanya kesalahan pengetikan dan kurang baiknya cara memperbanyak perangkat soal, sebaiknya guru memeriksa lebih dahulu perangkat soal yang diperbanyak untuk memperbaiki atau membuat ralat.
- d. Pada tes esai walaupun sedikit soalnya akan lebih baik bila diperbanyak dibandingkan dengan apabila soal tersebut didiktekan.

Ada beberapa faktor yang berhubungan dengan pelaksanaan tes di sekolah, antara lain:

1). *Faktor Fisik*

Faktor fisik meliputi hal-hal yang berhubungan dengan fisik murid dan pengawas, sedangkan faktor lingkungan meliputi meja, bangku dan ruangan suasana dalam dan luar ruangan pelaksanaan tes, diantaranya yang perlu diperhatikan adalah:

- a. Tes hendaknya dilaksanakan dimana peserta mempunyai kondisi yang baik. Pada pagi hari adalah waktu yang cocok dalam melaksanakan

ujian mengingat murid masih diharapkan merasa senang/tenang. Bila dalam satu hari itu, ada beberapa mata ujian hendaknya antara ujian yang satu dengan mata ujian yang lain diberi waktu yang cukup, untuk istirahat dalam memulihkan kembali kondisi peserta setelah mengikuti mata ujian sebelumnya.

- b. Pada umumnya setiap peserta ujian mempunyai rasa takut atau tertekan sebelum mereka memulai/memasuki ruangan ujian. Untuk itu hendaknya sebelum peserta memulai ujian, pengawas terlebih dahulu menciptakan suasana tenang pada diri peserta dengan memberrikan bimbingan secara kelompok.
- c. Pengawas hendaknya tidak menunjukkan sikap-sikap yang kurang baik selama dan sebelum ujian berlangsung. Seperti yang disebut di atas (butir) b, bila pengawas menunjukkan sikap otoriter, kejam, sombong atau yang lain, akan mengakibatkan rasa takut yang menghantui peserta akan bertambah kuat. Pengawas yang tidak dalam kondisi baik akan condong bersikap yang kurang baik. Disamping itu kondisi fisik pengawas yang kurang baik akan condong kurang memperdulikan pelaksanaan pengawasan ujian.

- d. Rasa bersaing diantara peserta akan mengurangi keinginan untuk bekerja sama. Hal ini perlu ditekankan pengawas sebelum memulai ujian dengan kata lain dalam bimbingannya sebelum ujian dimulai hendaknya ditumbuhkan rasa bersaing dalam diri peserta.
- e. Pelaksanaan tes hendaknya sudah diketahui peserta jauh sebelumnya. Hal ini perlu agar peserta dapat menyiapkan diri sebaik-baiknya, baik mental, fisik dan segala alat/bahan yang diperlukan dalam pelaksanaan tes.
- f. Bangku dan meja dimana peserta akan duduk mengerjakan ujian hendaknya memadai dilihat dari kemungkinan untuk menulis dari keamanan dalam duduk. Dalam hubungan dengan meja tulis yang dipergunakan terutama dalam tes esai harus sebaik mungkin. Kondisi ruangan ujian hendaknya sehat. Pertukaran udara ruangan hendaknya memenuhi syarat kesehatan. Luasnya ruangan ujian harus disesuaikan dengan jumlah peserta ujian yang memungkinkan anak sukar untuk bekerja sama, hal ini terutama dalam tes objektif. Pemakaian ruangan yang sangat luas misalnya memungkinkan memuat peserta 80 orang, malah sampai ratusan orang kurang baik. Apalagi

penggunaan ruang terbuka, merupakan kondisi ruangan yang kurang baik. Biasanya ruangan sebaiknya cocok untuk menampung peserta paling banyak 20 orang dengan dua orang pengawas. Bila melebihi hal itu ketenangan dan objektivitas ujian sudah sukar dicapai.

## *2). Disiplin dan Pengawasan*

Disiplin dan pengawasan berhubungan dengan peraturan-peraturan/tata tertib yang ditetapkan berhubungan dalam pelaksanaan tes. Peraturan-peraturan/ tata-tertib pelaksanaan tes, biasanya dilakukan secara rinci dan tertulis yang bertujuan cukup penting.

Beberapa ketentuan/tata-tertib pelaksanaan tes yang cukup penting adalah:

### *a). Kehadiran peserta tes*

Kehadiran peserta tes biasanya diharapkan hadir di lokasi pelaksanaan 10 menit sebelum tes dimulai. Hal ini dimaksudkan untuk mengurangi keterlambatan peserta tes, dan memberikan ketenangan pada peserta sebelum memasuki ruangan serta waktu yang di pergunakan menyampaikan beberapa pengumuman dan lain sebagainya.

*b. Penyediaan baban-bahan*

Mungkin saja sesuatu tes memerlukan alat tulis-menulis tertentu misalnya pensil 2-B (biasanya dipergunakan bagi tes yang pemeriksaannya dilakukan oleh komputer).

*c. Tanda peserta ujian*

Untuk menghindarkan penipuan peserta. Tanda peserta sebaiknya memuat beberapa item identitas, misahlya nama, nomor ujian dan mungkin tanggal lahir serta alamat. Pemuatan beberapa item identitas ini menghindarkan nama yang serupa. Disamping itu bagi tes penerimaan/seleksi diperlukan penempatan foto peserta.

*d. Larangan-larangan*

Hal ini mencakup larangan sebelum dan selama ujian berlangsung, misalnya: larangan membawa kertas, buku-buku dan lainnya selain alat tulis menulis ke dalam ruangan ujian, larangan saling pinjam-meminjam alat selama ujian berlangsung, dan larangan berkomunikasi antara peserta selama ujian berlangsung

*e. Tanda-tanda waktu pelaksanaan tes*

Tes yang terdiri dari beberapa ruangan perlu memakai tanda-tanda waktu pelaksanaan tes untuk memungkinkan tes dimulai dan diakhiri secara serentak,

untuk hal ini dipergunakan suatu bunyi, misalnya bunyi lonceng.

Pada umumnya tanda waktu itu meliputi: aba-aba memasuki ruangan, aba-aba memulai ujian, aba-aba kesiapan mengakhiri ujian (biasanya 5 -10 menit) sebelum ujian berakhir, dan aba-aba akhir waktu pelaksanaan tes.

#### *f. Sangsi-sangsi*

Peserta hendaknya mengetahui sanksi-sanksi atas pelanggaran tata tertib tes, misalnya: sangsi atas keterlambatan hadir, sangsi atas tidak membawa tanda peserta, Sangsi bila bekerja sama, dan lain-lain. Hal-hal yang berhubungan dengan pengawasan selain tata tertib bagi murid diperlukan ketentuan-ketentuan bagi pengawas, misalnya: tanda pengenal pengawas, hal ini menghindarkan pengawas yang bukan orangnya, atau menghindarkan turut campurnya orang lain dalam pelaksanaan ujian. Kehadiran pengawas pada lokasi ujian, sebaiknya pengawas sudah hadir di lokasi ujian  $\pm$  20 menit sebelum ujian dimulai, yang dimaksudkan untuk perencanaan penggantian pengawas yang berhalangan hadir dan juga untuk dipergunakan dalam menyampaikan hal-hal tertentu sesuai dengan ketentuan pengawas dan juga dalam membagi/menerima bahan-bahan ujian. Ketentuan ruang gerak pengawas, misalnv pengawas

tidak boleh memasuki ruangan ujian lain, pengawas tidak diperkenankan terlalu banyak bergerak di dalam ruang ujian selama ujian berlangsung.

## **B. Pemeriksaan Jawaban Tes**

Sesuai dengan bentuk tes, pemeriksaan jawaban secara umum terbagi atas tes esai dan tes objektif.

### **(1). Pemeriksaan jawaban tes esai**

- a. Kunci jawaban tes esai haruslah lebih dahulu direncanakan dan sebaiknya kunci ini telah dibuat segera setelah butir soal ditulis. Kunci jawaban berisi pokok-pokok jawaban yang dikehendaki dan untuk setiap pokok jawaban ditentukan nilai tertinggi bagi jawaban yang paling mendekati kebenaran/sesuai dengan yang dikehendaki.
- b. Skor untuk setiap soal memakai sistem "weight", yaitu perbandingan berat harga penilaian dari setiap soal adalah berdasarkan tingkat kesukaran soal. Untuk baiknya tingkat berat harga penilaian tingkat kesukaran soal dibedakan atas 3 tingkatan yaitu sukar, sedang dan mudah. Sebaiknya dihindarkan tingkat kesukaran yang ekstrim yaitu bukti soal yang sangat sukar atau sangat mudah. Dengan adanya bobot nilai untuk setiap tingkatan kesukaran tes maka perbedaan besarnya harga

setiap butir soal tidak perlu lagi dibuat. Jadi skor untuk setiap butir soal didistribusikan secara merata, jadi bila dipergunakan skor dari 0 s/d 100, maka setiap jawaban untuk setiap butir, soal diskor paling tinggi adalah 20, yakni hasil pembagian dari skor tertinggi (100) dibagi dengan jumlah butir soal (5).

- c. Rentangan yang dipakai dalam menskor tes secara keseluruhan atau menskor setiap soal hendaknya sesederhana mungkin, jadi tidak terlalu kecil atau terlalu besar. Hal ini dimaksudkan agar memudahkan pemeriksaan dalam mengelompokkan jawaban, karena rentangan skor yang terlalu kecil akan menghasilkan skor-skor ekstrim. sebaliknya rentangan skor yang terlalu banyak akan menghasilkan skor yang sangat homogen. Besarnya rentangan skor yang dipergunakan untuk setiap butir soal adalah 10, angka untuk keseluruhan tes adalah 100.
- d. Setiap jawaban harus dibaca secara keseluruhan untuk mendapatkan inti jawaban dan kualitas jawaban, barulah kemudian mengklasifikasikannya pada rentangan kebenaran jawaban dan seterusnya memberi skor.

- e. Pergunakanlah *Whole method* dalam memeriksa lembar jawaban, yaitu pemeriksaan jawaban dari suatu nomor untuk semua lembar jawaban. Hal ini dimaksudkan untuk dapat membandingkan jawaban dari seluruh peserta' ujian dan menghindari pemberian skor yang berbeda bagi jawaban yang sebenarnya mempunyai isi yang sama, dan juga dapat memudahkan pemeriksaan serta setiap murid mendapat sikap/perlakuan yang sama dari guru dalam objektifitas pemeriksaan. Hal ini sukar diperoleh bila pemeriksaan dilakukan sekaligus, untuk setiap jawaban dalam setiap pertanyaan, karena, murid yang dinilai pada saat-saat terakhir akan cenderung mendapat penilaian yang kurang objektif mengingat semakin kurangnya kondisi pemeriksaan.
- f. Bagi ujian yang sangat penting, biasanya pemeriksaan untuk satu pekerjaan dilakukan dua kali oleh dua orang guru. Skor yang dipakai adalah rata-rata dari kedua penilaian guru tersebut dengan catatan perbedaan nilai dari kedua guru tidak jauh berbeda (paling besar dua angka).
- g. Waktu yang digunakan untuk memeriksa jawaban hendaknya benar-benar direncanakan, jadi bukan

"dimana ada kesempatan" atau "sambil lalu" seperti kebiasaan yang sering terjadi, waktu yang disediakan harus seimbang dengan jumlah pekerjaan yang akan diperiksa. Dengan kata lain, sebaiknya jangan dilakukan pemeriksaan ujian apabila kondisi fisik dan psikis pemeriksa kurang baik.

## **(2) Pemeriksaan jawaban tes objektif**

- a. Pemeriksaan jawaban tes objektif jauh lebih mudah dari pemeriksaan jawaban tes esai. Biasanya pemeriksaan dilakukan dengan memakai lembar kunci jawaban. Lembar kunci jawaban adalah salah satu dari lembar jawaban yang tidak terpakai dimana huruf jawaban yang benar untuk setiap soal dibuang sehingga terdapat lobang pada tempat jawaban. Dengan menetapkan lembar kunci jawaban diatas lembar jawaban peserta dapat dihitung berapa jumlah jawaban yang benar, selanjutnya dihitung jumlah butir soal yang tidak dijawab. Dengan mengetahui keduanya dapat pula dihitung jumlah jawaban yang salah.
- b. Pada saat ini pemeriksaan jawaban tes objektif dilakukan dengan memakai komputer. Pemakaian alat ini hendaknya didasarkan atas jumlah peserta yang sangat banyak, dan mempertinggi ketelitian pemeriksaan. Hal ini dimaksudkan mengingat pembiayaan pelaksanaan tes sangat tinggi, karena, memerlukan alat tulis-menulis khusus. Disamping itu sedikit kesalahan dalam mengikuti petunjuk mengakibatkan kerugian besar pada peserta tes.

- c. Menerka jawaban adalah satu kelamahan yang terdapat dalam tes objektif. Untuk itu dalam mencari skor seftap peserta dalam tes objektif dipergunakan rumus yang bertujuan mengurangi pengaruh terkaan terhadap skor tes.

Rumus yang dimaksud adalah:

$$Sc = R - \frac{W}{Op - 1}, \text{ dimana :}$$

- Sc = skor akhir  
R = Right (jumlah jawaban benar)  
W = Wrong (jumlah jawaban salah)  
Op = Option (jumlah kemungkinan jawaban)  
1 = Angka satu (bilangan tetap)

Rumus ini berlaku *untuk tes pilihan ganda* dan untuk tes benar-salah, karena kemungkinan jawaban adalah 2 buah maka rumus tersebut menjadi,:

$$Sc = R - W$$

Bagi tes menjodohkan rumus koreksi tersebut diatas tidak dipergunakan, jadi tidak ada pengurangan skor bagi jawaban yang salah.

Skor untuk tes menjodohkan ditentukan oleh jumlah jawaban (pasangan) yang benar, jadi:  $Sc = \text{Jumlah jawaban yang benar}$

- d. Dalam satu perangkat ujian, sering dipergunakan beberapa bentuk, tes mungkin terdapat tes benar-salah, tes menjodohkan dan tes pilihan ganda malah ada satu atau dua nomor pertanyaan terdiri dari tes esai. Mengingat tingkat kesukaran soal berhubungan dengan bentuk tes, ditambah pula dengan jumlah butir soal yang tidak sama maka sebaiknya untuk menetapkan skor akhir peserta, dipergunakan lagi *sistim weight* sesuai dengan tingkat kesukarannya bentuk soal dan banyaknya pertanyaan untuk setiap bentuk soal.

## **Rangkuman**

1. Dalam merencanakan seperangkat tes hasil belajar yang pertama harus dilakukan adalah menentukan tujuan tes, melakukan analisis kurikulum, analisis buku pelajaran, dan meninjau kembali program pengajaran (program tahunan, semesteran dan satuan-satuan pelajaran yang disusun).
2. Sebelum menyusun butir soal harus dilakukan perencanaan-perencanaan butir soal, kisi-kisi butir soal yang merupakan pedoman dalam menyusun butir-butir soal. Petunjuk mengerjakan soal dan jumlah waktu untuk mengerjakan soal harus ditetapkan setepat mungkin. Jumlah waktu yang

terlalu lama atau terlalu cepat akan mengurangi kualitas dan objektivitas perangkat tes.

3. Dalam menyajikan tes harus diperhatikan hal yang berhubungan dengan penggandaan butir soal, serta persyaratan lingkungan yang baik (termasuk kondisi anak). Disiplin dan pengawasan penyajian tes harus ditetapkan dan diketahui/ dipahami oleh peserta ujian dan pengawas ujian (terutama sekali dimaksudkan pada ujian-ujian yang dilaksanakan secara bersamaan dan serentak seperti ujian UAN atau EBTANAS).
4. Untuk meningkatkan sifat objektivitas tes, maka seharusnya penyusunan butir, soal harus diikuti oleh pedoman pemeriksaan dan penentuan nilai serta kunci jawaban dari perangkat tes. Dengan kata lain bahwa sebelum ujian dilaksanakan pedoman pemeriksaan, penentuan nilai serta kunci jawaban ujian telah ditetapkan.

