

KONTRUKSI TES



OLEH :

LODIANA AYU, S.Psi



**FAKULTAS PSIKOLOGI
UNIVERSITAS MEDAN AREA
M E D A N**

KATA PENGANTAR

Perkembangan instrumen pengukuran yang sedemikian pesat di bidang Psikologi dan pendidikan akhir-akhir ini menuntut perhatian, agar hasil evaluasi yang diharapkan dapat memberikan informasi yang akurat. Keberadaan alat evaluasi dalam bidang psikologi adalah integral dalam tindakan intervensi. Untuk membantu para praktisi psikologi, khususnya para tester dalam hal pelaksanaan kegiatan assessment dirasa perlu membekali mereka dengan pengetahuan praktis terhadap pemilihan instrumen-instrumen yang baik dari kaca mata konstruksi suatu alat tes.

Lebih dari pada itu diktat ini akan memberikan informasi dasar dalam merencanakan pembuatan alat-alat tes yang dibutuhkan. Sering seorang praktisi psikologi diminta untuk melakukan evaluasi terhadap gejala-gejala perilaku yang khusus yang kadang-kadang instrumen standar belum tersedia. Pada kondisi tersebut kiranya diktat ini dapat berguna dalam membantu merencanakan pembuatan alat evaluasi yang dimaksud.

Mahasiswa Fakultas Psikologi yang merupakan subjek utama pengguna diktat ini, akan mendapatkan gambaran bagaimana teknik-teknik penyusunan alat evaluasi baik dalam konteks kebutuhan belajar mata kuliah Konstruksi Tes maupun dalam kegiatan penelitian tugas akhir yang kelak akan dijalaninya pada saat pengambilan data penelitian.

Selain teknik-teknik penyusunan alat ukur, dalam buku diktat ini terdapat penilaian portofolio. Penilaian teknik portofolio memang belum umum dilaksanakan di sekolah-sekolah maupun Perguruan Tinggi di Indonesia. Namun mengingat fungsinya yang cukup baik dalam memvalidasi hasil penilaian yang formal dirasa perlu memperkenalkan sekilas tentang penilaian portofolio yang sebenarnya sudah waktunya dikembangkan di institusi pendidikan di negara kita untuk meningkatkan cara belajar siswa aktif.

Mudah-mudahan buku diktat ini dapat dipergunakan mahasiswa dan dosen dalam proses pembuatan alat ukur. Kritik dan saran yang disampaikan dalam penyempurnaan buku diktat ini sangat kami harapkan.

Medan, November 2002

Penulis

DAFTAR ISI

KATA PENGANTAR	i
DAFTAR ISI	ii
BAB I. PENDAHULUAN	1
A. Permasalahan	2
B. Pengertian	2
C. Tujuan & Fungsi Evaluasi/Testing	3
D. Klasifikasi Tes	4
BAB II. SYARAT-SYARAT TES YANG BAIK	6
1. Harus Valid	6
2. Harus Reliabel	8
3. Harus Distandardisasikan	9
4. Harus Obyektif	9
5. Harus Diskriminatif	9
6. Harus Komprehensif	16
7. Harus Praktis	16
BAB III. PENGEMBANGAN ALAT TES	17
A. Pengembangan Tes Pilihan Ganda	17
B. Pengembangan Tes Benar-Salah	21
C. Pengembangan Tes Menjodohkan	24
D. Pengembangan Tes Melengkapi	25
E. Pengembangan Tes Esai	26
F. Pengembangan Tes Lisan	30
G. Pengembangan Tes Tindakan	32
BAB IV. ADMINISTRASI TES	35
A. Penyiapan Perangkat Tes	35
B. Format-format Perencanaan	36
C. Petunjuk Tes	39
D. Waktu Tes	41
E. Penyajian Tes	41

BAB V. TEKNIK PENGOLAHAN DAN PENGGUNAAN HASIL PENGUKURAN.....	52
KURAN.....	52
A. Penyajian Skor	52
B. Ukuran Tendensi Sentral	55
C. Ukuran Variabilitas	57
D. Pembuatan Nilai/Konversi Skor Menjadi Nilai	60
E. Laporan Hasil Penilaian	65
BAB VI. KAWASAN PENGUKURAN	68
1. Pengetahuan	68
2. Pemahaman	68
3. Penerapan	69
4. Analisis	69
5. Sintesis	69
6. Evaluasi	69
BAB VII. PENILAIAN PORTOFOLIO	73
1. Makna Penilaian Portofolio	73
2. Fungsi Portofolio	74
3. Bentuk Portofolio	75
4. Portofolio dan Testing	77
5. Portofolio Didalam Kelas	77
6. Pengembangan Portofolio	78
7. Konfrensi Portofolio	79
8. Penilaian Portofolio	79
DAFTAR PUSTAKA	81

BAB I

PENDAHULUAN

A. Permasalahan

Istilah tes telah sedemikian populernya diberbagai kalangan masyarakat kita sehingga bukan lagi sesuatu yang asing. Boleh dikata hampir semua orang pernah mendengar, membicarakan, atau bahkan pernah menjalani tes dalam situasi dan keperluan yang berbeda. Misalnya pada saat seorang memasuki jenjang pendidikan tertentu, atau juga pada saat seorang pelamar menjalani tes dalam rangka seleksi dan penempatan serta pada saat seorang staf dipromosikan pada posisi atau tingkat jabatan tertentu dan untuk tujuan-tujuan lainnya.

Namun apabila ditanya mengenai apakah sebenarnya tes itu, kiranya tidak banyak orang yang benar-benar memahami dan dapat menjelaskannya secara baik. Apa yang terbayang dalam pikiran masing-masing orang apabila mendengar kata "tes" belum tentu merupakan konsep atau gambaran yang tepat, walaupun biasanya memang menggambarkan situasi yang kurang lebih serupa.

Dalam bidang psikologi terdapat berbagai bentuk dan jenis tes standar yang digunakan, yang sebahagian sudah mulai dikenal masyarakat banyak. Misalnya tes IQ untuk menentukan tingkat kecerdasan seseorang, Tes Bakat untuk mengenal kemampuan-kemampuan khusus, tes minat, tes kepribadian dalam arti sempit serta tes lainnya. Namun masih banyak gejala-gejala psikologis yang belum dapat dideteksi secara memadai, hal ini karena terbatasnya alat-alat tes yang ada sekarang. Tes yang tersedia sekarang kebanyakan merupakan hasil adaptasi dari Negara-negara maju yang kebanyakan tidak sesuai dengan budaya maupun iklim masyarakat Indonesia. Revisi yang terus menerus dituntut agar sesuai dengan perkembangan dan agar dapat memberikan gambaran yang akurat tentang perilaku yang khas keIndonesiaan.