# **BAB IV**

## **PENGOLAHAN DATA HASIL PENILAIAN**

### **Data Hasil Penilaian**

#### **A. Pendahuluan**

Data hasil pengukuran melalui alat penilaian tertentu, misalnya tes, baik tes objektif maupun tes esai, berupa data kuantitatif, yaitu angka-angka atau bilangan numerik. Angka atau bilangan tersebut adalah skor hasil pengukuran yang biasa disebut skor mentah. Agar skor mentah ini mempunyai makna nilai sehingga dapat ditafsirkan untuk menentukan prestasi atau kemampuan siswa, perlu diolah menjadi skor masak melalui teknik statistika. Proses mengubah skor mentah menjadi skor masak dengan menggunakan teknik statistika disebut pengolahan data. Pengolahan data dimaksudkan untuk (a) menentukan posisi dan prestasi atau nilai siswa dibandingkan dengan kelompoknya dan untuk (b) menentukan batas kelulusan berdasarkan kriteria yang ditentukan. Berikut ini akan dibahas cara-cara mengolah

data hasil pengukuran melalui alat-alat penilaian yang lazim digunakan, yakni tes prestasi belajar.

## **B. Batas Kelulusan**

Batas kelulusan hasil penilaian mempunyai kaitan erat dengan kedua sistem penilaian tersebut. Dengan demikian, ada batas kelulusan yang berorientasi kepada penilaian acuan norma, yakni batas lulus aktual dan batas lulus ideal. Batas lulus tersebut mengisaratkan penggunaan nilai rata-rata kelas dan simpangan baku. Disamping itu ada pula batas kelulusan yang berorientasi kepada sistem penilaian acuan patokan, yakni batas lulus purposif (ditentukan berdasarkan kriteria tertentu).

### **1. Batas Lulus Aktual**

Batas lulus aktual didasarkan atas nilai rata-rata aktual atau nilai rata-rata yang dapat dicapai oleh kelompok siswa. Unsur yang diperlukan untuk menetapkan batas lulus aktual adalah nilai rata-rata aktual dan simpangan baku aktual. Biasanya skor yang dinyatakan lulus adalah skor di atas  $(\bar{X} + 0,25 \text{ SD})$  di mana  $\bar{X}$  = nilai rata-rata kelas dan SD adalah simpangan baku atau deviasi standar.

Contoh:

Misalkan kelas II SMA diberi tes bahasa Inggris dengan menggunakan bentuk pilihan ganda sebanyak 60 pertanyaan. Setiap pertanyaan yang dijawab benar diberi skor satu sehingga skor maksimal yang mungkin dicapai siswa sebanyak 60. kemudian dihitung nilai rata-ratanya dari semua siswa yang ada di kelas tersebut, misalnya 25 dan simpangan bakunya (S) adalah 8,0. Dengan demikian, skor yang dinyatakan lulus adalah  $25 + 0,25(8,0) = 27$ . Skor-skor di atas 27 dinyatakan lulus sedangkan skor di bawah 27 dinyatakan gagal atau tidak lulus.

## **2. Batas lulus idelal**

Batas lulus ideal hampir sama dengan batas lulus aktual, yakni menentukan batas lulus dengan menggunakan nilai rata-rata dan simpangan baku ideal. Nilai rata-rata dan simpangan baku dalam batas lulus ideal mudah dihitung yakni menggunakan aturan sebagai berikut: Nilai rata-rata ideal adalah setengah dari maksimum skor. Simpangan baku ideal adalah sepertiga dari nilai rata-rata ideal.

Contoh:

Kembali kepada contoh dalam batas lulus aktual. skor maksimum yang mungkin dicapai dari tes bahasa Inggris adalah 60. rata-rata idealnya

adalah setengah dari 60, yakni 30, sedangkan simpangan batas adalah sepertiga dari rata-rata ideal, yakni 10.

Batas lulusnya adalah  $30 + 0,25 (10) = 32,50$ .

### **3. Batas lulus purposif**

Batas lulus purposif mengacu kepada penilaian acuan patokan sehingga tidak perlu menghitung nilai rata-rata dan simpangan baku. Dalam hal ini ditentukan kriterianya, misalnya 75%. Artinya skor yang dinyatakan lulus adalah skor di atas 75% dari skor maksimum. Dalam contoh di atas maka batas lulusnya adalah 75% dari skor maksimum. Dalam contoh di atas maka batas lulusnya adalah 75% dari 60, yakni 45. skor yang besarnya di atas 45 dinyatakan lulus dan yang berada dibawahnya dinyatakan gagal.

Maka tinggi kriteria kelulusannya, maka makin tinggi pula kualitas hasil belajar yang dituntutnya. Sebaliknya, makin rendah kriterianya, makin rendah pula kualitas hasil belajar yang dihasilkannya.

Ketiga batas lulus di atas sering digunakan oleh para guru di sekolah terutama dalam penilaian sumatif atau ujian akhir tahun.

### C. Kecenderungan Memusat dan Keragaman

Dalam uraian mengenai batas lulus telah disinggung pentingnya nilai rata-rata dan simpangan baku atau deviasi standar. Istilah-istilah tersebut menjadi salah satu pembahasan dalam statistika. Nilai rata-rata merupakan salah satu ukuran kecenderungan memusat di samping ukuran-ukuran lainnya, yakni media dan modus. Sedangkan simpangan baku menjadi salah satu pembahasan atau bagian dari ukuran keragaman atau disversi di samping ukuran lainnya, yakni rentangan dan variansi.

Pada bagian ini akan dijelaskan ukuran-ukuran tersebut, baik pengertian-nya maupun cara perhitungannya serta penggunaannya untuk kepentingan pengolahan data hasil penilaian.

#### 1. Ukuran kecenderungan memusat

Ada tiga ukuran kecenderungan memusat yang paling banyak digunakan, yakni *modus*, *median*, dan *mean* (nilai rata-rata).

**Modus** adalah skor yang paling banyak frekuensinya sehingga tidak perlu dihitung, cukup dilihat dari penyebaran skor, kemudian dicari skor mana yang permunculannya paling sering. Oleh sebab itu, modus

merupakan ukuran kecenderungan memusat yang paling sederhana. Perhatikan contoh berikut ini:

Skor hasil tes matematika dari 7 orang siswa adalah sebagai berikut:

6

7

7

6 dari data tersebut dengan mudah dapat ditentukan

7 modusnya, yakni 7, sebab skor 7 paling banyak di

8 peroleh siswa artinya mempunyai frekuensi paling

7 tinggi

Dengan demikian, melalui modus dapat diketahui secara cepat tingkat penguasaan bahan pengajaran atau hasil belajar dari sebagian besar siswa.

**Median** adalah titik tengah dari data yang telah diurutkan sehingga membatasi, setengahnya dibawahnya dan setengah lagi berada di atasnya.

Contoh:

Kita kembali kepada data diatas, yakni skor hasil tes matematika dari 7 orang siswa. Data tersebut, jika diurutkan dari skor terendah menuju skor tertinggi, menjadi sbb:

6

6

7 Skor 7 yang dilingkari adalah median sebab menjadi titik

7 tengah distribusi skor yang telah diurutkan sehingga membatasi, setengahnya berada di bawahnya dan setengah lagi

7 di atasnya. Apabila titik tengah tersebut jatuh pada dua

8 skor, maka titik tengahnya adalah jumlah dua skor dibagi dua.

Contoh:

6

6 Median dari data nilai di samping adalah 6 dan 7, sehingga

6

jika dijumlahkan dan dibagi dua nilainya menjadi 6,5

7

7

8

**Mean** atau rata-rata diperoleh dengan menjumlahkan seluruh skor dibagi dengan banyaknya subjek. Secara sederhana rumusnya adalah:

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

$\bar{X}$  = rata-rata (mean)  
 $\sum X$  = jumlah seluruh skor  
 $N$  = banyaknya subjek

Kita kembali kepada data diatas, yakni skor hasil tes matematika dari 7 orang siswa. Rata-rata untuk data adalah 6,85.

6	
6	
7	Jumlah seluruh skor adalah 48
7	Banyaknya subjek adalah 7
7	Rata-ratanya (mean) adalah $48/7 = 6,85$
7	
<u>8</u>	+
48	

Dapat juga dihitung dengan menggunakan rumus:

$$\bar{X} = \frac{1}{N} \sum_{i=1}^n f_i X_i$$

Dimana  $f_i X_i$  adalah frekuensi x

Data di atas dapat dihitung menjadi:

$$\bar{X} = \frac{1}{7} (2 \times 6 + 4 \times 7 + 1 \times 8) = 6,85$$

Cara di atas tidak praktis jika dihadapkan dengan jumlah subjek yang banyak dan skor-skor yang besar. Oleh karena itu perlu dilakukan pengelompokan skor ke dalam interval tertentu,

kemudian dibuat distribusi akor dalam bentuk tabel distribusi.

Contoh:

Diketahui data skor hasil tes ujian matematika dari 30 orang siswa.

20	26	30	35	40	45
28	32	33	39	34	39
37	38	44	36	42	40
44	43	41	44	49	48
46	47	54	50	57	62

Langkah yang ditempuh adalah sebagai berikut:

- Tentukan kelas interval, biasanya dengan menggunakan bilangan ganjil. Misalnya kita pakai kelas interval 5
- Buatlah tabel distribusi frekuensi dengan menggunakan kelas interval 5 mulai dari kelompok skor terendah sampai kelompok skor tertinggi.
- Hitunglah frekuensi skor pada setiap kelompok dengan cara membuat turus (tally).

Dari data di atas dapat dibuat tabel distribusi frekuensi seperti tampak pada tabel 4.1. berikut.

Tabel 1. Distribusi frekuensi skor data nilai matematika

Kelompok Skor (Interval =5)	Turus	Frekuensi
20 – 24	/	1
25 – 29	//	2
30 – 34	////	4
35 – 39	//// /	6
40 – 44	//// ///	8
45 – 49	////	5
50 - 54	//	2
55 – 59	/	1
60 - 64	/	1
Jumlah		30

Dengan menggunakan tabel distribusi frekuensi skor, dapat menghitung atau mencari nilai rata-rata dan median dengan lebih praktis. Rumus untuk mencari nilai rata-rata data yang dikelompokkan antara lain adalah:

$$(1) \bar{X} = \sum \frac{f \cdot X}{N}$$

Dapat juga digunakan rumus:

$$(2) \bar{X} = u + i \left( \sum \frac{f \cdot d}{n} \right)$$

Contoh pada Tabel 4.1, maka akan digunakan menjadi Tabel 4.2, sebagai berikut:

Kelompok Skor (Interval =5)	Titik Tengah (X)	f	fX
20 – 24	22	1	22
25 – 29	27	2	4
30 – 34	32	4	128
35 – 39	37	6	222
40 – 44	42	8	336
45 – 49	47	5	235
50 - 54	52	2	102
55 – 59	57	1	57
60 - 64	62	1	62
Jumlah		30	1220

Titik tengah adalah jumlah skor bawah interval dengan skor atas interval dibagi dua. Dalam contoh untuk interval ke satu adalah  $\frac{20 + 24}{2} = 22$ , dst.

Dengan rumus di atas dapat dihitung nilai rata-ratanya sebagai berikut:

$$\bar{X} = \sum \frac{f \cdot X}{N} = \frac{1220}{30} = 40,67$$

Dengan menggunakan rumus kedua diperlukan tabel distribusi tersendiri yang unsur-unsurnya berbeda dengan tabel di atas. Dalam tabel tersebut (Table 4.3) deviasi nol ditempatkan pada kelompok nilai yang mempunyai frekuensi terbanyak, yaitu pada kelompok 40-44, Dengan

demikian, pada kelompok skor di bawahnya, deviasi menjadi bilangan negatif yang diurutkan mulai dari -1. Sedangkan pada kelompok skor di atasnya, deviasi menjadi bilangan positif, diurutkan mulai dari 1. Notasi  $u$  dalam rumus adalah titik tengah dari kelompok skor yang memiliki deviasi nol, yaitu 42.

Kelompok Skor (Interval =5)	f	Deviasi (d)	fd
20 – 24	1	-4	-4
25 – 29	2	-3	-6
30 – 34	4	-2	-8
35 – 39	6	-1	-6
40 – 44	8	0	0
45 – 49	5	1	5
50 - 54	2	2	4
55 – 59	1	3	3
60 - 64	1	4	4
Jumlah			-8
			30

Keterangan:  $f$  = frekuensi  
 $d$  = deviasi (penyimpangan skor dari rata-rata)

Rumus untuk mencari rata-rata adalah:

$$\begin{aligned} \bar{X} &= u + i \left( \sum \frac{f \cdot d}{n} \right) = 42 + 5 \left( \sum \frac{-8}{30} \right) = 42 - (40/30) \\ &= 40,67 \end{aligned}$$

Cara ini ternyata lebih praktis dan mudah dalam menghitungnya.

Bagaimana untuk mencari median dari data yang dikelompokkan. Seperti mencari rata-rata, untuk menghitung median dapat digunakan rumus sebagai berikut:

$$Me = L_1 + i \left( \frac{\left( \frac{1}{2}n - (\sum f_1) \right)}{f_{med}} \right)$$

dimana: Me = median yang dicari  
 $L_1$  = batas bawah kelas median  
*i* = interval  
 $(\sum f_1)$  = jumlah frekuensi kelas yang lebih rendah dari kelas median  
*n* = banyaknya subjek

Tabel 4.4. Distribusi skor

Kelompok Skor (Interval =5)	f	cf
20 – 24	1	30
25 – 29	2	29
30 – 34	4	27
35 – 39	6	23
40 – 44	8	17
45 – 49	5	9
50 - 54	2	4
55 – 59	1	2
60 - 64	1	1
Jumlah		
30		

Keterangan: cf adalah frekuensi kumulatif, diperoleh dengan menjumlahkan frekuensi dari bawah ke atas

1 + 1 = 2, 2 + 2 = 4, 4 + 5 = 9, dst

Kelas median ada pada kelompok skor 40-44 sebab setengah dari n, yaitu  $n_i = 30/2 = 15$ , ada pada cf = 17.  $L_1$  dihitung dari batas bawah kelompok skor 40-44, yaitu 39,5. Interval adalah 5. frekuensi median adalah 8. ( $\sum f_1$ ) adalah 9. maka hasilnya adalah:

$$Me = L_1 + i \left( \frac{\left( \frac{1}{2}n - (\sum f_1) \right)}{f_{med}} \right) = 39,5 + 5 \left( \frac{(15 - 9)}{8} \right) = 39,5 + (5)(0,75) = 43,25$$

## 2. Ukuran Keragaman

Ukuran keragaman yang paling sederhana adalah *rank*, yakni selisih skor tertinggi dengan skor terendah. Sedangkan ukuran keragaman lain yang paling banyak digunakan adalah simpangan baku dan variansi. Variansi adalah pangkat dua simpangan baku. Notasi simpangan baku adalah s, sedangkan notasi variansi adalah  $s^2$ . simpangan adalah penyimpangan nilai dari rata-ratanya. Makin besar simpangan, makin beragam nilai atau skor yang diperoleh siswa. Sebaliknya, makin kecil simpangan berarti skor-skor tersebut atau skor yang dicapai

cenderung homogen atau merata. Tes yang baik tentunya mempunyai simpangan baku yang kecil.

Cara menghitung simpangan baku ada untuk data yang tidak dikelompokkan dan ada untuk data yang dikelompokkan. Cara menghitung simpangan baku untuk data yang tidak dikelompokkan menggunakan rumus sebagai berikut:

$$S = \sqrt{\frac{\sum (X - \bar{X})^2}{n}} \quad \begin{array}{l} X = \text{skor nilai yang dicapai, } \bar{X} \\ = \text{rata-rata} \end{array}$$

Contoh:

X	$X - \bar{X}$	$(X - \bar{X})^2$
7	7-7=0	0
8	8-7=1	1
9	9-7=2	4
6	6-7=-1	1
5	5-7=-2	4
$\Sigma =$ 35	===	$\Sigma =$ 10

$$\begin{array}{l} \Sigma X = 35, \quad n = 5 \\ \bar{X} = 35/5 = 7 \end{array}$$

$$S = \sqrt{\frac{10}{5}} = \sqrt{2} = 1,41$$

Untuk data yang dikelompokkan digunakan rumus:

$$S = i \sqrt{\frac{\sum fd^2}{n} - \left( \frac{\sum fd}{n} \right)^2}$$

Dengan diketahuinya simpangan baku, variansi dapat dihitung, yaitu pangkat dua dari simpangan baku.