Kebutuhan akan pengembangan test sebagai alat evaluasi atau dalam istilah psikologi alat diagnostik sangat menuntut banyak perhatian para ahli dan praktisi. Dalam dunia industri seorang manager sudah mulai menyadari betapa sentralnya fungsi pengembangan sumber daya manusia untuk meningkatkan produktivitas usaha. Faktor latihan dan pengembangan memiliki pengaruh apabila mutu Sumber daya manusianya memiliki kemampuan untuk dikembangkan. Kemampuan sumber daya manusia baik yang berupa potensi yang laten maupun potensi yang actual harus dapat diidentifikasi secara tepat agar dapat menempatkan seseorang pada posisi yang tepat.

Demikian juga halnya dalam konteks konseling dan psikoterapi, fungsi tes sangat membantu memberi informasi seberapa baik misalnya seseorang melakukan penyesuaian diri terhadap problem-problem yang dihadapinya dilingkungan sekitar, seberapa besar toleransi seseorang terhadap stress, tingkat inisiatif dari klien, dan berbagai informasi lainnya.

Dalam dunia pendidikan tes tentunya menjadi bagian yang integral dengan proses belajar mengajar, sehingga sebagai seorang pendidik dituntut untuk dapat senantiasa merancang tes yang memadai dalam mengevaluasi penguasaan siswa terhadap materi yang diberikan dikelas. Otomatis Guru tidak hanya berperan sebagai fasilitator dan motivator, tapi juga sebagai sebagai tester dalam arti luas. Kesalahan dalam merancang tes dapat berakibat langsung maupun tidak langsung terhadap motivasi belajar siswa, karena nilai yang diperoleh siswa merupakan cermin bagi kemampuan dirinya untuk dapat menguasai materi-materi tertentu sehingga lebih jauh akan berpengaruh terhadap pembentukan konsep dirinya.

Pada kasus-kasus yang khusus guru juga dituntut untuk menjadi konselor yang diharapkan mengenal lebih jauh tentang potensialiti maupun keterbatasan-keterbatasan dari murid-muridnya agar dia dapat menyusun strategi yang lebih tepat dalam proses pembelajaran disekolah. Diantara murid tentunya ada siswa yang perlu mendapat perhatian khusus.

B. Pengertian

Kata test berasal dari bahasa latin "Testum" yang diartikan sebagai alat untuk mengukur tanah. Dalam bahasa Perancis yaitu ukuran yang dipergunakan untuk membedakan emas dan perak dari logam-logam yang lain. Perkataan tes dalam bidang psikologi pertama sekali digunakan oleh J.M. Cattell (1890) yang bertujuan untuk menentukan (mengukur) aspek-aspek tertentu dari kepribadian.

Berbicara masalah Tes sama halnya dengan membicarakan penilaian atau evaluasi. Evaluasi yang dalam bahasa Inggris Evaluation dibatasi sebagai *Evaluation rever to the process to determining the value of something* (Etwin Wand & Gerald W. Brown) yaitu suatu tindakan atau suatu proses untuk menentukan nilai dari pada sesuatu. Pengertian nilai dalam defenisi diatas dikaitkan dengan makna yang diberikan terhadap sesuatu. Sesuatu yang sama mungkin saja memiliki nilai yang berbeda. Katakanlah seseorang menjawab benar 6 dari 10 pertanyaan/soal yang diajukan dalam ulangan/ujian. Untuk itu ia mendapat skor 60 (mungkin saja skor untuk jawaban atas enam pertanyaan itu kurang atau lebih dari 60, tetapi untuk kasus ini katakana siswa diberi skor 70). Melalui proses tertentu skor 60 itu diubah menjadi angka 7 dalam rentangan angka 1 – 10 seperti yang berlaku di sekolah. Dalam kesempatan lain mungkin saja skor 60 itu hanya diberi nilai 4 atau 5.

Skor angka yang diberikan merupakan perwakilan nilai yang ditentukan berdasarkan aturan yang berlaku. Jadi dalam hal ini angka digunakan sebagai lambing nilai (mis : 1 – 10 di sekolah SD-SLTA, A,B,C,D,E di PT, dll). Adapun arti dari suatu angka/nilai hanya dapat ditetapkan berdasarkan keterhubungannya dengan suatu kondisi tertentu. Sebagai contoh misalnya disuatu Perguruan tinggi sebagian terbesar Mahasiswanya memiliki prestasi tinggi. Indek Prestasi setiap kelas 3,0 atau lebih. Oleh karena itu IP 2,5 yang diperoleh seorang Mahasiswa di Perguruan Tinggi itu memiliki arti yang berbeda dari IP 2,5 di Perguruan Tinggi lain yang Indek Prestasi rata-rata kelasnya adalah 2,4.

Sedangkan pengukuran yang dalam bahasa Inggris *Measurement* dibatasi *Measurement mens the act or process or exestaining the exent or quantity of something* yaitu suatu tindakan atau proses untuk membuat luas atau kuantitas daripada sesuatu.

Dari batasan evaluasi (penilaian) dan pengukuran yang disebutkan diatas terdapat perbedaan, dimana penilaian akan memberikan jawaban terhadap pertanyaan "what value" sedangkan pengukuran akan memberikan jawaban terhadap pertanyaan "How much". Dalam pengertian evaluasi terdapat skor dan nilai. Kedua-duanya dapat berbentuk angka. Angka pertama yaitu 60 disebut dengan skor, sedangkan angka kedua yaitu 7, berdasarkan rentang nilai 1 -- 10 disebut nilai.

Skor adalah hasil dari suatu pengukuran, sedangkan nilai merupakan perpaduan antara hasil pengukuran dengan standar patokan tertentu atau konversi hasil pengukuran kedalam kriteria tertentu dalam rentang angka yang tersedia. Oleh karena itu dikenal *makna* dari setiap angka yang disebut nilai. Lagi pula nilai dapat dikonversi dalam bentuk hurup atau symbol lain.