Contoh: Tabel 4.5. Distribusi skor

Kelompok Skor (Interval =5)	f	Deviasi (d)	fd	fd <sup>2</sup>
20 – 24	1	-4	-4	16
25 – 29	2	-3	-6	18
30 – 34	4	-2	-8	16
35 – 39	6	-1	-6	6
40 – 44	8	0	0	0
45 – 49	5	1	5	5
50 – 54	2	2	4	8
55 – 59	1	3	3	9
60 – 64	1	4	4	16
Jumlah			-8	94
30				

Maka simpangan baku adalah:

$$\begin{aligned}
 S &= i \sqrt{\frac{\sum fd^2}{n} - \left(\frac{\sum fd}{n}\right)^2} \\
 &= 5 \sqrt{\frac{94}{30} - \left(\frac{-8}{30}\right)^2} = 8,75
 \end{aligned}$$

Dari perhitungan tersebut maka variansi dapat dicari yaitu  $(8,75)^2 = 76,56$

Nilai rata-rata dan simpangan baku sangat diperlukan untuk mengolah data hasil tes maupun untuk keperluan analisis lebih lanjut. Untuk keperluan penilaian, rata-rata dan simpangan baku dapat digunakan dalam :

- a) menentukan batas kelulusan, terutama batas lulus aktual dan batas lulus ideal.
- b) Membuat konversi nilai
- c) Mengubah skor mentah ke dalam skor baku seperti skor z dan skor T. Kedua-duanya akan dibahas kemudian.
- d) Menentukan atau menghitung korelasi, signifikansi, dan lain-lain

Oleh sebab itu, pemahaman dan keterampilan menghitung nilai rata-rata dan simpangan baku sangat diperlukan.

## **Tranformasi Data Penilaian**

### **A. Skor Baku (skor z dan skor T)**

Jika akan membandingkan dua sebaran skor yang berbeda standar yang digunakannya, misalnya yang satu menggunakan nilai standar sepuluh dan yang satu lagi menggunakan standar 100, sebaiknya dilakukan transformasi atau mengubah skor mentah ke dalam skor baku. Ada dua macam skor baku, yakni skor z dan skor T. Skor z dapat dihitung dengan membagi selisih skor dan nilai rata-ratanya dengan simpangan bakunya.

$$Z = \frac{X - \bar{X}}{S}$$

Contoh :

Iwan memperoleh skor 75 dari skor maksimum 100. rata-rata kelas atau mean adalah 60 dan simpangan bakunya 10.

$$\text{Skor z iwan adalah } \frac{75 - 60}{10} = 1,5$$

Misalkan kita akan membandingkan dua skor yang berbeda rentangnya. Skor pertama menggunakan 0-10 dan yang satu lagi menggunakan rentangan 0-100. Iwan memperoleh skor matematika 6.5 dalam standar 0-10. Rata-rata kelas adalah 6. Simpangan bakunya adalah 0.8. Sedangkan skor Bahasa Inggris sebesar 80 dari rentangan 0-100. rata-rata kelas untuk Bahasa Inggris adalah 75 dengan simpangan baku 10.

Pertanyaannya: dalam pelajaran manakah Iwan lebih unggul?

Untuk menjawab pertanyaan tersebut kita gunakan skor z.

$$\text{Skor z untuk Matematika adalah } \frac{6,5 - 6}{0,8} = 0,625$$

$$\text{Skor z untuk bahasa Inggris } \frac{80 - 75}{10} = 0,50$$

Dengan membandingkan skor z di atas dapat disimpulkan bahwa Iwan lebih unggul dalam hal matematika dari pada dalam bahasa Inggris.

Kelemahan skor z ialah berhadapan dengan bilangan negatif dan bilangan pecahan sehingga kurang praktis. Untuk itu dapat digunakan skor baku lainnya, yakni skor T.

Skor T diperoleh dengan mengalihkan skor z kepada bilangan 10, kemudian ditambah dengan bilangan 50 sehingga diperoleh skor dalam rentangan 0-100.

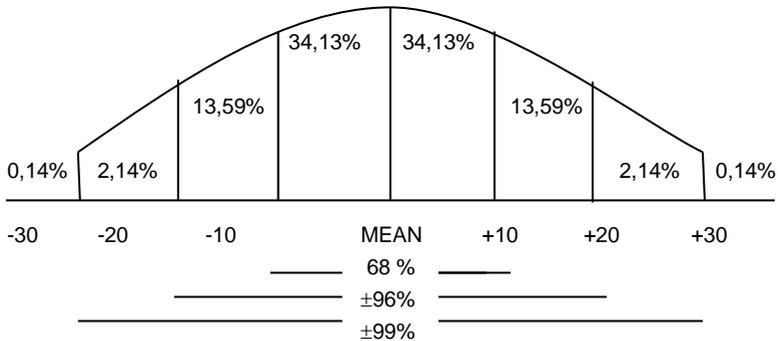
Contoh di atas adalah mengenai skor z Iwan dalam matematika dan Bahasa Inggris. Jika digunakan skor T, maka skor Iwan menjadi:

Skor T matematika adalah  $(0,625 \times 10) + 50 = 56,25$ .

Skor T bahasa Inggris adalah  $(0,5) (10) + 50 = 55,0$ .

Nilai rata-rata, simpangan baku, dan skor z sangat penting kegunaannya untuk melukiskan sebaran skor dalam kurva normal. Kurva normal adalah distribusi atau penyebaran teoritis dari skor dalam hal sebagian besar skor yang diperoleh subjek mengelompok di sekitar rata-ratanya. Frekuensi skor subjek akan menjadi semakin kecil sewaktu bergerak menjauhi rata-rata, baik ke arah

kiri yang lebih kecil dari rata-rata maupun ke sebelah kanan yang lebih besar dari rata-rata. Kira-kira 34 persen dari subjek dalam sebaran kurva normal terletak di antara rata-rata dan satu simpangan baku di atas di bawah rata-rata. Di antara satu simpangan baku dan dua simpangan baku dari rata-rata di kedua sisinya (kiri dan kanan) memuat sekitar sekitar 14 persen. Jika dilukiskan akan tampak seperti dalam gambar berikut ini.



Melalui kurva normal kita dapat menghitung persentase subjek di bawah dan di atas setiap skor z dengan melihat tabel kurva normal. Misalnya skor z Martina dalam matematika adalah 0,625 atau 0,125 di atas rata-rata. Tabel A distribusi normal menunjukkan bahwa 0,125 z adalah 0,394, artinya 39,44 persen dari subjek terletak di antara nilai z dari rata-rata sebaran skor tersebut. Karena skor z tersebut bilangan positif, berarti di atas rata-rata. Rata-rata adalah 50%. Dengan demikian, bilangan tadi

ditambah dengan 50% sehingga menjadi 89,44%. Ini berarti bahwa 89,44% dari subjek berada dibawah skor z 0,625.

Apabila skor z itu bilangan negatif, maka perhitungannya harus di balik, yakni mengurangkannya dari 50%. Misalnya skor z adalah -0,40. daerah antara rata-rata dengan skor z = 0,40 adalah 0,3446. dengan mengurangkan 34,46% dari 50% diperoleh 15,54%. Artinya, sekitar 15,54% dari subjek berada di bawah nilai z = -0,40.

## **B. Konversi Nilai**

Seperti telah dijelaskan sebelumnya bahwa standar yang sering digunakan dalam menilai hasil belajar dapat dibedakan ke dalam beberapa kategori, yaitu:

- a. standar seratus (0 – 100)
- b. standar sepuluh (0 – 10), dan
- c. standar empat (1 – 4) atau dengan huruf (A-B-C-D)

Sedangkan skor baku baik skor z maupun skor T, jarang digunakan. Standar-standar tersebut (z dan T) hanya digunakan untuk keperluan khusus, misalnya untuk menganalisis kecakapan seseorang dibandingkan dengan orang lain dan membandingkan dua skor yang berbeda standarnya.

Konversi nilai dapat dilakukan dari standar seratus ke standar sepuluh dan ke standar empat, atau bisa juga dari standar sepuluh ke standar seratus atau ke standar empat. Dalam konversi nilai digunakan dua cara, yakni cara yang menggunakan rata-rata dan simpangan baku dan cara tanpa menggunakan rata-rata dan simpangan baku.

### 1. Konversi tanpa menggunakan nilai rata-rata dan simpangan baku.

Cara ini sangat sederhana, yaitu dengan menentukan kriteria sebagai dasar untuk melakukan konversi nilai. Misalnya dengan menggunakan kriteria dalam bentuk persentase.

Tabel 4.6. Kriteria nilai konversi

Persentase Jawaban (%)	Nilai Konversi		
	Huruf	Standar 10	Standar 100
90 – 99	A	9	4
80 – 89	B	8	3
70 – 79	C	7	2
60 -69	D	6	1
Kurang dari 60	Gagal	Gagal	Gagal
		Nilai 10 bila mencapai 100 %	

Contoh penggunaan:

Misalkan kepada siswa berikan tes fisika dalam bentuk tes objektif pilihan ganda sebanyak 60 soal. Jawaban

yang benar diberi skor satu sehingga skor maksimum yang dicapai siswa adalah 60. Berdasarkan kriteria di atas, konversi nilai dalam standar huruf, standar sepuluh dan standar empat adalah sebagai berikut:

Tabel 4.6. Kriteria nilai konversi

Skor Mentah	Nilai Konversi		
	Huruf	Standar 10	Standar 4
54 – 60	A	9/10	4
48 – 53	B	8	3
42 – 47	C	7	2
36 – 41	D	6	1
Kurang dari 36	Gagal	Gagal	Gagal
		Nilai 10 bila mencapai 60	

## 2. Konversi nilai dengan menggunakan nilai rata-rata dan simpangan baku

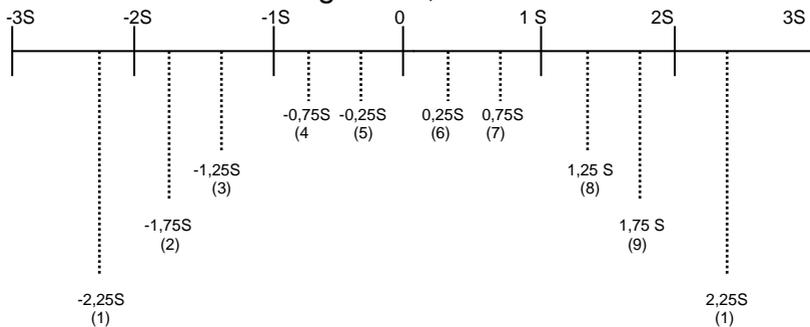
Konversi nilai ini perlu dihitung terlebih dahulu, harga nilai rata-rata dan simpangan baku yang diperoleh siswa, kemudian terhadap nilai-nilai atau skor mentah tersebut dilakukan konversi. Konversi biasanya dilakukan terhadap standar 10 atau standar huruf dan atau standar empat.

Kriteria yang digunakan untuk melakukan konversi skor mentah ke dalam standar 10 adalah sebagai berikut:

$$\begin{aligned}
 M + 2,25 S &= 10 \\
 M + 1,75 S &= 9 \\
 M + 1,25 S &= 8 \\
 M + 0,75 S &= 7 \\
 M + 0,25 S &= 6 \\
 M + S &= 5 \\
 M - 0,25 S &= 4 \\
 M - 0,75 S &= 3 \\
 M - 1,25 S &= 2 \\
 M - 2,25 S &= 1
 \end{aligned}$$

M = nilai rata-rata  
S = simpangan baku (deviasi standar)

Jika dilukiskan dalam gambar, kriteria tersebut adalah



Nilai rata-rata dan simpangan baku dapat dicari dengan menggunakan rumus-rumus yang telah dijelaskan di atas.

Contoh:

Tes diberikan kepada siswa dalam bentuk tes objektif sebanyak 90 soal. Setiap soal yang dijawab benar diberi skor satu sehingga skor maksimum yang dapat

dicapai siswa adalah 90. setelah diperiksa, ternyata skor yang paling tinggi mencapai 50 dan skor terendah 30. nilai rata-rata (setelah dihitung) adalah 40 dan simpangan bakunya 4,0. dengan menggunakan rumus atau kriteria di atas, diperoleh nilai dalam standar sepuluh sebagai berikut

Skor mentah	Standar 10
$40 + (2,25)(4,0) = 49$	10
$40 + (1,75)(4,0) = 47$	9
$40 + (1,25)(4,0) = 45$	8
$40 + (0,75)(4,0) = 43$	7
$40 + (0,25)(4,0) = 41$	6 (batas lulusnya)
$40 - (0,25)(4,0) = 39$	5
$40 - (0,75)(4,0) = 27$	4
$40 - (1,25)(4,0) = 35$	3
$40 - (1,75)(4,0) = 33$	2
$40 - (2,25)(4,0) = 31$	1

Nilai rata-rata dan simpangan baku di atas dihitung dari skor yang di peroleh siswa sebagaimana adanya. Dengan kata lain ialah skor aktual sehingga nilai rata-rata dan simpangan baku yang diperolehnya adalah nilai rata-rata dan simpangan bakunya aktual.

Kriteria konversi nilai di atas berlaku juga untuk batas lulus ideal yang menggunakan nilai rata-rata dan simpangan baku ideal. Jika menggunakan batas lulus ideal, dari contoh di atas dapat dicari nilai rata-rata dan simpangan baku idealnya.

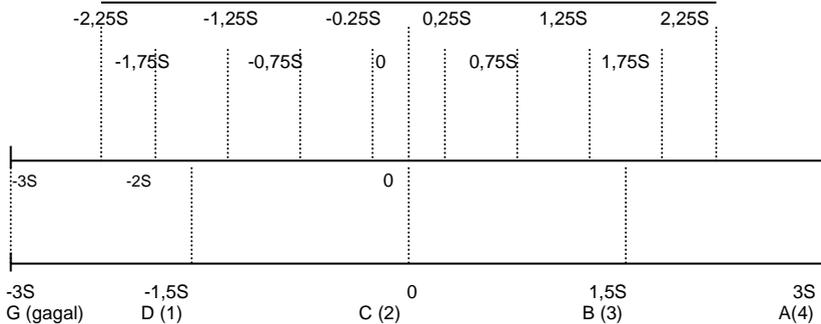
- Skor minimum yang mungkin dicapai adalah 90
- Nilai rata-rata idealnya adalah setengah dari 90 = 45
- Simpangan bakunya adalah sepertiga dari 45 = 15

Dengan menggunakan nilai rata-rata dan simpangan baku ideal tersebut, skor mentah yang ditransformasikan ke dalam standar sepuluh menjadi sebagai berikut:

Skor mentah	Standar 10
$45 + (2,25)(15) = 78,75$	10
$45 + (1,75)(15) = 68,75$	9
$45 + (1,25)(15) = 63,75$	8
$45 + (0,75)(15) = 53,75$	7
$45 + (0,25)(15) = 48,75$	6
$45 - (0,25)(15) = 41,25$	5
$45 - (0,75)(15) = 36,25$	4
$45 - (1,25)(15) = 26,25$	3
$45 - (1,75)(15) = 21,25$	2
$45 - (2,25)(15) = 11,25$	1

Konversi lainnya adalah konversi skor mentah ke dalam standar huruf (A-B-C-D) dan standar empat (4-3-2-1). Dalam standar ini huruf A setara dengan 4, artinya istimewa; huruf B setara dengan 3 artinya memuaskan; huruf C setara dengan 2, artinya cukup ; huruf D setara dengan 1, artinya kurang.

Kriteria yang digunakan pada dasarnya tidak berbeda dengan kriteria untuk konversi nilai ke dalam standar 10. perhatikan gambar pada halaman berikutnya.



Dalam gambar tersebut rentangan skala unit deviasi (SUD) menjadi empat kategori, yakni A(4), B(3), C(2) dan D (1), masing-masing dengan jarak 1,5 S, yakni diperoleh dari 6 skala unit deviasi dibagi empat kategori atau  $6/4 = 1,5$ .