Walaupun ada perbedaan antara pengukuran dengan penilaian (evaluasi) namun kedua hal tersebut tidak dapat dipisahkan karena antara pengukuran dan penilaian tersebut terdapat hubungan yang sangat erat. Sebab untuk dapat mengadakan penilaian yang tepat harus terlebih dahulu mengadakan pengukuran. Misalnya untuk menilai apakah seorang anak dapat membaca dengan lancar atau tidak, maka kita perlu mengukur berapa jumlah kata-kata yang dapat dibacanya dalam tempo satu menit, berapa kesalahan-kesalahan yang dibuatnya, dan sebagainya. Sebaliknya hasil pengukuran-pengukuran yang dilakukan tidak akan memberi arti apa-apa kalau tidak dihubungkan dengan penilaian. Misalnya apabila berdasarkan suatu pengukuran kita ketahui bahwa seorang anak dapat membaca dengan kecepatan 50 kata dalam satu menit. Apakah dapat kita katakana bahwa anak itu cukup lancar membaca atau tidak ? Tentu saja kita belum bisa mengatakan anak tersebut lancar membaca, tanpa kita ketahui criteria penilaiannya. Jika kecepatan membaca anak-anak yang sekolah dengan anak tadi pada umumnya ada 35 kata tiap menit, maka anak tadi dapat kita katakana anak yang lancar dalam membaca. Tetapi kalau kecepatan membaca anak-anak yang sekelas dengan anak tadi umumnya adalah 65 kata tiap menit, maka anak tadi tergolong anak lambat.

C. Tujuan & Fungsi Evaluasi/Testing

Pada dasarnya tujuan melakukan evaluasi atau testing adalah untuk mendapatkan informasi yang dapat digunakan dalam memberikan penilaian tentang obyek/keadaan yang diukur. Penilaian tersebut dijadikan sebagai dasar pertimbangan untuk mengambil keputusan. Seorang Guru melakukan penilaian terhadap siswa untuk dapat mengambil suatu kesimpulan tentang apa yang perlu dipersiapkan, dibenahi terutama yang berkenaan dengan setiap komponen proses belajar mengajar. Seorang psikolog melakukan psikotes/evaluasi terhadap pelamar dalam rangka mengambil kesimpulan tentang yang berkaitan dengan personality pelamar sehingga dapat dijadikan sebagai

pertimbangan dalam seleksi dan placement. Secara singkat tujuan tersebut dapat dibedakan dalam 2 kategori fungsi yaitu :

1. Untuk Research

Dalam setiap lapangan ilmu pengetahuan, penelitian adalah kegiatan yang mutlak harus dilakukan. Dalam lapangan psikologi testing untuk tujuan riset dapat bermacam-macam bentuknya a.l :

a. Riset pembuatan test

Berbagai tes psikologi disusun berdasarkan riset yang secara ilmiah benar-benar dapat dipertanggungjawabkan.

b. Riset untuk eksplorasi

Tes psikologi sering digunakan untuk mengetahui sifat-sifat psikologis lapisan masyarakat tertentu.

c. Riset untuk menerangkan dan menunjukkan penyelesaian problem-problem sosial tertentu.

2. Untuk Diagnosis Psikologik

Untuk sebagian besar tujuan testing adalah untuk dapat membuat diagnosis secara psikologis. Diagnosis sendiri bukanlah tujuan akhir, melainkan baru merupakan pangkal kerja dari rangkaian tindakan tertentu. Tujuannya dapat bermacam-macam a.l:

a. Untuk kepentingan seleksi

b. Untuk penempatan jabatan

c. Untuk psikoterapi

d. Untuk bimbingan penyluhan & karir

D. Klasifikasi Tes

Tes psikologi itu sangat banyak macam ragamnya, pengklasifikasiannya dapat ditinjau dari beberapa hal :

1. Jumlah testee yang dapat diukur

a. Individual test

b. Group test

2. Penyelesaian soal tes

a. Verbal test

b. Non Verbal test

3. Cara menilai tes

a. Alternative test

b. Graduil test

4. Fungsi psikologis yang dijadikan Sasaran tes

a. Tes perhatian

b. Tes fantasi

c. Tes ingatan

d. Tes kemauan

e. dsb

5. Isi & waktu yang digunakan
 - a. Speed test
 - b. Power test
6. Rancangan teorinya
 - a. Peroyective test
 - b. Non proyective test
7. Bentuk tes
 - a. True false test (benar-salah)
 - b. Multiple choice test (pilihan berganda)
 - c. Fill in test (isian)
 - d. Matching test (mencari pasangan)
 - e. Completion test (penyempurnaan)
 - f. Object arrangement test (mengatur obyak)
 - g. Digit span test (deret angka)
 - h. Block design test (rancangan balok)
 - i. Asociation test (assosiasi)
 - j. Inventori test
 - k. Gravis test
 - l. Proyeksi test
 - m. Test Situasi
 - n. dll
8. Berdasarkan penciptanya
 - a. Tes Weschsler
 - b. Tes Kraepelin
 - c. Tes Binet
 - d. Tes Rorschach
9. Cara penyajiannya
 - a. Tes lisan
 - b. Tes Tulisan
 - c. Tes perbuatan
10. Sasaran pengukuran
 - a. Cognitive test (Inteligensi, Bakat)
 - b. Non Cognitive Test (Kepribadian, Minat)
11. Penggolongan umum
 - a. Tes Inteligensi
 - b. Tes Bakat
 - c. Tes kepribadian
 - d. Tes prestasi

BAB II SYARAT-SYARAT TES YANG BAIK

Akurasi hasil pengukuran dan penelitian erat hubungannya dengan baik tidaknya teknik dan jenis serta alat penilaian/pengukuran yang dipergunakan. Dengan demikian alat pengukuran haruslah memenuhi syarat yang baik. Seorang pendidik dan penguji harus memahami syarat-syarat tes yang baik serta dapat menganalisis kebaikan serta keterbatasan alat ukur yang disusun. Dengan mengetahui tingkat kebaikan dan keterbatasan alat pengukuran tersebut, penafsiran hasil dari pengukuran lebih berhati-hati sehingga akan memberikan hasil yang lebih akurat dan objektif. Sebelum kita mempergunakan suatu alat tes ada beberapa syarat yang harus kita pertimbangkan kualitasnya yaitu :

I. Harus Valid

Validitas tes adalah hal yang sesuai tidaknya isi tes dengan criteria/tujuan tes. Makin tinggi validitas suatu alat tes, maka semakin sesuai data yang dihasilkannya sesuai dengan tujuan pengukuran. Untuk memudahkan pengertian dapat dijelaskan sebagai berikut : Kalau kita hendak mengukur tinggi seseorang, maka ukuran yang cocok adalah meter, dan bila kita hendak mengukur berat sesuatu benda ukuran yang cocok adalah timbangan yang satuannya disebut gram, kilogram dan sebagainya. Demikian juga halnya dengan tes, apabila kita mau mengukur pengetahuan (ingatan) dari anak, maka item pertanyaan berikut tidak cocok (tidak valid) :

- Mengapa tumbuhan yang tidak kena sinar matahari tidak subur ?
- Apa akibat dari pengangguran ?