Dengan demikian, secara sederhana untuk nilai C berada pada nilai rata-rata atau deviasi standar nol. Selanjutnya, nilai B ada pada  $M + 1,5 S$  nilai A adalah  $M + 3 S$  sedangkan nilai D adalah  $M - 1,5 S$

Mengingat kedudukan nilai A,B, C dan D adalah titik tengah, maka perlu dicari batas bawah (Bb) dan batas atas (Ba) untuk setiap nilai tersebut. Ukuran atau kriterianya adalah sebagai berikut:

Nilai	Batas bawah	Batas atas
D	$M - 1,5 S$	$M - 0,5 S$ (Bb.C)
C	$M - 0,5 S$	$M + 0,5 S$
B	Ba C	$M + 1,5 S$
A	Ba B	$M + 3,0 S$

Contoh:

Apabila berdasarkan perhitungan diperoleh nilai rata-rata ( $M$ ) = 40 dan simpangan baku ( $S$ ) = 10<sub>3</sub> maka konversi nilainya menjadi sbb:

Skor	Nilai	Keterangan:
25-35	D (1)	Batas bawah D = $40 - 1,5 (10)$ = 25
36 -45	C(2)	Batas atas D = $40 - 0,5 (10)$ = 35
46-55	B(3)	
56 – 70 keatas	A(4)	

Contoh diatas dengan memecahkan soal pada halaman berikutnya:

Contoh penggunaan:

Tes objektif bentuk pilihan ganda bilangan studi IPA diberikan kepada 20 orang siswa SMA kelas III. Jumlah soal sebanyak 45 pertanyaan. Setiap jawaban yang benar diberi skor 2 sehingga skor maksimal yang dapat di capai siswa adalah 90. setelah diperiksa, hasilnya adalah sebagai berikut:

69	59	49	52	60	37	45
55	68	41	48	51	56	30
43	53	62	40	46	51	

*Pertanyaan:*

1. Dengan menggunakan batas lulus aktual, tentukan batas lulusnya dan lakukan konversi nilai ke dalam standar 10, standar huruf dan standar 4.
2. Tentukan pula bila menggunakan batas lulus ideal.

Jawaban :

Dari data diketahui bahwa jumlah siswa ( $n$ ) adalah 20. Skor tertinggi adalah 69, dan skor terendah adalah 30 dari skor maksimum 90. Untuk dapat menjawab pertanyaan tersebut diperlukan menghitung nilai rata-rata ( $M$ ) dan simpangan baku ( $S$ ) aktual. Dengan menggunakan rumus nilai rata-rata dan simpangan baku untuk data yang dikelompokkan, dapat diketahui kedua harga tersebut. Untuk itu perlu membuat tabel distribusi frekuensi skor sebagai berikut:

Tabel 4.8. Distribusi skor

Kelompok Skor (Interval =5)	f	Deviasi (d)	fd	fd <sup>2</sup>
30 – 34	1	-4	-4	16
35 – 39	2	-3	-6	18
40 – 44	3	-2	-6	12
45 – 49	3	-1	-3	3
50 – 54	4	0	0	0
55 – 59	3	1	3	3
60 – 64	2	2	4	8
65 – 69	2	3	6	18
Jumlah	20		-6	78

Maka simpangan baku adalah: 
$$S = i \sqrt{\frac{\sum fd^2}{n} - \left(\frac{\sum fd}{n}\right)^2} =$$

$$= 5 \sqrt{\frac{78}{20} - \left(\frac{-6}{20}\right)^2} = 9,75$$

Rata-rata adalah 
$$\bar{X} = u + i \left(\sum \frac{f \cdot d}{n}\right)$$

$$= 52 + 5 \left(\frac{-6}{20}\right) = 52 - 1,5 = 50,5$$

Dari perhitungan tersebut diketahui bahwa simpangan baku 9,75 dibulatkan 10, dan rata-rata sebesar 50,5,

a) Konversi nilai dengan persen (secara sederhana).

Tanpa menghitung nilai rata-rata dan simpangan baku, skor diatas dapat ditransformasikan ke dalam standar 10 dan standar hurup sebagai berikut:

Tabel 4.9. Kriteria nilai konversi

Persentase %	Skor dicapai Maks. 69	Nilai Konversi		
		Huruf	Standar 10	Standar 4
90 – 99	63 – 69	A	9	4
80 – 89	56 – 62	B	8	3
70 – 79	49 – 55	C	7	2
60 – 69	42 – 48	D	6	1
Kurang dari 60 %	Kurang dari 42	Gagal Tidak ada	Gagal Nilai 10 (diatas 69)	Gagal

- b) Konversi nilai ke dalam standar 10 dengan menggunakan nilai rata-rata dari simpangan baku:  $M= 50,5$  dan  $S=10$

Skor mentah	Standar 10
$50,5 + (2,25)(10) = 73$	10
$50,5 + (1,75)(10) = 68$	9
$50,5 + (1,25)(10) = 63$	8
$50,5 + (0,75)(10) = 58$	7
$50,5 + (0,25)(10) = 53$	6(batas lulusnya)
$50,5 - (0,25)(10) = 48$	5
$50,5 - (0,75)(10) = 43$	4
$50,5 - (1,25)(10) = 38$	3
$50,5 - (1,75)(10) = 33$	2
$50,5 - (2,25)(10) = 28$	1

Dengan demikian, perolehan nilai dari 20 orang siswa adalah sebagai berikut:

69 = 9	56 = 6	51 = 5	46 = 4	40 = 3
68 = 9	55 = 6	50 = 5	45 = 4	39 = 3
62 = 7	53 = 6	49 = 5	43 = 4	37 = 2
60 = 7	52 = 5	48 = 5	41 = 3	30 = 1

Yang dinyatakan lulus dari nilai 6 keatas ada 8 orang.

c) Konversi ke dalam standar huruf dan standar 4

Dengan nilai rata-rata 50 dan simpangan baku = 9, maka nilai yang diperoleh sebagai berikut:

Skor Mentah	Skor dicapai Maks. 69	Nilai Konversi	
		Huruf	Standar 4
$50,5 - 1,5(10)$	$35,5 - 45,0$	D	1
$50,5 - 0,5(10)$	$45,5 - 55,0$	C	2
$50,5 + 0,5(10)$	$55,5 - 65,0$	B	3
$50,5 + 1,5(10)$	$65,5 - 75,0$ Ke atas	A	4

Dengan demikian, perolehan nilai dari 20 orang siswa adalah sebagai berikut:

69 = A	56 = B	51 = C	46 = C	40 = D
68 = A	55 = C	50 = C	45 = D	39 = D
62 = B	53 = C	49 = C	43 = D	37 = D
60 = B	52 = C	48 = C	41 = D	30 = G

### C. Pengolahan Data Hasil Non Tes

Telah dijelaskan bahwa hasil dan proses belajar tidak hanya bisa diukur dengan tes, tetapi juga dengan alat-alat ukur bukan tes seperti kuesioner, wawancara, observasi dan skala. Dalam mengolah data hasil dari nontes, disamping digunakan cara-cara seperti pada pengolahan data yang menggunakan tes (terutama jika datanya bersifat interval dalam bentuk skor nilai), juga dapat digunakan cara-cara lain seperti *persentase*, *modus*, *peringkat*, terutama bila hasil pengukuran menghasilkan data nominal atau ordinal. Berikut ini akan dijelaskan secara umum cara pengolahan tersebut.

Pada umumnya data hasil nontes bertujuan untuk mendeskripsikan hasil pengukuran sehingga dapat dilihat kecenderungan jawaban responden melalui alat ukur tsb. Misalnya bagaimana kecenderungan jawaban yang diperoleh dari wawancara, kuesioner, observasi, skala atau sosiometri.

### **1. Pengolahan data hasil wawancara dan kuesioner**

Dari data hasil wawancara dan atau kuesioner pada umumnya dicari frekuensi jawaban responden untuk setiap alternatif yang ada pada setiap soal. Frekuensi yang paling tinggi ditafsirkan sebagai kecenderungan jawaban alat ukur tersebut. Sebaliknya, frekuensi yang paling rendah dapat di tafsirkan sebagai kecenderungan jawaban yang tidak menggambarkan pendapat ke-banyakan responden. Sebagai contoh:

Melalui kuesioner ataupun wawancara diungkapkan pandangan siswa mengenai guru yang diharapkan dalam:

1. kemampuan mengajar
2. hubungan dengan siswa

Kuesioner atau wawancara diajukan kepada 40 orang siswa dengan pertanyaan sebagai berikut:

1. Guru yang saya harapkan adalah guru yang:
  - a. Menguasai bahan pelajaran atau pandai dalam bidang ilmunya.

- b. Cara menjelaskan bahannya dapat saya pahami sekalipun tidak begitu pandai.
  - c. Pandai dalam bidang ilmunya dan dapat menjelaskannya ke pada siswa dengan baik.
2. Pada waktu mengajarkan bahan pelajaran:
- a. Sebaiknya dimulai dari yang umum, kemudian dibahas secara khusus.
  - b. Sebaiknya dimulai dari yang khusus, kemudian menuju kepada yang umum.
  - c. Dimulai dari mana saya asal dijelaskan secara sistematis.
3. Menurut pendapat saya, hubungan guru dengan siswa di dalam kelas:
- a. Harus menjaga jarak agar tidak kehilangan wibawa.
  - b. Tidak perlu menjaga jarak asal dalam batas-batas pendidikan.
  - c. Mencerminkan hubungan orang tua dengan anak-anaknya.
4. Untuk membina hubungan guru dengan siswa, sebaiknya guru berusaha dalam:
- a. memahami pribadi para siswanya.
  - b. Melibatkan diri dalam berbagai kegiatan yang dilaksanakan oleh siswa.
  - c. Bergaul dengan siswa dalam berbagai kesempatan.

Kuesioner yang telah diisi oleh siswa kemudian diperiksa dan diolah dengan menghitung frekuensi jawaban seluruh siswa terhadap setiap pertanyaan tersebut.

Misalnya hasil pemeriksaan tersebut adalah sebagai berikut:

Tabel 4.10. Frekuensi jawaban siswa tentang masalah kemampuan guru mengajar (n=40)

Masalah yang diungkapkan	f	%	Peringkat jawaban
1. Kemampuan mengajar			
1.1. Kemampuan mengajar:			
a. Menguasai bahan	4	10	3
b. Mampu menjelaskan bahan pelajaran	12	30	2
c. Menguasai bahan dan mampu menjelaskannya	24	60	1
1.2. Prosedur mengajarkan bahan pelajaran:			
a. Dimulai dari yang umum	10	25	2
b. Dimulai dari yang khusus	6	15	3
c. Harus sistematis	24	60	1

Dari data pada tabel 4.10, dapat ditafsirkan bahwa kemampuan guru mengajar yang diharapkan oleh siswa adalah guru yang menguasai pelajaran dan dapat menjelaskannya agar dapat dipahami para siswa, sedangkan prosedur mengajar yang diharapkan oleh siswa harus sistematis. Cara lain dalam mengolah data di atas dengan menggunakan rumus Chi Kuadrat ( $\chi^2$ ).

$$\text{Rumus yang digunakan: } \chi^2 = \sum \left( \frac{fo - fe}{fe} \right)^2$$

Dalam Chi Kuadrat, yang dicari ialah adakah perbezaan yang berarti di antara frekuensi hasil pengamatan atau jawapan nyata ( $f_o$ ) dengan frekuensi jawapan yang diharapkan ( $f_e$ ). Jika ada perbezaan, artinya jawapan tersebut betul-betul adanya, bukan karena faktor kebetulan.

Contoh: Kita ambil jawapan nomor 1 dari Tabel 4.10.

Jawaban	$f_o$	$f_e$	$\frac{(f_o - f_e)^2}{f_e}$
a. Menguasai bahan	4	13,3	6,00
b. Mampu menjelaskan bahan pelajaran	12	13,3	0,13
c. Menguasai bahan dan mampu menjelaskannya	24	13,3	8,61
			$\chi^2 = 15,24$

Keterangan:

- $f_e$  : 13,3 diperoleh dari  $40/3 = 13,3$
- Harga  $\chi^2 = 12,54$  kemudian dibandingkan dengan harga chi kuadrat tabel untuk tingkat kepercayaan 0,05 (misalnya) dengan derajat bebas 3 - 1 (banyaknya alternatif jawapan 3)
- Harga  $\chi^2$  dalam tabel = 5,99

Dengan demikian,  $\chi^2=15,24 > 5,99$  sehingga perbezaan itu cukup berarti. Ini berarti bahwa interpretasi yang menyatakan bahwa guru yang diharapkan adalah guru yang menguasai bahan dan dapat menjelaskannya kepada siswa adalah sah sebagai kesimpulan dari data tersebut.

Tabel 4.11. Frekuensi jawaban siswa mengenai hubungan guru dan siswa

Masalah yang diungkapkan	f	%	Peringkat jawaban
1. Hubungan guru-siswa			
1.1. Sifat hubungan:			
a. Menjaga jarak	4	10	3
b. Tidak menjaga jarak	10	25	2
c. Hubungan orang tua-anak	26	65	1
1.2. Upaya membina hubungan:			
a. Memahami siswa	8	20	3
b. Turut serta dalam kegiatan siswa	12	35	2
c. Bergaul dengan siswa	20	50	1

Dari jawaban diatas dapat ditafsirkan bahwa hubungan guru siswa dalam kelas harus mencerminkan hubungan orang tua dengan anaknya, dan untuk itu guru dituntut bergaul dengan para siswa dalam berbagai kesempatan. Dari contoh diatas, data diolah dengan cara mencari persentase jawaban yang paling banyak atau modus jawaban siswa. Persentase dihitung dengan rumus  $f/N \times 100$ . N adalah banyaknya siswa yang menjawab pertanyaan, dalam hal ini 40.

## 2. Pengolahan data hasil observasi

Pengolahan data hasil observasi sangat bergantung pada pedoman observasinya, terutama dalam mencatat hasil observasi. Hasil observasi yang dinyatakan dalam bentuk pernyataan-pernyataan sebagaimana adanya yang tampak dari perilaku yang

diobservasi, diolah dengan melakukan analisis dan interpretasi seluruh hasil amatan tersebut. Dengan kata lain, dengan menggunakan analisis kualitatif. Sudah barang tentu sifatnya subjektif, yakni dipengaruhi oleh pengamatannya.

Namun, ada pula observasi yang hasil pengamatannya diberi nilai atau disediakan skala nilai, misalnya dengan huruf A, B, C, dan D atau dengan angka 4, 3, 2, dan 1 yang tersebut bermakna sebagai skala nilai.

Contoh :

**Observasi Kemampuan Guru dalam Mengajar**

Nama Guru: .....	Pendidikan: .....			
	Nilai Pengamatan			
Aspek yang diamati	4	3	2	1
1. Penguasaan bahan		v		
2. Kemampuan menjelaskan bahan	v			
3. Hubungan dengan siswa		v		
4. Penguasaan kelas	v			
5. Keaktifan belajar siswa		v		

Pengamat

Dari contoh di atas, skor hasil observasi adalah:

$$3 + 4 + 3 + 4 + 3 = 17.$$

Nilai rata-rata untuk kelima aspek tsb adalah  $17/5 = 3,4$ . skor ini cukup tinggi sebab maksimum rata-rata atau skor maksimum untuk setiap aspek adalah 4 atau 20 untuk semua aspek (5 x 4).

Skor ini bisa juga dikonversikan ke dalam bentuk standar 100 atau standar 10.

- Konversi ke dalam standar 100 adalah

$$\frac{17}{20} \times 100 = 85$$

- Konversi ke dalam standar 10 adalah  $\frac{17}{20} \times 10 = 8,5$

Jika dibuat interpretasi untuk setiap aspek, maka dapat disimpulkan bahwa guru tersebut sangat istimewa dalam hal kemampuan menjelaskan dan penguasaan kelas, sedangkan dalam penguasaan bahan, komunikasi dengan siswa dan dalam mengaktifkan siswa termasuk memuaskan.