Demikian juga halnya tes psikologi yang disusun disuatu negara tidak selalu cocok untuk dipakai di negara lain, hal ini disebabkan latar belakang kebudayaan dari kedua negara tersebut tidak sama, misalnya : pertanyaan "apakah apel itu ?" di negara barat adalah pertanyaan yang tingkat kesukarannya rendah, sedang di Indonesia pertanyaan itu adalah termasuk sukar karena buah apel bukan buah-buahan yang biasa (umum) dijumpai. Dengan demikian harus dimodifikasi misalnya dengan pertanyaan "Apakah pisang itu"

Suatu perangkat tes yang menghasilkan skor tidak berkorelasi dengan skor tes lain dimana tes lain itu diperkirakan sudah baik, maka tes yang bersangkutan dapat dikatakan tidak valid. Memperhatikan ada beberapa factor yang mempengaruhi tes, maka ada bermacam-macam jenis validitas sesuai dengan factor dimaksud. Berikut ini adalah macam-macam jenis validitas tersebut :

a. Face validity

Face validity disebut juga validitas tampang, validitas semu. Disebut demikian karena sebenarnya jenis ini tidak menunjukkan kesesuaian hasil tes dengan criteria, tetapi menyangkut "tampang" tes itu yaitu bentuk luar seperti bentuk penulisan item, kertas yang dipergunakan dan lain-lain.

Memang jenis validitas ini nampaknya nilainya rendah karena bentuk tes yang menarik, dicetak rapi dan kertas yang dipergunakan baik belum tentu isi tes itu valid. Namun demikian validitas semu perlu mendapat perhatian. Tes yang bentuk luarnya ditata rapi membangkitkan suatu kesan bahwa tes itu baik sehingga dengan demikian peserta akan bersungguh-sungguh dalam bekerja.

Di Kampus para dosen sering terikat tuntutan masalah efisiensi, termasuk dalam pembuatan naskah ujian, sehingga dapat mengurangi efektivitas.

Contoh :

1). Harga 5 buah jambu 125 rupiah.

Harga sebuah jambu adalah : a. 25 rupiah, b. 20 rupiah, c. 1,25 rupiah, d. 35 rupiah

2). Harga 1 buah buku 600 rupiah.

Harga 1 lusian buku adalah :

a. 6.000,- rupiah

b. 7.000,- rupiah

c. 7.200,- rupiah

d. 9.000,- rupiah

Contoh item 1 penulisannya efisien tetapi “tampang”nya kurang baik, sedangkan

item 2 lebih memenuhi face validity.

b. Content validity

Content validity sering disebut Logical validity dan ada yang menyebut Curricular validity. Content validity menyangkut perumusan item-item dan kesesuaian isi/makna item-item dengan tujuan pengukuran. Perumusan kalimat-kalimat dalam perangkat tes harus memenuhi tata bahasa yang baik, serta mudah dipahami.

c. Empirical validity (validitas empiris) ialah kecocokan hasil yang digambarkan oleh tes. Suatu perangkat tes yang mempunyai validitas empiris tinggi apabila distribusi skor yang dihasilkan tes itu mempunyai korelasi yang tinggi dengan distribusi skor yang dihasilkan “penilaian lain”. Penilaian lain dalam hal ini haruslah penilaian yang dianggap sudah “baik”.

c. Predictive validity

Predictive validity (validitas ramal) bermakna taraf keserasian suatu tes dalam meramalkan tingkah laku (keberhasilan) peserta yang dites pada masa yang akan datang. Seorang murid yang berhasil dalam mencapai angka tertinggi dalam tes matematika yang dibuat oleh guru A akan memperoleh angka tertinggi pula pada tes matematika yang dibuat oleh guru B apabila tes yang dibuat oleh kedua guru tersebut mempunyai predictive validity yang tinggi. Seorang calon pelamar pekerja yang memiliki skor yang tinggi dalam tes kecerdasan akan dapat menyelesaikan problem-problem pekerjaannya kelak apabila tes inteligensi yang diberikan memiliki predictive validity.

Dalam tes seleksi masuk suatu sekolah atau penerimaan kerja hendaknya mempunyai predictive validity yang tinggi. Dengan kata lain seseorang yang lulus dalam suatu seleksi, akan sukses dalam mengikuti program yang diberikan bila alat seleksi yang dipergunakan mempunyai predictive validity yang tinggi.

d. Factorial validity (Construct validity)

Faktorial validity (validitas factor) menunjuk kepada kemampuan perangkat tes untuk menggambarkan kelakuan yang dites dan sampai dimana hubungan antara hasil yang digambarkannya dengan kelakuan yang lain. Jadi tes hendaknya mempunyai item-item yang dapat mengungkapkan semua tingkah laku sebagai manifestasi dari factor yang diteliti. Suatu perangkat tes yang bertujuan untuk

mengukur kemampuan murid dalam menganalisa sesuatu pokok masalah, maka item-item yang disusun harus dapat mengungkapkan manifestasi dari kemampuan menganalisa secara lengkap, misalnya termasuk di dalamnya kemampuan membandingkan, memberi kritik, menunjukkan baik buruknya sesuatu dan lain-lain. Kedua validity yang terakhir ini sangat dituntut oleh tes yang standar seperti tes psikologi.

2. Harus Reliable

Reliable artinya dapat dipercaya. Jadi reliability adalah dimana suatu test menunjukkan masalah ketetapan (consistency) dari hasil tes tersebut. Dengan demikian reliabilitas adalah syarat tes yang menyangkut kepercayaan kita tentang ketetapan hasil bila dilakukan dua atau beberapa kali pengukuran pada kelompok yang sama.

Ketetapan yang dimaksud bukanlah kesamaan skor tetapi mencakup ketetapan kedudukan anak dalam kelompoknya.

Syarat reliabilitas sangat dituntut dalam tes yang standar misalnya tes inteligensi, bakat, minat dan lain-lain. Dalam tes hasil belajar yang disusun oleh guru pada umumnya kurang memperhatikan reliabilitas mengingat tes yang disusun hanya dipakai satu kali saja. Tetapi tes yang dipakai umum misalnya perangkat tes EBTANAS/EBTA, hendaknya mempunyai syarat reliabilitas mengingat tes yang bersangkutan mungkin dipergunakan pada masa yang akan datang.

Ada 3 masalah pokok yang mempengaruhi reliabilitas tes, yaitu :

a. Masalah yang bersangkutan dengan perangkat tes.

Panjang pendeknya kalimat item tes.

Item yang kalimatnya terlalu pendek apalagi sangat mudah diingat akan mengurangi reliabilitas tes karena setelah selesai pelaksanaan yang pertama murid akan membahasnya dan bila dilaksanakan untuk kedua kalinya semua murid telah dapat menjawabnya dengan benar. Demikian juga item yang kalimatnya terlalu panjang akan menimbulkan keraguan dalam memberi jawaban atau murid tidak berusaha menjawabnya.

b. Tingkat kesukaran tes

Item yang sangat sukar dan sangat mudah akan mengurangi tingkat reliabilitas tes, karena item yang sangat sukar akan selalu salah dijawab oleh kemungkinan seluruh peserta, sebaiknya item yang sangat mudah akan selalu benar dijawab oleh kemungkinan seluruh peserta tes. Sehingga bila dalam perangkat tes terdapat banyak item yang kesukarannya ekstrim akan mengurangi variasi skor dengan demikian akan mengurangi tingkat reliabilitas.

c. Kesesuaian waktu pelaksanaan tes

Jumlah waktu yang dipergunakan hendaknya disesuaikan dengan jumlah item dan tingkat kesukaran item. Waktu yang "agak lama", dimana setengah atau lebih dari murid yang dites sudah selesai sebelum waktu yang ditentukan habis, akan mengurangi sifat reliabilitas apalagi kalau metode perhitungannya melalui tes-retest.

Seorang anak yang sudah menyelesaikan pekerjaannya sepuluh menit sebelum waktu tes habis pada pelaksanaan pertama, mungkin akan menghabiskan waktu

yang tersedia pada pelaksanaan kedua, sehingga waktu yang dipergunakan tidak tepat sejajar lagi, untuk kedua pelaksanaan tes itu.

3. Harus Distandardisasikan

Standardisasi suatu alat tes bertujuan supaya setiap testee yang dites tersebut mendapat *perlakuan yang benar-benar sama*. Skor yang diperoleh dalam pengukuran/tes bersifat relatif. Skor tertentu punya arti kalau skor tersebut dibandingkan satu sama lain. Skor 50 bisa diartikan sangat baik atau sebaliknya.

Yang perlu distandardisasikan adalah :

a. Materi tes

Yaitu berupa bahan untuk membuat tes (kertas, karton, pewarna, balok-balok, dsb) dan item-itemnya (kata-kata, gambar, tanda-tanda, ukuran, dsb)

b. Penyelenggaraan tes

Yaitu berupa perlengkapan pada saat penyelenggaraan (meja, kursi, alat tulis, cara penyajian, waktu yang disediakan, dsb) serta situasi (suhu, ketenangan, penerangan)

c. Skoring tes

Yaitu cara-cara dan pertimbangan untuk menentukan skor. Batas-batas yang digunakan (Buatkan kuncinya)

d. Interpretasi hasil tes

Yaitu terhadap hasil testing yang sama harus diberikan interpretasi yang sama.

4. Harus Obyektif

Obyektivita sesuatu tes ditinjau dari segi apakah tester maupun interpreter mempunyai pengaruh terhadap penilaian hasil testing. Jadi yang obyektif itu adalah penilaiannya. Tes yang obyektif akan memberikan hasil yang sama kalau dinilai oleh tester yang berlainan. Tes yang standard umumnya telah mempunyai pedoman jawaban yang benar, begitu juga cara-cara menskoring serta menginterpretasi telah tersedia norma-norma yang dikembangkan. Namun demikian perlu juga dijaga bahwa test yang standar pun mungkin terjadi penyimpangan dalam hal scoring. Petugas scoring yang kurang memahami rumus scoring yang telah ditetapkan atau kurang bersifat jujur karena satu dan lain hal akan mengurangi sifat objektivitas tes ini.

5. Harus Diskriminatif

Yaitu kemampuan tes tersebut untuk menangkap gejala yang spesifik dari tiap responden sehingga menunjukkan perbedaan-perbedaan (diskriminasi) gejala tersebut dari individu yang satu dengan individu lainnya. Tes itu mampu menunjukkan perbedaan-perbedaan yang kecil dari sifat tertentu pada individu-individu yang berbeda. Ini dinamakan daya pembeda (*discriminating power*) yang biasa disingkat DP.

Daya pembeda adalah kemampuan aitem dalam membedakan antara peserta tes yang menjawab benar dari kelompok peserta yang pandai dengan jumlah peserta yang menjawab benar dari kelompok peserta yang lemah.

Berikut ini adalah satu contoh perhitungan daya pembeda dari bermacam teknik yang dipergunakan. Langkah-langkah untuk melakukan analisis item dalam menetapkan daya pembeda adalah sebagai berikut :

- 1). Susunlah lembaran jawaban dimulai dari lembaran jawaban yang memperoleh skor paling tinggi sampai kepada lembar jawaban yang memperoleh skor paling rendah.