### **3. Pengolahan data skala penilaian atau skala sikap**

Data hasil skala, baik skala penilaian maupun skala sikap yang berbentuk skor atau data interval, pengolahannya hampir sama dengan pengolahan data hasil observasi yang menggunakan skor atau nilai dalam pengamatannya. Dengan demikian, untuk setiap siswa yang diukur melalui skala penilaian atau skala sikap bisa ditentukan:

- a. perolehan skor dari seluruh butir ditentukan;
- b. skor rata-rata dari setiap pertanyaan dengan membagi jumlah skor oleh banyaknya pertanyaan,
- c. interpretasi terhadap pertanyaan mana yang positif atau baik dan pertanyaan atau aspek mana yang negatif atau kurang baik.

Lebih jauh lagi data hasil skala penilaian dan skala sikap sebenarnya menyerupai data hasil tes, yakni diperolehnya data interval dalam bentuk skor total untuk setiap siswa. Dengan demikian, dapat diolah seperti mengolah data hasil tes. Misalnya dicari nilai rata-rata atau simpangan baku, bergantung pada tujuan pengolahan data tersebut.

Uraian pengolahan data di atas terbatas pada hal-hal yang sederhana dengan maksud dapat dipraktekkan dalam tugas sehari-hari. Sudah tentu dalam pelaksanaannya diperhatikan ketentuan-ketentuan yang berlaku di sekolah, khususnya dalam sistem penilaian, baik dalam hal sistem pemberian angka maupun dalam menentukan batas kelulusannya. Hal ini perlu diingatkan karena dalam praktek di sekolah belum ada keseragaman.

# **BABV MENGANALISIS TES**

## **Menganalisis Tes**

### **A. Pendahuluan**

Menganalisis tes, merupakan upaya untuk mengetahui tingkat kebaikan tes yang akan digunakan. Dalam melakukan analisis tes sangat berkaitan dengan analisis butir soal. Analisis butir soal atau analisis item adalah pengkajian pertanyaan-pertanyaan tes agar diperoleh perangkat pertanyaan yang memiliki kualitas yang memadai. Ada dua jenis analisis butir soal, yakni analisis tingkat kesukaran soal dan analisis daya pembeda di samping validitas dan reliabilitas. Menganalisis tingkat kesukaran soal artinya mengkaji soal-soal tes dari segi kesulitannya sehingga dapat diperoleh soal-soal mana yang termasuk mudah, sedang, dan sukar. Sedangkan menganalisis daya pembeda artinya mengkaji soal-soal tes dari segi kesanggupan tes tersebut dalam membedakan siswa yang termasuk ke dalam kategori lemah atau rendah dan kategori kuat atau tinggi prestasinya. Sedangkan validitas dan reliabilitas mengkaji kesulitan dan kejelasan pertanyaan tes.

## **B. Analisis Tingkat Kesukaran**

Asumsi yang digunakan untuk memperoleh kualitas soal yang baik, di samping memenuhi validitas dan reliabilitas, adalah adanya keseimbangan dari tingkat kesulitan soal tersebut. Keseimbangan yang dimaksudkan adalah adanya soal-soal yang termasuk mudah, sedang dan sukar secara proporsional. Tingkat kesukaran soal dipandang dari kesanggupan atau kemampuan siswa dalam menjawab, bukan dilihat dari sudut guru sebagai pembuat soal. Persoalan yang penting dalam melakukan analisis tingkat kesukaran soal adalah penentuan proporsi dan kriteria soal yang termasuk mudah, sedang dan sukar.

Ada beberapa dasar pertimbangan dalam menentukan proporsi jumlah soal kategori mudah, sedang dan sukar. Pertimbangan pertama adalah adanya keseimbangan, yakni jumlah soal sama untuk ketiga kategori tersebut. Artinya soal mudah, sedang dan sukar jumlahnya seimbang. Misalnya tes objektif pilihan berganda dalam pelajaran matematika di susun sebanyak 60 pertanyaan. Dari ke-60 pertanyaan tersebut, soal kategori mudah sebanyak 20, kategori sedang 20 dan kategori sukar 20. Pertimbangan kedua proporsi jumlah soal untuk ketiga kategori tersebut didasarkan atas kurva normal. Artinya, sebagian besar

soal berada dalam kategori sedang, sebagian lagi termasuk ke dalam kategori mudah dan sukar dengan proporsi yang seimbang.

Perbandingan antara soal mudah-sedang-sukar bisa dibuat 3-4-3. artinya, 30% soal kategori mudah, 40% soal kategori sedang, dan 30% lagi soal soal kategori sukar. Misalnya, dari 60 pertanyaan pilihan ganda terdapat 18 soal kategori mudah, 24 soal kategori sedang, dan 18 soal kategori sukar. Perbandingan lain yang termasuk sejenis dengan proporsi di atas misalnya 3-5-2. artinya, 30% soal kategori mudah, 50% soal kategori sedang, dan 20% soal kategori sukar.

Persoalan lain adalah menentukan kriteria soal, yaitu ukuran untuk menentukan apakah soal tersebut termasuk mudah, sedang atau sukar. Dalam menentukan kriteria ini digunakan *judgment* dari guru berdasarkan per-timbangan-pertimbangan tertentu. Pertimbangan tersebut antara lain adalah:

- a. Abilitas yang diukur dalam pertanyaan tersebut. Misalnya untuk bidang kognitif, aspek pengetahuan atau ingatan dan pemahaman termasuk kategori mudah, aspek penerapan dan analitis termasuk kategori sedang, dan aspek sintesis dan evaluasi termasuk kategori sukar.

- b. Sifat materi yang diujikan atau ditanyakan. Misalnya ada fakta, konsep, prinsip dan hukum, serta generalisasi. Fakta termasuk ke dalam kategori mudah, konsep dan prinsip termasuk ke dalam kategori sedang, dan generalisasi (menarik kesimpulan) termasuk ke dalam kategori sukar.
- c. Isi bahan yang ditanyakan sesuai dengan bidang keilmuannya, baik luasnya maupun kedalamannya. Tentang persoalan ini bahan yang akan diujikan, guru sendiri harus sudah bisa menentukan mana yang termasuk mudah-sedang-sukar. Dengan kata lain, untuk menentukan kesulitan isi bahan, kewenangan ada pada guru itu sendiri.
- d. Bentuk soal. Misalnya dalam tes objektif, tipe soal pilihan benar-salah lebih mudah dari pada pilihan berganda dengan option tiga atau empat. Menjodohkan relatif lebih sulit dari pada pilihan berganda jika terdapat lima atau lebih yang harus dipasangkan.

Sesungguhnya demikian, keempat pertimbangan di atas tidak mutlak sebab bergantung pada isi bahan yang ditanyakan. Kadang-kadang soal benar-salah untuk aspek tertentu lebih sulit dari pada pilihan berganda untuk aspek lainnya. Demikian juga soal yang mengungkapkan kemampuan analisis dalam hal tertentu

lebih mudah dari pada soal yang mengungkapkan pemahaman. Dengan demikian, judgment ada pada guru yang bersangkutan setelah ia menentukan ruang lingkup materi yang akan diujikan, baik luas maupun kedalamannya.

Hal yang sama berlaku dalam menyusun tes uraian (esai). Artinya soal-soal jenis esai hendaknya memperhatikan pula tingkat kesukaran soal. Mengingat sifatnya, menentukan tingkat kesukaran soal tes uraian jauh lebih mudah dari pada tes objektif. Melalui analisis abilitas yang diukur serta isi dan sifat bahan yang ditanyakan, dalam tes uraian dapat dengan mudah menentukan tingkat kesukaran.

Setelah *judgment* dilakukan oleh guru, kemudian soal tersebut diuji-cobakan dan dianalisis apakah *judgment* tersebut sesuai atau tidak. Misalnya soal nomor 5 termasuk ke dalam kategori mudah, soal nomor 7 kategori sedang, dan nomor 9 kategori sukar. Setelah dilakukan uji coba, hasilnya dianalisis apakah nomor-nomor soal itu sesuai dengan judgment tersebut. Cara melakukan analisis untuk menentukan tingkat kesukaran soal adalah dengan menggunakan rumus sebagai berikut:

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

Keterangan:

I = indeks kesulitan untuk setiap butir soal

B = banyaknya siswa yang menjawab benar setiap butir soal

N = banyaknya siswa yang memberikan jawaban pada soal yang dimaksudkan

Kriteria yang digunakan adalah makin kecil indeks yang diperoleh, makin sulit soal tersebut. Sebaliknya, makin besar indeks yang diperoleh, makin mudah soal tersebut. Kriteria indeks kesulitan soal itu adalah sbb:

0 - 0,30 = soal kategori sukar

0,31 - 0,70 = soal kategori sedang

0,71 - 1,00 = soal kategori mudah

Contoh:

Guru IPS memberikan 10 pertanyaan pilihan berganda dengan komposisi 3 soal mudah, 4 soal sedang, dan 3 soal sukar. Jika dilukiskan, susunan soalnya adalah sebagai berikut:

No. Soal	Abilitas yang diukur	Tingkat kesukaran soal
1	Pengetahuan	Mudah
2	Aplikasi	Sedang
3	Pemahaman	Mudah
4	Analisis	Sedang
5	Evaluasi	Sukar
6	Sintesis	Sukar
7	Pemahaman	Mudah
8	Aplikasi	Sedang
9	Analisis	Sedang
10	Sintesis	Sukar

Kemudian soal tersebut diberikan kepada 20 orang siswa, dan tidak seorang pun yang tidak mengisi seluruh pertanyaan tersebut. Setelah diperiksa, hasilnya adalah sebagai berikut:

No. soal	Banyaknya siswa yang menjawab (N)	Banyaknya siswa yang menjawab betul (B)	Indeks $\frac{B}{N}$	Kategori soal
1	20	18	0,9	mudah
2	20	12	0,6	sedang
3	20	10	0,5	sedang
4	20	20	1,0	mudah
5	20	6	0,3	sukar
6	20	4	0,2	sukar
7	20	16	0,8	mudah
8	20	11	0,55	sedang
9	20	17	0,85	mudah
10	20	5	0,25	sukar

Dari sebaran di atas ternyata ada *tiga soal* yang meleset, yakni soal nomor 3 yang semula diproyeksikan

ke dalam kategori mudah, setelah dicoba ternyata termasuk ke dalam kategori sedang. Demikian, juga soal nomor 4 yang semula diproyeksikan sedang ternyata termasuk ke dalam kategori mudah. Nomor soal 9 semula diproyeksikan sedang, ternyata termasuk ke dalam kategori mudah. Sedangkan tujuh soal lainnya sesuai dengan proyeksi semula. Atas dasar tersebut, ketiga soal di atas harus diperbaiki kembali, adalah

- soal no. 3 diturunkan ke dalam kategori mudah,
- soal no. 4 dinaikkan ke dalam kategori sedang,
- soal no. 9 dinaikkan ke dalam kategori sedang.

Cara lain dalam melakukan analisis tingkat kesukaran soal adalah dengan menggunakan Tabel Rose dan Stanley.

Kriterianya yang digunakan seperti disajikan pada tabel berikut:

Persentase	Option				Kategori
	2	3	4	5	
16	0,16n	0,213n	0,24n	0,256n	Mudah
50	0,50n	0,667n	0,75n	0,800n	Sedang
84	0,84n	0,200n	1,26n	1,344n	Sukar

*Keterangan:*

- option 2 adalah bentuk benar-salah
- option 3, 4, dan 5 adalah bentuk pilihan berganda
- n adalah 27% dari banyaknya siswa yang mengikuti tes

Dalam menghitung indeks kesukaran soal, rumusnya adalah:

<b>SR + ST</b>
----------------

- SR adalah siswa yang menjawab salah dari kelompok rendah
- ST adalah siswa yang menjawab salah dari kelompok tinggi.

Mengingat  $n$  adalah 27%, maka siswa dari kelompok rendah maupun kelompok tinggi 27% dari banyaknya peserta. Penentuan siswa kelompok rendah dan tinggi dilakukan berdasarkan peringkat skor yang diperoleh dari tes tersebut. Ambillah 27% dari kelompok tinggi dan 27% dari kelompok rendah.

Berikut ini akan dijelaskan sebuah contoh.

Tes pilihan berganda dengan *option empat* diberikan kepada 28 siswa kelas III SMA. Jumlah soal sebanyak 15. Setelah diperiksa hasilnya diurutkan sebagai berikut:

No Siswa	Skor	Peringkat
1	12	1
2	11	2,5
3	11	2,5
4	10	4
5	9	5
6	8	7,5
7	8	7,5
8	8	7,5
9	8	7,5
10	7	10,5
11	7	10,5
12	6	14
13	6	14
14	6	14
15	6	14
16	6	14
17	5	19
18	5	19

Siswa kategori tinggi  
(27 % dari 20) = 8 orang

19	5	19
20	5	19
21	5	19
22	4	23
23	4	23
24	4	23
35	3	26
26	3	26
27	3	26
28	2	28

} Siswa kategori kurang atau lemah  
(27% dari 28 orang = 8 orang)

Dari hasil sebaran skor peringkat di atas, siswa yang termasuk ke dalam kategori tinggi diambil 27% dari 28 orang, yaitu 8 orang. Siswa tersebut adalah nomor 1 s.d. 8. Sedangkan siswa kategori rendah adalah no 21 s.d. 28.

Setelah hasil jawaban kategori siswa di atas diperiksa, hasilnya adalah sebagai berikut:

Untuk siswa kelompok rendah (8 orang siswa)		Untuk siswa kelompok tinggi (8 orang siswa)		Keterangan
No Soal	Jumlah siswa yang menjawab salah kelompok rendah (SR)	Jumlah siswa yang menjawab salah kelompok tinggi (ST)	SR + ST	
1	4	1	5	sedang
2	4	1	5	sedang
3	3	1	4	Mudah
4	5	2	7	Sedang
5	2	1	3	Mudah
6	4	2	6	Sedang
7	2	1	3	Mudah
8	5	1	6	Sedang
9	5	3	8	Sukar
10	4	2	6	Sedang
11	3	1	4	Mudah
12	3	1	4	Mudah
13	2	1	3	mudah
14	4	1	5	sedang
15	5	3	8	Sukar

Dari table Ros dan Stanley, untuk pilihan ganda dengan *option* 4, kriterianya adalah  $0,24n$  (soal mudah),  $0,73n$

(soal sedang), dan  $1,26n$  (soal sukar). Telah diketahui bahwa  $n=27\%$  dan  $N$ , yaitu  $\frac{27}{100} \times 28 = 8$ , sehingga susunan soal adalah:

- soal mudah kriterianya  $8 \times 0,24 = 1,92$
- soal sedang kriterianya  $8 \times 0,75 = 6$
- soal sukar kriterianya  $8 \times 1,26 = 10,08$
- 

Dengan interpolasi diperoleh kriteria (dibulatkan):

- 0 – 4,0 = mudah
- 5,0 – 7,0 = sedang
- 8,0 – ke atas = sukar

Indeks kesukaran butir ini dapat ditentukan dengan cara menentukan perbandingan antara jumlah subjek yang jawabannya betul dengan seluruh subjek yang menjawab butir. Bila suatu soal (butir) diberi skor lulus atau gagal maka pengukurannya diukur dengan persentase siswa berhasil menjawabnya. Azasnya berlawanan : yaitu makin mudah butir soal, makin tinggi indeks kesukaran



tersebut, hasilnya sama saja. Dengan demikian, tes yang tidak memiliki daya pembeda, tidak akan menghasilkan gambaran hasil yang sesuai dengan kemampuan siswa yang sebenarnya. Sungguh aneh bila anak panai tidak lulus, tetapi anak bodoh lulus dengan baik tanpa dilakukan manipulasi oleh si penilai atau di luar faktor kebetulan.

Cara yang biasa dilakukan dalam analisis daya pembeda adalah dengan menggunakan tabel atau kriteria dari Rose dan Stanley seperti dalam analisis tingkat kesukaran soal.

Rumusnya adalah:

<b>SR - ST</b>
----------------

Contoh:

No. Soal	Jumlah siswa yang menjawab salah kelompok rendah (SR)	Jumlah siswa yang menjawab salah kelompok Tinggi (ST)	SR - ST	Ket.
1	6	1	5	
2	6	1	5	
3	5	2	3	
4	6	1	5	
5	2	1	1	
6	5	1	4	
7	2	1	1	
8	7	1	6	
9	7	1	6	
10	4	2	2	
11	3	1	2	
12	6	1	2	
13	2	1	5	
14	6	1	1	
15	5	2	3	
N = 30 orang		N= 27% dari 30 = 8		

Kriteria yang digunakan dari Tabel Ross dan Stanley adalah sebagai berikut:

Jumlah testi (N)	n (27 % x N)	Option			
		2	3	4	5
28 – 31	8	4	5	5	5
32 – 35	9	5	5	5	5
36 - 38	10	5	5	5	5

dan seterusnya, lihat lampiran

Kriteria pengujian daya pembeda adalah sebagai berikut: Bila  $SR - ST$  sama atau lebih besar dari nilai tabel, artinya butir soal itu mempunyai daya pembeda.

Berdasarkan data di atas, batas pengujian adalah 5, yaitu yang pertama dalam tabel di atas dengan jumlah N (28-31),  $n=8$ , pada option A, maka dapat disimpulkan sebagai berikut:

No. Item	SR - ST	Batas Nilai Tabel	Keterangan
1	5	5	Diterima
2	5	5	Diterima
3	3	5	Ditolak
4	5	5	Diterima
5	1	5	Ditolak
6	4	5	Ditolak
7	1	5	Ditolak
8	6	5	Diterima
9	6	5	Diterima
10	2	5	Ditolak
No. Item	SR - ST	Batas Nilai Tabel	Keterangan
11	2	5	Ditolak
12	5	5	Diterima
13	1	5	Ditolak
14	5	5	Diterima
15	3	5	Ditolak

Dari kesimpulan di atas hanya soal no: 1, 2, 4, 8, 9, 12, dan 14 yang memenuhi daya pembeda, sedangkan soal nomor lainnya tidak memiliki daya pembeda.

Dari contoh di atas dapat disimpulkan bahwa cara menghitung daya pembeda adalah dengan menempuh langkah sebagai berikut:

- a) Memeriksa jawaban soal semua siswa peserta tes
- b) Membuat daftar peringkat hasil tes berdasarkan skor yang dicapainya
- c) Menentukan jumlah sampel sebanyak 27% dari jumlah peserta tes untuk kelompok siswa pandai (peringkat atas) dan 27% untuk kelompok siswa kurang (peringkat bawah).
- d) Melakukan analisis butir soal, yakni menghitung jumlah siswa yang menjawab salah dari semua nomor soal, baik pada kelompok pandai maupun pada kelompok kurang.
- e) Menghitung selisih jumlah siswa yang salah menjawab pada kelompok kurang dengan kelompok pandai (SR-ST).
- f) Membandingkan nilai selisih yang diperoleh dengan nilai Tabel Ross & Stanley.
- g) Menentukan ada tidaknya daya pembeda pada setiap nomor soal dengan kriteria “memiliki daya pembeda” bila nilai selisih jumlah siswa yang

menjawab salah antara kelompok kurang dengan kelompok pandai (SR-ST) *sama atau lebih besar dari* nilai tabel.

Butir soal yang tidak memiliki daya pembeda diduga terlalu mudah atau terlalu sukar sehingga perlu diperbaiki atau diganti dengan pertanyaan lain. Idealnya semua butir soal memiliki daya pembeda dan tingkat kesukaran. Tes yang telah dibakukan, di samping memenuhi validitas dan reliabilitas, juga memenuhi tingkat kesukaran dan daya pembeda.

## **Validitas dan Reliabilitas Tes**

### **A. Analisis Validitas**

Validitas tes dibedakan antara (a) validitas isi, (b) validitas bangun pengertian, (c) validitas ramalan, dan (d) validitas kesamaan. Dua validitas yang disebut pertama, yakni validitas isi dan bangun pengertian, dapat dibuat melalui upaya penyusunan tes tanpa harus dilakukan pengujian statistik. Sedangkan untuk validitas kesamaan dan validitas ramalan dilakukan pengujian statistika

melalui uji korelasi. Validitas ramalan pada hakikatnya ingin melihat apakah tes sebagai alat penilaian dapat digunakan untuk meramalkan suatu ciri atau perilaku tertentu. Umpamanya apakah tes prestasi belajar tertentu dapat digunakan untuk meramalkan keterampilan tertentu jika misalnya terdapat korelasi positif yang berarti di antara hasil tes prestasi belajar tertentu dapat digunakan untuk meramalkan keterampilan tertentu jika misalnya terdapat korelasi positif yang berarti diantara hasil tes prestasi belajar dengan keterampilan memecahkan masalah. Sedangkan validitas kesamaan dilakukan dengan mengorelasikan tes yang dibuat oleh guru dengan tes yang sudah baku dalam bidang studi yang sama dan untuk tingkat yang sama pula. Dengan demikian, kedua validitas ini dihadapkan pada pengujian statistika yakni korelasi.

Ada dua jenis korelasi yang biasa digunakan, yakni korelasi momen produk (product moment) atau metode Pearson yang diberi notasi “r” dan korelasi tata jenjang (*rank correlation*) atau metode Spearman yang diberi notasi “rho”. Rumus yang digunakan adalah sbb:

$$r = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\left[ \sum X^2 - \frac{(\sum X)^2}{N} \right] \left[ \sum Y^2 - \frac{(\sum Y)^2}{N} \right]}$$

$$rho = 1 - \frac{6 \sum D^2}{N(N^2 - 1)}$$

Berikut ini akan dijelaskan sebuah contoh.

Misalkan guru IPA membuat soal untuk tes IPA bagi siswa kelas III dalam bentuk pilihan berganda. Ia bermaksud menguji apakah tes yang dibuatnya mempunyai validitas kesamaan atau tidak. Kemudian ia mengorelasikan hasil tes yang dibuatnya dengan soal-soal EBTA bidang studi IPA.

Setelah kedua tes tersebut diujikan kepada siswa kelas III yang sama, kemudian dicari indeks korelasi dari kedua hasil tes tersebut. Misalkan hasil tes buatan guru diberi notasi X dan datanya sbb:

Nama siswa : A B C D E F G H I Y K L M N O

Skor : 14 18 18 17 19 19 19 20 22 23 24 27 24 18 28

Sedangkan hasil tes Ebta bidang studi yang sama diberi notasi Y dan datanya untuk siswa yang sama adalah sbb :

Nama Siswa : A B C D E F G H I Y K L M N O

Skor : 12 15 17 15 16 17 18 18 20 20 22 24 24 18 26

Pertanyaan: Apakah kedua tes yang dibuat guru (X) mempunyai validitas kesamaan? Dengan perkataan lain apakah tes X berkorelasi positif dengan tes Y.

Untuk menjawab pertanyaan tersebut perlu dicari koefisien korelasi XY dengan menggunakan salah satu rumus yang telah dijelaskan di atas. Dengan

menggunakan korelasi momen produk dilakukan langkah perhitungan sebagai berikut:

Tabel 5.1. Distribusi Skor

Siswa	Skor X	Skore Y	XY	X <sup>2</sup>	Y <sup>2</sup>
A	14	12	168	196	144
B	18	15	270	324	225
C	18	17	306	324	289
D	17	17	289	289	289
E	18	16	288	324	256
F	19	17	323	361	289
G	19	18	342	361	324
H	20	18	360	400	324
I	22	20	440	484	400
J	23	20	460	529	400
K	24	22	528	576	484
L	27	24	648	729	576
M	24	24	576	576	576
N	18	18	324	324	324
O	28	26	728	784	676
Jumlah	309	284	6050	6581	5576

$$\begin{aligned}
 r &= \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\left[ \sum X^2 - \frac{(\sum X)^2}{N} \right] \left[ \sum Y^2 - \frac{(\sum Y)^2}{N} \right]} = \frac{6050 - \frac{(309)(284)}{15}}{\left[ 6581 - \frac{(309)^2}{15} \right] \left[ 5576 - \frac{(284)^2}{15} \right]} \\
 &= \frac{6050 - 5850,4}{(6581 - 6365,4)(5576 - 5377,06)} \\
 &= \frac{199,6}{(215,6)(198,94)} = 0,96
 \end{aligned}$$

Dengan demikian, koefisien korelasi adalah 0,96, suatu indeks korelasi yang termasuk tinggi sehingga tes X yang dibuat oleh guru memiliki validitas kesamaan. Koefisien

korelasi dikatakan baik atau tinggi atau sempurna bila mendakati harga satu.

Selanjutnya apakah koefisien korelasi 0,96 berarti atau tidak harus diuji melalui uji t pada signifikansi tertentu dengan derajat kebebasan n-2. Rumusnya adalah:

$$t = \sqrt{n-2} \cdot \frac{r}{\sqrt{1-r^2}}$$

dalam contoh di atas, maka: n = 15

$$r = 0,96$$

$$t = \sqrt{15-2} \cdot \frac{0,96}{\sqrt{1-0,96^2}}$$

$$\sqrt{13} \cdot \frac{0,96}{0,28} = 12,33$$

Nilai t, tersebut jika dibandingkan dengan harga t tabel distribusi t untuk tingkat kepercayaan 0,05 (misalnya) dan derajat kebebasan (15-2). Harga tabel t adalah 2,60. Ternyata harga t hitung lebih besar dari harga t tabel (12,33 > 2,60). Dengan demikian korelasi tersebut berarti. Dalam korelasi yang menggunakan metode Spearman atau korelasi tata jenjang, skor-skor yang akan dikorelasikan terlebih dahulu dibuat berperingkat dari peringkat pertama sampai peringkat terakhir melalui pengurutan skor tertinggi sampai terendah. Dengan

menggunakan data yang sama dengan di atas dapat dibuat peringkat sebagai berikut:

Tabel 5.2. Peringkat Skor

Skor X	Peringkat	Skor Y	Peringkat
28	1	26	1
27	2	24	2,5 $\frac{2+3}{2}$
24	3,5 $\frac{3+4}{2}$	24	2,5 $\frac{2+3}{2}$
24	3,5 $\frac{3+4}{2}$	22	4
23	5	20	5,5 $\frac{5+6}{2}$
22	6	20	5,5 $\frac{5+6}{2}$
20	7	18	8
19	8,5 $\frac{8+9}{2}$	18	8 $\frac{7+8+9}{3}$
19	8,5 $\frac{8+9}{2}$	18	8 $\frac{7+8+9}{3}$
18	11,5	17	11
18	11,5	17	11
	$\frac{10+11+12+13}{4}$		$\frac{10+11+12}{3}$
18	11,5 $\frac{10+11+12+13}{4}$	17	11 $\frac{10+11+12}{3}$
18	11,5	16	13
17	14	15	14
14	15	12	15

Urutan peringkat skor diatas kemudian dimasukkan kedalam tabel seperti pada contoh berikut:

Tabel 5.3. Distribusi Skor

Skor X	Peringkat (RX)	Skor Y	Peringkat (RX)	D (Rx - Ry)	D <sup>2</sup>
14	15	12	15	0	0
18	11,5	15	14	-2,5	6,25
18	11,5	17	11	0,5	0,25
17	14	17	11	3	9
18	11,5	16	13	-1,5	2,25
19	8,5	17	11	-2,5	6,25
19	8,5	18	8	0,5	0,25
20	7	18	8	-1	1
22	6	18	8,5	0,5	0,25
23	5	20	5,5	-0,5	0,25
Skor X	Peringkat (RX)	Skor Y	Peringkat (RX)	D (Rx - Ry)	D <sup>2</sup>
24	3,5	20	4	-0,5	0,25
27	2	22	2,5	-0,5	0,25
24	3,5	24	2,5	1	1
18	11,5	18	8	3,5	12,25
28	1	26	1	0	0
Jumlah					39,5

$$D^2 = 39,5$$

$$N = 15$$

$$rho = 1 - \frac{6 \sum D^2}{N(N^2 - 1)} = 1 - \frac{6 \cdot (39,5)}{15(224)} = 1 - \frac{237}{3360} = 1 - 0,07 = 0,93$$

Berdasarkan perhitungan di atas diperoleh indeks korelasi 0,93. Suatu indeks korelasi yang sangat tinggi, sehingga dapat dikatakan ada korelasi yang sempurna. Ini berarti bahwa tes yang dibuat memiliki validitas kesamaan terhadap tes sejenis yang dianggap baku.

Indeks korelasi yang diperoleh dari kedua cara tersebut relatif sama, yaitu 0,96 dan 0,93. Perbedaan 0,03 disebabkan oleh pembulatan angka, terutama dalam perhitungan korelasi product moment.

## **B. Analisis Reliabilitas**

Reliabilitas (reliability, kepercayaan) menunjuk pada pengertian apakah sebuah tes dapat mengukur sesuatu yang diukur secara konsisten dari waktu ke waktu. Jadi, kata kunci untuk syarat kualifikasi suatu tes adalah konsistensi, keajegan, atau tidak berubah-ubah. Misalnya, alat ukur yang berupa alat penimbang dengan satuan berat gr (gram), ons, dan kg (kilogram) dapat dipergunakan secara konsisten untuk mengukur satuan

berat sesuatu oleh siapa pun dan kapanpun. Jadi alat ukur tersebut dapat dinyatakan sebagai alat pengukur yang reliabel. Sebaliknya, jika untuk mengukur satuan berat dengan hanya dijinjing dengan tangan, misalnya dengan membandingkan antara jinjingan pertama, kedua dan seterusnya, hasil satuan berat yang diperoleh tidak dapat dipertanggungjawabkan karena jinjingan itu sendiri tidak dapat dirasakan secara konsisten dan selisih berat yang relatif kecil. Jadi, alat pengukuran berat suatu benda yang diperoleh dengan cara menjinjing tidak dapat dinyatakan reliabel. Data-data hasil penelitian yang akan diolah dengan suatu teknik statistik yang diperoleh lewat pengukuran, instrumen pengukuran yang dipergunakan harus memenuhi syarat reliabilitas.

Analisis reliabilitas suatu tes dan atau alat ukur lainnya, termasuk nontes, pada hakikatnya menguji keajegan pertanyaan tes apabila diberikan berulang kali pada objek yang sama. Suatu tes dikatakan reliabel atau ajeg apabila beberapa kali pengujian menunjukkan hasil yang relatif sama. Pengujian suatu tes bisa dilakukan terhadap objek yang sama pada waktu yang berlainan dengan selang waktu yang tidak terlalu lama dan juga terlalu singkat, bisa juga dilakukan dengan membandingkan hasil pengujian dari tes yang setara.

Beberapa cara menguji reliabilitas alat ukur, antara lain tes ulang, pecahan setara, belah dua, dan

kesamaan rasional. Analisis reliabilitas suatu tes dan atau alat ukur lainnya pada umumnya menggunakan teknik korelasi seperti dalam analisis validitas. Dengan demikian, teknik – teknik perhitungannya sama dengan apa yang telah dijelaskan pada contoh-contoh menghitung indeks korelasi validitas pada butir tiga di atas. Perbedaannya terletak dalam hal tes atau alat ukur yang dikorelasikannya. Misalnya untuk jenis tes ulang, suatu tes diberikan dua kali kepada objek (siswa) yang sama. Hasil tes yang pertama ( $X_1$ ) kemudian dikorelasikan dengan tes yang kedua ( $X_2$ ). Berbeda dengan teknik belah dua, yaitu dua perangkat tes yang setara atau sederajat, baik isinya maupun tingkat kesukarannya, kemudian diberikan kepada objek yang sama. Hasil tes perangkat pertama kemudian dikorelasikan dengan hasil tes perangkat kedua.

Contoh:

Data dalam tabel 5.4, disajikan data tentang hasil analisis terhadap sebuah alat ukur tes hasil belajar, yang terdiri dari 20 butir soal, dan jumlah peserta tes sebanyak 16 orang.