Skor tes (X)	Frekuensi f	
38	5	} Kelompok Pandai 27 % skor tertinggi (dapat ditetapkan dengan nilai P_{83})
37	6	
35	8	
33	10	
32,5	2	
32	13	
-	-	} 46 % frekuensi peserta yang nilainya di tengah tidak dijadikan data analisis.
-	-	
-	-	
-	-	
10	15	} Kelompok Lemah 27 % skor terendah (dapat ditetapkan dengan nilai P_{27})
9	9	
8	8	
7	5	
6	3	
5	1	

Gambar Ilustrasi Penentuan Subyek Untuk Analisa Butir

- 2). Tentukan kelompok peserta yang pandai dan kelompok peserta yang lemah menurut besarnya skor total tes (X) individu, masing-masing 27 % frekuensi dari susunan teratas adalah kelompok pandai dan 27 % frekuensi dari susunan bawah adalah kelompok lemah, sisa lembar jawaban sebesar 46 % disisihkan dari perhitungan analisis item. Penentuan 27 % menurut Kelly (1939) adalah sample yang cukup representatif.
- 3). Selanjutnya buat table distribusi jawaban masing-masing untuk kelompok pandai (disingkat H) dan untuk kelompok lemah (disingkat L).

Tabel 2.1 Distribusi Jawaban Oleh Kelompok Pandai

No.urut	Nama Peserta	Nomor Butir Soal						
		1	2	3	4	5	6	dst
1.	Pamungkas Satria	√	√	√	x	√	√	dst
2.		√	√	√	√	x	√	dst
3.								
4.								
5.								
6.								
7.								
dst								
mis Nil = 40								
Jumlah jawaban betul mis :		35	32	35	30	25	20	dst
Persentase jawaban betul		87,5	80	87,5	75	62,5	50	dst

Tanda : √ = jawaban betul
 X = jawaban salah
 - = tidak dikerjakan

Tabel 2.2 Distribusi Jawaban Oleh Kelompok Lemah

No.urut	Nama Peserta	Nomor Butir Soal						
		1	2	3	4	5	6	dst
1.	Bejo Ucok	√	x	√	x	√	x	dst
2.		x	-	√	√	x	x	dst
3.								
4.								
5.								
6.								
7.								
dst.								
Mis Nil = 40								
Jumlah jawaban betul mis :		10	20	5	10	8	6	dst
Persentase jawaban betul		25	50	12,5	25	20	15	dst

- 4). Setelah dilakukan pendataan jawaban setiap peserta pada setiap nomor, selanjutnya dijumlahkan jawaban betul (R_{H} dan R_{L}) untuk setiap item dalam setiap kelompok dan seterusnya dihitung persentase jawaban betul (P_{H} dan P_{L}) untuk masing-masing nomor item dalam setiap kelompok dengan memakai rumus :

$$P_{H} = \frac{R_{H}}{N_{H}} \times 100 \%$$

P_{H} = Persentase jawaban betul untuk kelompok pandai (percentage high)

R_{H} = Jumlah jawaban betul untuk kelompok pandai (right high)

N_{H} = Jumlah semua peserta dalam kelompok pandai (27 % dari seluruh peserta).

Rumus diatas adalah untuk menghitung persentase pada kelompok pandai sedang pada kelompok lemah prinsip rumus adalah sama, yakni :

$$P_{L} = \frac{R_{L}}{N_{L}} \times 100 \%$$

P_{L} = Persentase jawaban betul untuk kelompok lemah (percentage lower)

R_{L} = Jumlah jawaban betul untuk kelompok lemah (right lower)

N_{L} = Jumlah semua peserta dalam kelompok lemah (27 % dari seluruh peserta).

Tabel 2.3 Jumlah Dan Persentase Tiap Jawaban Item Yang Betul Pada 2 Kelompok

No.soal	Kel. Pandai		Kel. Lemah	
	R_{H}	P_{H}	R_{L}	P_{L}
1.	35	87,5 %	10	25 %
2.	32	80 %	20	50 %
3.	35	87,5 %	5	12,5 %
4.	30	75 %	10	25 %
5.	25	62,5 %	8	20 %
6.	20	50 %	6	15 %
7.				
8.				
9.				
10.				
11.				
dst	dst	dst	dst	dst

5. Hasil hitungan langkah keempat dipakai sebagai dasar perhitungan pada format analisa butir soal.

Format analisis butir soal sekaligus untuk menentukan tingkat kesukaran butir soal disamping daya beda..

Tabel 2.4 Format Analisis Butir Soal

No.soal	Kel. Pandai		Kel. Lemah		Tingkat Kesukaran		Daya Beda	
	R ₁₁	P ₁₁	R _{1.}	P _{1.}	proporsi	%	proporsi	%
1.	35	87,5 %	10	25 %	p = 0,5625	56,25 %	d = 0,625	62,5 %
2.	32	80	20	50 %	p = 0,65	65 %	d = 0,30	
3.	35	87,5	5	12,5	p = 0,5			
4.	30	75	10	25	p =			
5.	25	62,5	8	20	p =			
6.	20	50	6	15	p =			
7.								
8.								
9.								
10.								
11.								
dst	dst	dst	dst	dst				

Untuk menghitung daya pembeda setiap item digunakan rumus sebagai berikut :

$$d = \frac{R_{11}}{N_{11}} - \frac{R_{1.}}{N_{1.}}$$

d = Indeks daya beda

R₁₁ = Jumlah jawaban betul untuk kelompok pandai (right hight)

N₁₁ = Jumlah semua peserta dalam kelompok pandai

R_{1.} = Jumlah jawaban betul untuk kelompok lemah (right lower)

N_{1.} = Jumlah semua peserta dalam kelompok lemah

Sebagai contoh untuk menghitung daya pembeda (d) dari item nomor 1 adalah sebagai berikut :

$$d = \frac{35}{40} - \frac{10}{40} = 0,625, \text{ dan item no.2} = 0,30$$

Daya beda maksimal tercapai apabila seluruh peserta dalam kelompok pandai dapat menjawab suatu item dengan benar (RH = NH), sedangkan peserta dalam kelompok lemah tidak seorangpun yang dapat menjawab dengan benar (RL = 0). Dalam hal ini harga d = 1 - 0 = 1.

Indeks daya beda sebesar 0 akan terjadi apabila proporsi jawaban betul dari kelompok pandai dan dari kelompok lemah sama besarnya, yaitu ketika indeks kesukaran bagi kelompok pandai sama besar dengan indeks kesukaran bagi kelompok lemah.