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

$$= \frac{16(740) - (105)(110)}{\sqrt{[(16)(727) - (105)^2][(16)(786) - (110)^2]}} = \frac{290}{\sqrt{(607)(476)}} = 0,539$$

Hasil penghitungan koefisien korelasi ( $r_{gg}$ ) antarparuh soal tersebut kemudian dimasukkan ke dalam rumus Spearman-Brown untuk meramalkan besarnya koefisien reliabilitas untuk seluruh soal (N soal) dengan mempergunakan rumus:

$$r_{ns} = \frac{2r_{gg}}{1 + r_{gg}} = \frac{2(0,539)}{1 + 0,539} = 0,7$$

Untuk menafsirkan arti koefisien tersebut kita perlu mengkonsultasikannya dengan tabel nilai-nilai kritis r (Tabel Lampiran 2). Tabel nilai-nilai kritis r dengan N =16 menunjukkan bahwa r pada taraf signifikansi 5% adalah  $\geq 0,497$ , sedang taraf signifikansi 1%  $\geq 0.623$ . jadi, korelasi N soal di atas signifikan, dan karenanya instrumen penelitian yang sedang diuji reliabilitasnya tersebut dinyatakan reliabel.

## **Rangkuman**

Analisis tes soal bertujuan untuk memperoleh kualitas soal yang baik sehingga dapat diketahui gambaran tentang perstasi siswa yang sebenarnya. Ada beberapa cara melakukan analisis tes melalui analisis butir soal, yaitu analisis tingkat kesukaran, analisis daya pembeda, analisis validitas, dan analisis reliabilitas. Analisis tingkat kesukaran soal bertujuan untuk dapat membedakan soal-soal dalam kategori mudah, sedang, dan sukar. Analisis daya pembeda mengkaji apakah soal tersebut punya kemampuan dalam membedakan siswa termasuk kedalam kategori memiliki kemampuan tinggi dan kemampuan rendah. Dengan demikian, soal yang memiliki daya pembeda, jika diberikan kepada siswa berkemampuan tinggi, hasilnya menunjukkan lebih tinggi daripada jika diberikan pada siswa berkemampuan rendah. Analisis validitas bertujuan untuk mengkaji kesahihan alat ukur atau soal dalam menilai apa yang seharusnya diukur atau untuk mengkaji ketepatan soal tes sebagai alat ukur. Sedangkan reliabilitas mengkaji keajegan atau ketepatan hasil tes manakala tes tersebut diujikan kepada siswa yang sama lebih dari satu kali, adat dari dua perangkat tes yang setara kepada objek yang sama. Analisis butir soal dengan keempat cara di atas menggunakan pendekatan teknik statistik.

# BAB VI

## PELAPORAN DAN PEMANFAATAN HASIL PENILAIAN

### Pelaporan Hasil Penilaian

#### A. Pendahuluan

Penilaian merupakan suatu kewajiban bagi setiap guru atau dosen terhadap terhadap hasil belajar para siswa dan mahasiswanya. Hal tersebut dilakukan karena menilai hasil belajar siswa menjadi bagian integral dari tugasnya sebagai pengajar. Penilaian mungkin dilakukan sebelum pengajaran dimulai (*prates*), pada saat berlangsungnya proses belajar mengajar untuk mengetahui pemahaman siswa, atau yang jelas dilakukan pada akhir pengajaran (*pascates*). Pelaksanaan penilaian mungkin berbeda untuk penilaian proses belajar mengajar, mungkin tidak semua pengajar melakukannya karena mereka tidak menaruh perhatian terhadap proses belajar mengajar, atau tidak merasa perlu melakukannya, atau mungkin juga mereka tidak dapat melakukannya.

Penilaian hasil belajar yang dilakukan, baik penilaian formatif maupun penilaian sumatif, sangat bervariasi pelaksanaannya. Ada guru atau dosen yang sengaja mempersiapkannya dengan baik dalam hal menentukan apa yang seharusnya dinilai, bagaimana penilaian itu harus dilakukan dan tindakan apa yang harus dilakukan setelah penilaian itu dilaksanakan. Namun, ada pula guru dan dosen yang melaksanakan penilaian tersebut semata-mata untuk memenuhi kelengkapan tugas mengajarnya, bahkan tak peduli apapun hasil tindakan penilaian yang dilaksanakannya. Pengajar profesional yang memandang tugasnya sebagai keahlian khusus yang tidak dimiliki oleh profesi lain, hasil penilaian yang dilakukannya akan menjadi menjadi alat uji bagi keberhasilan dirinya sebagai pengajar sehingga senantiasa dimanfaatkan untuk perbaikan dan penyempurnaan tugas-tugas profesinya. Pengajar akan merasa selalu tidak puas dengan hasil belajar yang dicapai para siswa sehingga hasil penilaian itu selalu dikaji untuk mencari usaha dan cara baru dalam tindakan mengajarnya agar diperoleh hasil belajar siswa yang lebih baik. Kondisi inilah yang diduga belum sepenuhnya dihayati oleh para guru di sekolah sehingga

tidak mengherankan apabila tugas mengajar cenderung bersifat rutin.

Kegiatan mengajar sebaiknya dimulai dari hasil penilaian sebelumnya, artinya dimulai dari apa yang telah dicapai siswa, bukan dari apa yang seharusnya dipelajari siswa. Ini berarti bahwa pengajar harus memanfaatkan hasil-hasil penilaian yang telah dilakukannya. Ia perlu melakukan kajian terhadap hasil belajar siswa yang diperoleh melalui penilaian. Kajian terutama dilakukan untuk mengetahui apa yang telah dan yang belum dapat dicapainya, dalam hal apa atau bagian mana dari program belajar yang belum dapat dikuasainya, mengapa hal itu belum dikuasainya. Hasil kajian di atas sangat bermanfaat bagi guru atau dosen sebagai bahan untuk menentukan dari mana harus memulai mengajar pada program berikutnya. Ada beberapa pertanyaan yang dapat diajukan antara lain:

- a) Berapa banyak siswa yang pencapaian hasil belajarnya termasuk tinggi sedang dan rendah? Hal ini akan memberi petunjuk secara umum tentang keberhasilan pengajaran.
- b) Siswa dengan karakteristik yang bagaimana yang menunjukkan pencapaian hasil belajar yang termasuk tinggi-sedang-rendah? Hal ini akan

memberi petunjuk bagi upaya memahami kemampuan para siswa.

- c) Mengapa siswa-siswa tersebut mencapai hasil belajar yang rendah? Jawaban terhadap pertanyaan ini memberi petunjuk kepada guru dalam melakukan diagnosis kesulitan belajar.
- d) Apakah tinggi-rendahnya pencapaian hasil belajar siswa berkaitan dengan tindakan guru atau dosen dalam hal mengajar? Jawaban terhadap pertanyaan ini menjadi dasar untuk menilai kemampuan guru dalam melaksanakan tugas mengajarnya.
- e) Aspek manakah yang dirasakan masih lemah atau kurang meyakinkan dari pengalaman mengajar sebelumnya? Jawaban terhadap pertanyaan ini memberi petunjuk tentang upaya perbaikan dan penyempurnaan tugasnya sebagai pengajar.

Pertanyaan-pertanyaan tersebut, dan mungkin masih banyak lagi pertanyaan lain yang senada, akan memberikan informasi yang berharga untuk dilaporkan dan dimanfaatkan dalam upaya membina dan mengembangkan penyelenggaraan pendidikan sehingga dari waktu ke waktu menunjukkan perkembangan dan kemajuan hasil-hasil pendidikan yang dicapainya.

## **B. Pelaporan Data Hasil Penilaian**

Pelaporan data hasil penilaian diperlukan bukan semata-mata untuk kepentingan diri guru yang bersangkutan, melainkan juga harus di dimanfaatkan oleh semua pihak yang terlibat dalam penyelenggaraan pendidikan di sekolah. Data hasil penilaian yang dilakukan dan ada pada guru perlu dilaporkan agar dapat dimanfaatkan bagi kepentingan pendidikan. Melalui laporan hasil penilaian tersebut, semua pihak dapat mengetahui kemampuan dan perkembangan siswa, sekaligus dapat mengetahui tingkat keberhasilan pendidikan di sekolahnya. Atas dasar itu pula semua pihak dapat menentukan langkah dan upaya yang harus dilakukan dalam meningkatkan kualitas proses dan hasil pendidikan di sekolahnya.

Sehubungan dengan pelaporan data hasil penilaian tersebut timbul pertanyaan:

- Apa yang seharusnya dilaporkan?
- Kepada siswa laporan itu diberikan?
- Bagaimana cara menyusun laporan tersebut?

Dalam sebuah laporan data hasil penilaian bukan hanya mengenai prestasi atau hasil belajar, melainkan juga mengenai kemajuan dan perkembangan belajar siswa di sekolah seperti motivasi belajar, disiplin, kesulitan belajar, atau sikap siswa terhadap mata

pelajaran. Maka guru perlu mencatat perkembangan dan kemajuan belajar siswa secara teratur dan berkelanjutan.

Hasil belajar yang dicapai siswa itu hendaknya dilaporkan secara menyeluruh, baik sebagai data mentah berupa skor-skor yang diperoleh siswa maupun sebagai data masak yang telah diolah dalam bentuk nilai-nilai siswa sesuai dengan ketentuan yang berlaku di sekolah, misalnya nilai dalam standar huruf atau nilai dalam standar angka. Lebih jauh lagi dilakukan interpretasi terhadap nilai yang diperoleh siswa, misalnya kedudukan siswa dibandingkan dengan kelompoknya atau posisi siswa dibandingkan dengan kriteria yang telah ditentukan. Dengan demikian dapat diketahui tingkat keberhasilan siswa, baik dilihat dari kelompoknya maupun dari tujuan yang harus dicapainya. Sedangkan data mengenai perkembangan belajar siswa di sekolah dilaporkan dalam bentuk catatan-catatan khusus sebagai pelengkap data hasil belajarnya. Catatan khusus ini berkenaan dengan aspek-aspek perilaku siswa seperti kehadiran, disiplin belajar, motivasi belajar dan kesulitan belajar.

Data hasil penilaian dilaporkan kepada semua staf sekolah, yaitu kepada kepala sekolah, wali kelas, guru pembimbing, dan jika dipandang perlu kepada guru-guru lainnya. Kepada kepala sekolah dilaporkan prestasi atau

hasil belajar para siswa dalam bidang studi atau mata pelajaran yang dipegangnya, termasuk perkembangan belajar siswa selama mengikuti pendidikan di sekolah, khususnya dalam proses belajar mengajar. Hasil belajar siswa disampaikan dalam bentuk yang ringkas, tetapi cukup jelas sehingga dapat dengan mudah dipahami kepala sekolah. Melalui laporan tersebut kepala sekolah dapat menangkap maknanya sehingga ia mengetahui tingkat keberhasilan siswa dalam bidang studi tersebut.

Sebuah contoh bentuk laporan guru kepada kepala sekolah mengenai prestasi belajar siswanya. Kategori prestasi dibuat dalam tiga kategori, yaitu di atas rata-rata kelas, di sekitar rata-rata kelas, dan dibawah rata-rata kelas.

Bidang Studi : .....  
 Kelas/Semester : .....  
 Jumlah Siswa : .....

Kategori prestasi yang dicapai siswa			Banyaknya siswa yang mencapai nilai			
No	Skor	Nilai Standar	Rata-rata tes formatif		Hasil tes sumatif	
			n	%	n	%
1	-	-	-	-	-	-
2	-	-	-	-	-	-
...	....	....	....	....	....	....

Keterangan:

- n=banyaknya siswa
- Nilai standar = hurup (A, B, C, D, G) atau standar sepuluh (6, 7, 8, 9, 10)

.....,19....  
 Guru Bidang Studi

.....  
 NIP.....

Dengan laporan tersebut kepala sekolah dapat dengan mudah mengetahui keberhasilan siswa dalam bidang studi yang bersangkutan. Banyak sedikitnya siswa yang mencapai prestasi tinggi di kelasnya dapat dijadikan ukuran keberhasilan dari pengajaran bidang studi tersebut. Data hasil penilaian yang berkenaan dengan kegiatan belajar siswa disajikan tersendiri dalam bentuk perilaku atau kasus-kasus tertentu yang dianggap paling menonjol dan ada hubungannya dengan keberhasilan belajar siswa di sekolah. Berikut ini adalah contohnya laporannya yang perlu disampaikan oleh guru.

Bidang Studi : .....  
 Kelas/Semester : .....  
 Jumlah Siswa : .....

Kategori kasus siswa	Banyaknya	%	Keterangan
1. Sering membolos dalam mengikuti pelajaran			
2. Lambat belajar			
3. Kurang motivasi dan disiplin			
4. Tidak bisa menyesuaikan diri			
5. Mengalami kesulitan belajar			

Selain data tentang siswa yang mengalami kasus seperti di atas, sebaiknya ada pula laporan yang berisi siswa-siswa yang menunjukkan kelebihan-kelebihan dalam bidang tertentu, misalnya siswa yang berbakat, kreatif dan berprestasi tinggi. Dengan adanya laporan tersebut, kepala sekolah dapat menentukan berbagai upaya dalam meningkatkan mutu pendidikan di sekolahnya.

Laporan hasil penilaian kepada wali kelas berupa nilai masak untuk digunakan dalam mengisi nilai raport. Oleh sebab itu, harus lengkap untuk setiap siswa. Nilai hasil belajar yang dilaporkan sudah mempertimbangkan hasil tes formatif dan tes sumatif, termasuk catatan khusus yang dibuat oleh guru mengenai kemajuan belajar siswa yang ditunjukkan selama menempuh pengalaman belajarnya.

Guru pembimbing memerlukan laporan khusus dari setiap guru bidang studi atau mata pelajaran tentang kasus-kasus siswanya, terutama dalam kegiatan belajar dan kehidupan pribadi maupun kehidupan sosial siswa di sekolahnya. Oleh sebab itu, laporan dari guru sangat diperlukan dalam hal kesulitan belajar, motivasi dan disiplin belajar, penyesuaian diri, kasus-kasus kenakalan siswa, kehidupan pribadinya di samping prestasi belajar yang dicapainya. Data ini sangat diperlukan oleh guru pembimbing sebagai bahan dalam melaksanakan tugasnya, yakni memberi bantuan kepada siswa agar siswa dapat memecahkan masalah yang dihadapinya. Data yang dilaporkan oleh guru adalah catatan-catatan kasus siswa, baik secara perseorangan maupun secara kelompok. Dalam laporan tersebut dikemukakan nama siswa, latar belakang keluarganya, identitas dirinya, jenis

kasus yang dihadapinya, serta prestasi belajar yang dicapainya.

Laporan data hasil penilaian setidak-tidaknya disampaikan kepada kepala sekolah dan pihak lainnya, yakni wali kelas dan guru pembimbing, pada akhir caturwulan atau akhir semester, baik tertulis maupun lisan, dalam rapat guru yang diselenggarakan oleh kepala sekolah. Sungguhpun demikian, guru dapat saja melaporkan setiap saat apabila dipandang perlu, terutama kepada wali kelas dan guru pembimbing agar kasus-kasus siswa dapat segera diatasi sehingga memacu siswa untuk belajar lebih baik dan mencapai hasil yang lebih tinggi lagi. Ini berarti bahwa laporan data hasil penilaian guru bidang studi atau mata pelajaran kepada pihak lain di sekolah bukan sekadar memberi informasi tentang keadaan para siswanya, tetapi juga menjadi bagian penting dari guru dalam rangka meningkatkan tugas profesinya sebagai pengajar yang profesional. Dari saran dan pendapat para wali kelas ataupun guru pembimbing, guru dapat menentukan langkah selanjutnya dalam memperbaiki proses belajar mengajar, menyesuaikan strategi mengajarnya dengan kondisi dan masalah yang dihadapi siswanya, lebih memahami keadaan siswanya dan upaya lainnya untuk meningkatkan proses dan hasil belajar siswa di sekolah.

Oleh sebab itu, laporan penilaian hasil belajar dari guru bidang studi atau mata pelajaran kepada staf sekolah lainnya merupakan salah satu alat dalam memecahkan persoalan belajar para siswa dalam rangka meningkatkan kualitas hasil pendidikan di sekolah. Semakin sering tukar informasi hasil penilaian para guru dalam bidang studinya masing-masing, semakin baik mengingat diperolehnya data mengenai pribadi siswa, khususnya kemajuan belajarnya sebagai bahan bagi guru dalam melaksanakan proses belajar dan mengajar.

Hal lain yang menyangkut laporan data hasil penilaian yang dibuat oleh guru adalah dokumentasi data hasil penilaian. Dokumentasi ini hendaknya dibuat secara teratur dan sistematis oleh kepala sekolah sehingga mudah dipelajari dan dikaji manakala diperlukan.

### **C. Pemanfaatan data hasil penilaian**

Salah satu upaya untuk meningkatkan kualitas hasil pendidikan dapat dilakukan melalui pemanfaatan data hasil penilaian. Hasil penilaian, baik melalui tes maupun bukan tes, besar sekali manfaatnya bila dikaji dan digunakan untuk upaya perbaikan proses belajar mengajar. Kajian hasil penilaian formatif dan sumatif dapat memberikan gambaran tentang hasil belajar yang

dicapai siswa setelah ia menempuh proses belajar mengajar.

Tes formatif dilaksanakan pada saat berlangsungnya proses belajar mengajar, khususnya pada akhir pengajaran, sedangkan tes sumatif dilaksanakan pada akhir suatu satuan program, misalnya pada akhir caturwulan, semester dan sejenisnya. Pertanyaan biasanya diajukan secara lisan ataupun tertulis untuk tes formatif dan secara tertulis untuk tes sumatif. Jenis tes biasa berbentuk uraian ataupun objektif. Data hasil penilaian biasanya dalam bentuk skor sehingga bisa diketahui posisi siswa dalam kelompoknya ataupun posisi siswa jika dibandingkan dengan kriteria tertentu sehubungan dengan tujuan yang harus dikuasai siswa.

Secara umum manfaat hasil penilaian tersebut berguna bagi guru dan bagi siswa disamping bagi tenaga kependidikan lainnya, yaitu wali kelas, guru pembimbing dan mungkin bagi kepala sekolah.

### **1. Manfaat data penilaian hasil belajar formatif**

Data hasil penilaian formatif dapat diperoleh guru secara langsung pada akhir proses belajar mengajar berupa skor hasil pascates. Data ini, di samping menggambarkan penguasaan tujuan instruksional oleh

para siswa, juga memberi petunjuk kepada guru tentang keberhasilan dirinya dalam mengajar. Oleh sebab itu, data ini sangat bermanfaat bagi guru dalam upaya memperbaiki tindakan mengajar selanjutnya.

Misalkan terhadap sejumlah tujuan instruksional diberikan tes 10 pertanyaan seperti telah dirumuskan oleh guru dalam satuan pelajaran yang dibuatnya. Pertanyaan tersebut diajukan kepada siswa yang menjawab benar seluruhnya, salah satu, salah dua, salah tiga dan seterusnya. Seandainya kita menggunakan kriteria 70%, maka minimal keberhasilan pengajaran tersebut adalah 70% dari jumlah siswa yang diajar harus dapat menjawab benar tujuh soal. Apabila kurang dari kriteria tersebut, pengajaran belum berhasil. Kemudian hitung berapa banyak siswa yang belum mencapai kriteria tersebut dan pertanyaan mana yang pada umumnya salah di jawab oleh siswa. Pertanyaan yang salah dijawab oleh kebanyakan siswa menunjukkan tujuan instruksional khusus yang belum dicapai oleh pengajaran tersebut. Selanjutnya dikaji siapa-siapa saja yang menjawab salah pertanyaan-pertanyaan tersebut (belum menguasai tujuan instruksional). Apabila siswa yang telah menjawab pertanyaan tersebut kebanyakan dari siswa yang lemah, dapat dikatakan lumrah. Namun, jika berasal dari siswa

kategori sedang atau pandai, berarti ada sesuatu yang harus diperbaiki dalam hal proses belajar mengajar yang telah dilaksanakan. Di sinilah guru perlu melihat kembali satuan pelajaran dan menelusuri tindakan yang dilakukannya pada waktu mengajar. Dari kejian hasil penilaian ini guru dapat memetik manfaat dalam:

- 1) memperbaiki program pengajaran atau satuan pelajaran di masa mendatang, terutama dalam merumuskan tujuan instruksional, organisasi bahan, kegiatan belaja-mengajar dan pertanyaan penilaian.
- 2) Meninjau kembali dan memperbaiki tindakan mengajarnya dalam memilih dan menggunakan metode mengajar, mengembangkan kegiatan belajar siswa, bimbingan belajar, tugas dan latihan para siswa, dan lain-lain.
- 3) Mengulang kembali bahan pengajaran yang belum dikuasai para siswa sebelum melanjutkan dengan bahan baru, atau memberi penugasan kepada siswa untuk memperdalam bahan yang belum dikuasainya.
- 4) Melakukan diagnosis kesulitan belajar para siswa sehingga dapat di temukan faktor penyebab kegagalan siswa dalam menguasai tujuan instruksional. Hasil diagnosis ini dapat dijadikan

bahan dalam memberikan bantuan dan bimbingan belajar kepada para siswa.

Data hasil penilaian formatif, termasuk catatan kelemahan para siswa, sebaiknya dicatat dan didokumentasikan sehingga perkembangan kemajuan belajar siswa dari waktu ke waktu mudah dilihat. Data ini juga dapat dijadikan bahan pertimbangan bagi pengisian nilai raport siswa pada akhir caturwulan atau semester di samping hasil penilaian sumatif.

Apabila di sekolah telah ada guru pembimbing, data dan informasi kelemahan para siswa yang diperoleh melalui kajian hasil penilaian tersebut dapat disampaikan kepada guru pembimbing itu agar dapat dijadikan bahan dalam upayanya memberikan bimbingan kepada siswa tersebut. Kerja sama guru bidang studi dengan guru pembimbing sangat di harapkan dalam upaya meningkatkan kualitas pendidikan.

## **2. Manfaat data penilaian hasil belajar sumatif**

Penilaian sumatif dilaksanakan pada akhir program selama satu catur wulan atau semester untuk mengukur tingkat penguasaan hasil belajar para siswa. Bahan pertanyaan bersumber dari GBPP untuk catur wulan atau semester tersebut dituangkan dalam pertanyaan tes yang pada umumnya dibuat dalam

bentuk objektif. Bagaimanapun hasilnya yang diperoleh dari test sumatif tampaknya menjadi keputusan akhir mengingat tidak adanya kesempatan bagi guru untuk memperbaiki kekurangan para siswa pada semester tersebut. Perubahan baru bisa dilakukan pada tahun berikutnya atau sekedar bahan untuk penyempurnaan semester berikutnya. Oleh sebab itu, data hasil penilaian sumatif dapat digunakan oleh guru untuk:

- 1) membuat laporan kemajuan belajar siswa (dalam hal ini menentukan nilai prestasi belajar untuk mengisi raport siswa) setelah mempertimbangkan pula nilai dari hasil tes formatif dan kemajuan-kemajuan belajar lainnya dari setiap siswa.
- 2) Menata kembali seluruh pokok bahasan dan subpokok bahasan setelah melihat hasil tes sumatif terutama kelompok materi yang belum dikuasainya. Konsep esensi pokok bahasan yang belum dikuasai siswa dilihat kembali, baik dalam hal tingkat kesulitannya, ruang lingkup dan susunannya, waktu yang diperlukan maupun buku sumber yang relevan untuk dipelajari siswa. Hasil penataran tersebut berupa program belajar atau GBPP yang telah disempurnakan tanpa mengurangi ketentuan yang berlaku dalam kurikulum, minimal untuk

digunakan pada caturwulan atau semester yang sama tahun berikutnya.

- 3) Melakukan perbaikan dan penyempurnaan alat penilaian tes sumatif yang telah digunakan berdasarkan hasil-hasil yang telah diperoleh atau dicapai siswa. Soal-soal yang dijawab salah oleh sebagian besar siswa hendaknya dikaji ulang dari berbagai segi, yaitu dari tingkat kesulitan soal, konsep esensi yang ditanyakan, kebenaran jawaban dari pertanyaan, bahasa yang digunakan, relevansi pertanyaan dengan kemungkinan jawabannya, jumlah soal dan waktu yang disediakan, bentuk soal, dan lain-lain.
- 4) Merancang program belajar bagi siswa (GBPP) pada semester berikutnya berdasarkan hasil-hasil yang telah dicapai dari tes sumatif program belajar sebelumnya. Kajian dilakukan pada GBPP semester atau caturwulan selanjutnya dengan melihat ada tidaknya materi prasyarat yang harus dikuasai pada semester atau caturwulan sebelumnya. Jika ada, apakah materi prasyarat tersebut telah dikuasai siswa yang ditunjukkan oleh hasil tes sumatif tersebut.

Data hasil penilaian sumatif juga bermanfaat bagi kepala sekolah, administrator pendidikan, atau

supervisor pendidikan sebagian bahan dalam menentukan tingkat keberhasilan pendidikan di sekolah yang bersangkutan, termasuk kemampuan guru dalam melaksanakan tugas-tugas profesinya. Berdasarkan informasi dan bahan-bahan tes sumatif dapat ditetapkan upaya pembinaan pendidikan di sekolah, dan pembinaan guru dalam proses belajar mengajar. Dengan demikian, pemanfaatan data hasil penilaian sumatif sangat berguna bukan hanya bagi guru, melainkan juga bagi kepala sekolah dan supervisor pendidikan dalam rangka meningkatkan pembinaan pendidikan disekolah dalam hal perencanaan, pelaksanaan atau penyelenggaraan, dan dalam penilaian, pengawsan, atau pemantauan proses dan hasil pendidikan.

#### **D. Manfaat data hasil penilaian proses belajar mengajar**

Data hasil penilaian proses belajar mengajar sangat bermanfaat bagi guru, siswa dan kepala sekolah. Bagi guru ialah ia dapat mengetahui kemampuan dirinya sebagai pengajar, baik kekurangan maupun kelebihanannya. Guru juga dapat mengetahui pendapat dan aspirasi para siswanya dalam berbagai hal yang berkenaan dengan proses belajar mengajar. Berdasarkan informasi ini guru dapat memperbaiki dan

menyempurnakan kekurangannya dan mempertahankan atau meningkatkan kelebihan-kelebihannya.

Demikian juga bagi siswa, data hasil penilaian mengenai cara belajar, kesulitan belajar dan hubungan sosial dapat dijadikan bahan untuk meningkatkan upaya dan motivasi belajar yang lebih baik lagi.

Pada pokok lain, kepala sekolah dapat memikirkan upaya-upaya pembinaan para guru dan siswa berdasarkan pendapat, suara, aspirasi dari berbagai pihak (guru, siswa, orang tua), yaitu melengkapi sarana belajar, meningkatkan kemampuan profesional tenaga guru, pelayanan sekolah, perpustakaan sekolah, tata tertib sekolah, disiplin kerja, pengawasan dll. Ideal sekali apabila kepala sekolah melaksanakan sendiri penilaian proses belajar mengajar dengan menyiapkan terlebih dahulu alat-alat penilaian seperti kuesioner, pedoman observasi atau skala minat yang dibakukan. Dengan alat-alat penilaian tersebut setiap saat dapat digunakan dalam menilai kemampuan guru maupun siswa dalam proses belajar mengajarnya. Hasilnya diolah dan dianalisis sebagai bahan dan pertimbangan untuk melakukan pembinaan staf dan penyempurnaan penyelenggaraan pendidikan di sekolahnya.

### **E. Manfaat hasil penilaian bagi penelitian pendidikan**

Penelitian-penelitian pendidikan yang dilaksanakan oleh badan penelitian pendidikan dan oleh perguruan tinggi banyak mengambil data yang ada di sekolah. Berbagai jenis data dikumpulkan oleh para peneliti untuk keperluan pemecahan masalah yang ditelitinya. Penelitian pendidikan, baik yang sifatnya pemecahan masalah pendidikan, pengembangan ilmu maupun penelitian kebijakan dan pembangunan, memerlukan data informasi dari berbagai pihak seperti guru, kepala sekolah, siswa dan tenaga kependidikan lainnya, baik yang berkenaan dengan penyelenggaraan pendidikan pengajaran maupun dengan hasil-hasil pendidikan pengajaran. Sehubungan dengan itu, data hasil penilaian, baik penilaian hasil belajar maupun proses belajar mengajar akan sangat berguna bagi para peneliti dalam kegiatan penelitiannya.

Hasil penelitian sering dijadikan dasar dan bahan pertimbangan dalam menentukan kebijakan pendidikan. Itulah sebabnya data hasil penilaian yang salah atau keliru dan tidak objektif akibat pencatatan yang tidak teratur atau tidak lengkap, jika disampaikan kepada peneliti, akan menghasilkan temuan penelitian yang salah pula atau tidak sesuai dengan kondisi dan keadaan di lapangan. Sudah barang tentu temuan penelitian yang

dihasilkannya, jika digunakan sebagai bahan dalam membuat kebijaksanaan, akan berbahaya dan merugikan bagi pendidikan secara keseluruhan. Mengingat manfaat di atas, data hasil penilaian baik penilaian proses maupun penilaian hasil belajar, perlu didokumentasikan secara rapi dan teratur sehingga siap digunakan untuk berbagai kepentingan manakala diperlukan.

## **Rangkuman**

Hasil penilaian yang dibuat oleh guru dalam bidang studi atau mata pelajaran yang diajarkannya tidak hanya berguna bagi dirinya, dan bagi siswanya, tetapi juga harus dimanfaatkan oleh semua staf sekolah dalam rangka meningkatkan kualitas pendidikan di sek olah yang bersangkutan. Untuk itu, setiap guru bidang studi atau mata pelajaran perlu memberikan laporan tentang data hasil penilaian secara periodik kepada berbagai pihak, yakni kepala sekolah, wali kelas, guru pembimbing, dan juga kepada rekan guru lainnya. Pada tahap selanjutnya data ini sebaiknya dimanfaatkan bagi kepentingan tugas-tugas yang dipikulnya sesuai dengan peran masing-masing.

Laporan dan pemanfaatan data hasil penilaian mencakup data penilaian proses belajar-mengajar dan penilaian hasil belajar siswa. Bentuk dan isi laporan

dapat dibuat atau ditentukan oleh sekolahnya sesuai dengan kesepakatan dengan kriteria mudah, sederhana, dan bermakna untuk dipelajari dan dimengerti oleh semua pihak. Data laporan hasil penilaian hendaknya didokumentasikan secara teratur agar sewaktu-waktu dapat digunakan manakala diperlukan.

## DAFTAR PUSTAKA

- Arikunto, Suharsimi (2003), *Dasar-dasar Evaluasi Pendidikan*, Jakarta: Bumi Aksara.
- Arikunto, Suharsimi dan Safruddin A. Jabar (2004). *Evaluasi Program Pendidikan: Pedoman Teoritis Praktis Bagi Praktisi Pendidikan*. Jakarta: Bumi Aksara
- Brown, Federick E., (1976). *Principles of Education and Psychological Testing*. 2<sup>nd</sup> ed. New York: Holt, Rinehart & Winston.
- Cronbach, Lee J. (1984). *Essential of Psychological Testing*. New York: Harper & Row. Publishers.
- Gronlund, Norman E., (1985). *Measurement and Evaluation in Teaching*. New York: Macmillan Publishing Company.
- Popham, W.James (1980). *Modern Educational Measurement*. Englewood Cliffs, New Jersey : Printice Hall, Inc.
- Soekamto, Toeti. (1989). *Penilaian Hasil Belajar Mahasiswa*. Bahan Penataran Pelaksanaan Kegiatan Instruksional dan Penilaian Hasil Belajar Mahasiswa. BM-PTS Wilayah III.
- Soekartawi. (1995). *Monitoring dan Evaluasi Proyek Pendidikan*. Jakarta: Pustaka Jaya.
- Stufflebeam, D.L. (1972), *Educational Evaluation and Decision Making*, Itasca: Peacock Publisher, Inc.
- Sudjana, Nana, (1995), *Penilaian Hasil Proses Belajar Mengajar*, Bandung: Rosdakarya.
- Sudijono, Anas. (1998). *Pengantar Evaluasi Pendidikan*. Jakarta: Raja Grafindo Persada.

ISBN 978-623-91636-3-1



9

786239

163631