Secara matematik angka kisaran d mulai dari -1 sampai dengan +1, namun harga d yang positif saja yang memiliki arti, sedangkan nilai 0 saja sudah menunjukkan bahwa item yang bersangkutan tidak mempunyai daya diskriminasi, apalagi nilai negatif justru dapat memberikan informasi yang menyesatkan. Indeks daya beda yang ideal adalah sebesar mungkin mendekati angka 1 yang berarti item tersebut semakin mampu membedakan antara mereka yang menguasai bahan yang diujikan dengan mereka yang tidak menguasai bahan. Semakin kecil indeks daya beda (mendekati 0) berarti semakin tidak jelas fungsi item yang bersangkutan dalam membedakan mana subyek yang menguasai bahan pelajaran dengan subyek yang tidak tahu apa-apa.

Tingkat kesukaran (Difficulty Level)

Disamping daya pembeda, tingkat kesukaran butir soal tes perlu dihitung yang ditentukan dari semua peserta yang memberi jawaban betul pada kedua kelompok (pandai dan lemah). Rumus yang dipergunakan untuk menentukan tingkat kesukaran soal adalah :

$$p = \frac{n_i}{N}$$

p = Indeks kesukaran (dalam angka proporsi)

n_i = Banyaknya peserta yang menjawab item dengan benar

N = Banyaknya peserta yang menjawab item (kelompok pandai dan lemah)

Ada juga rumus yang digunakan dalam perhitungan persentase yaitu :

$$p = \frac{R_{11} + R_{12}}{N_{11} + N_{12}} \times 100 \% \quad \text{atau diringkas} \quad p = \frac{n_i}{N} \times 100 \%$$

p = tingkat kesukaran (dalam angka persentase)

$N_{11}, N_{12}, R_{11}, R_{12}$ adalah sama maksudnya dengan penjelasan dimuka.

Perhatikan bahwa dalam menghitung indeks kesukaran p tidak membedakan antara kelompok Pandai dengan kelompok Lemah. Seluruh peserta untuk perhitungan ini dijadikan satu.

Bila rumus tersebut dipergunakan dengan data yang terdapat pada contoh diatas dimana :

$$\text{Item (1) : } R_{11} = 35 ; R_{12} = 10$$

$$N_{11} = 40 ; N_{12} = 40$$

$$\text{Item (2) : } R_{11} = 32 ; R_{12} = 20$$

$$N_{11} = 40 ; N_{12} = 40$$

Maka :

Tingkat kesukaran item nomor (1) adalah :

$$p = \frac{R_{11} + R_{12}}{N_{11} + N_{12}} \times 100 \%$$

$$p = \frac{N_{11} + N_{1.}}{40 + 40} \times 100\% = 56,25\%$$

atau ;

$$p = \frac{n_i}{N}$$

$$p = \frac{45}{80} = 0,5625$$

Tingkat kesukaran item nomor (2) adalah :

$$p = \frac{R_{11} + R_{1.}}{N_{11} + N_{1.}} \times 100\%$$

$$p = \frac{32 + 20}{40 + 40} \times 100\% = 65\% \quad \text{atau} \quad p = \frac{52}{80} = 0,65$$

Semakin rendah nilai p (mendekati 0) justru semakin sukar bagi peserta untuk menjawabnya dengan benar, sebaliknya makin besar nilai p , yaitu mendekati 1, tingkat kesukaran item tersebut makin rendah atau kebanyakan peserta dapat menjawab dengan betul.

Bagaimana kaitan antara taraf kesukaran dengan daya beda item ? Item yang terlalu sukar sehingga tidak seorangpun dapat menjawabnya dengan benar yaitu $p = 0$, tidak akan mempunyai daya beda ($d = 0 - 0 = 0$), demikian juga apabila suatu item begitu mudahnya sehingga semua peserta dapat menjawabnya dengan benar, maka $d = 1 - 1 = 0$.

Daya beda yang baik pada umumnya terdapat pada item yang tidak terlalu mudah dan juga tidak terlalu sukar, yaitu apabila harga p berkisar antara 0,40 s/d 0,60, namun jangan pula menganggap secara otomatis bahwa setiap item yang nilai $p = 0,50$, pasti daya bedanya baik. Secara matematik nilai $d = 0$ pada kasus $p = 0,5$ dapat terjadi, namun secara logika hal itu sulit dibayangkan. Dengan demikian kualitas suatu item/daya beda tidak semata-mata ditentukan berdasarkan tingkat kesukaran item tersebut, namun sebaliknya tingkat kesukaran sudah barang pasti menentukan indeks daya beda.

6. Harus Comprehensive

Yaitu kemampuan tes itu untuk mengungkap gejala yang lebih luas dari variable yang diukur. Semakin banyak hal yang diungkap tes itu semakin baik.

7. Harus Praktis (Usebilitas)

Yaitu kemudahan dari segi penggunaan tes tersebut. Nilai suatu tes tergantung kepada kegunaannya.

Syarat usebilitas menyangkut masalah pelaksanaan/penggunaan tes itu yaitu berupa :

a. Biaya

Suatu tes yang mengeluarkan biaya yang banyak, misalnya pemakaian kertas-kertas yang mahal, soal yang sangat banyak sehingga memerlukan material yang banyak dalam pencetakannya, padahal aspek yang diungkapnya sempit sehingga tidak memenuhi keseimbangan biaya yang baik.

Demikian juga sebaliknya dengan test yang tujuannya penting sekali, misalnya untuk penyaringan (seleksi) pegawai, dicetak dengan bahan sederhana dengan tujuan efisiensi biaya sehingga tingkat validitasnya rendah, hal ini juga kurang memenuhi syarat praktis ditinjau dalam soal biaya. Tes yang memenuhi syarat usebilitas tentang biaya, haruslah sebanding biayanya dengan fungsi tes yang dilaksanakan.

b. Waktu

Lamanya pelaksanaan suatu tes berhubungan dengan tingkat kesukaran item serta banyaknya soal-soal yang disajikan. Tes yang pelaksanaan terlalu lama atau scoring yang terlalu lama begitu juga interpretasinya apalagi hanya untuk pengukuran yang sempit mempengaruhi usebilitas tes. Sangat dituntut kreativitas untuk membuat item tes seminim mungkin namun tetap mengungkap masalah yang luas.

c. Ruang/fasilitas yang dipergunakan.

Fasilitas yang dipergunakan dalam pelaksanaan tes kadang menuntut sedemikian rupa seperti ruangan besar, kursi, meja luas, alat-alat peraga dan lain-lain. Perlu untuk merencanakan agar pelaksanaan tes tidak terlalu rumit.

d. Kebudayaan

Suatu tes dari suatu daerah mungkin tidak cocok dengan daerah lain. Pada umumnya hal ini berpengaruh apabila latar belakang kebudayaan itu jauh berbeda. Item tes inteligensi dari Amerika misalnya sering tidak cocok dengan alam Indonesia. Untuk itu perlu diganti dengan item-item/istilah-istilah yang biasa dijumpai sehingga cocok dengan budaya setempat. Tentu sangat bijaksana jika dalam merancang item suatu tes terlepas dari pengaruh budaya, dalam arti tes tersebut diterima secara luas dalam komunitas yang heterogen.

BAB III PENGEMBANGAN ALAT TES

Berbagai jenis test dapat digunakan baik dalam kegiatan pendidikan maupun kegiatan diagnosis psikologis mulai dari tes tulisan, tes lisan sampai pada tes perbuatan. Masing-masing jenis tes ini memerlukan langkah dan patokan dalam perumusan butir soal dan penggunaannya. Sudah barang tentu kaidah-kaidah penulisan butir soal untuk masing-masing jenis dan bahkan bentuk tes adalah berbeda satu dengan lainnya.

A. Pengembangan Tes Pilihan Ganda

Prinsip-prinsip penyusunan butir soal

Jenis test tulisan yang sering digunakan adalah bentuk tes pilihan ganda. Pilihan ganda adalah salah satu bentuk dari jenis tes objektif yang menuntut peserta tes untuk dapat menentukan pilihan jawaban yang benar diantara alternatif jawaban yang disediakan. Salah satu keunggulan yang paling menonjol dibandingkan dengan tes objektif lainnya adalah bahwa bentuk tes pilihan ganda dapat mencakup *bahan yang luas dan kemungkinan menebak dengan benar lebih kecil*. Namun demikian dibalik itu penyusunannya membutuhkan waktu dan ketekunan disamping kemampuan dan keterampilan agar dapat menjamin mutu soalnya dengan baik. Ada beberapa kaidah penulisan butir soal yang perlu dipedomani yaitu :

1. Rumuskan dengan jelas pernyataan/pertanyaan dari item soal.

Contoh yang kurang baik :

Salah satu provinsi di Sumatera Utara.....

- a. Merupakan penghasil karet terbesar di Indonesia
- b. Berpenduduk paling banyak di Sumatera
- c. Mempunyai kebudayaan yang tinggi
- d. Mempertahankan adat istiadat dengan kuat

Permasalahan yang dipertanyakan dalam item tidak jelas, justru meragukan atau membingungkan subyek didik.

Contoh yang lebih baik :

Salah satu provinsi yang terbanyak penduduknya di Sumatera adalah....

- a. Sumatera Utara
- b. Sumatera Barat
- c. Jambi
- d. Sumatera Selatan

2. Kemungkinan Jawaban harus logis, baik isi maupun hubungannya dengan batang soal.

Contoh yang kurang baik :

Siapakah presiden yang pertama negara republik Indonesia ?

- a. Ir. Sukarno
- b. Hamzah Has
- c. Ferdinan Marcos
- d. George Buss

Option b tidak logis karena beliau bukan presiden dan saat ini adalah wakil presiden. Option C dan D terlalu jauh hubungannya dikaitkan dengan itemnya karena kedua nama tersebut tidak menggambarkan nama bangsa Indonesia.

3. Kemungkinan jawaban harus mendekati, baik dari segi fungsi maupun panjang pendeknya kalimat.

Contoh yang kurang baik :

Sesuai dengan teori usia bayi normal dalam kandungan adalah ...

- a. 30 minggu
- b. 36 minggu
- c. 70 minggu
- d. 110 minggu

Option c dan d tidak homogen karena usia bayi normal dikandung tidak ada yang selama itu atau dengan kata lain hal itu telah diluar batas waktu ibu mengandung.

Contoh yang lebih baik :

- a. 30 minggu
- b. 36 minggu
- c. 40 minggu
- d. 44 minggu

4. Kalau jawaban berbentuk bilangan, maka sebaiknya diurutkan dari besar ke kecil atau sebaliknya.

Urutan yang tidak teratur dapat membingungkan peserta test.

5. Hindarkan penggunaan pertanyaan yang bersifat negatif pada butir soal.

Contoh yang kurang baik :

Menurut teori pendidikan modern, pertumbuhan dan perkembangan manusia tidak dapat berlangsung dengan tanpa dipengaruhi ...

- a. Bawaan dan lingkungan
- b. Bawaan
- c. Lingkungan
- d. Pendidikan

Dengan menggunakan kata tidak dan tanpa dalam batang butir soal akan dapat membingungkan testee sehingga membutuhkan waktu untuk mengerjakan soalnya, akibatnya mempengaruhi reliabilitas soal.

Contoh yang lebih baik :

Menurut teori pendidikan modern, pertumbuhan dan perkembangan manusia dipengaruhi oleh.....

- a. Bawaan dan lingkungan
- b. Bawaan
- c. Lingkungan
- d. Pendidikan

6. Hindari penggunaan alternatif jawaban yang kemungkinannya "semua

jawaban salah” atau “semua jawaban benar”.

Contoh yang kurang baik :

Seseorang dapat dikatakan sukses dalam hidup karena memiliki...

- a. Usaha yang mapan
- b. Anak-anak yang Soleh dan cerdas
- c. Karir yang melejit
- d. Status sosial yang terpendang

Dengan kemungkinan semua jawaban diatas benar, tidak dapat dipantau apakah testee mengetahui jawaban yang salah dan sebaliknya. Bila butir soalnya menggunakan option yang semua jawaban diatas benar atau salah, sebaiknya digunakan dengan butir soal pengecualian, yaitu dengan membuatnya suatu option yang tidak benar dan selainnya benar atau sebaliknya dengan membuat hanya satu jawaban yang benar dan selainnya salah. Penjelasan ini ada kaitannya dengan kaidah berikut.

7. Beri tanda tertentu pada item yang menggunakan kata pengecualian

Apabila menggunakan kata-kata pengecualian (kecuali, selain) pada butir item maka kata tersebut sebaiknya diberi tanda seperti ; garis bawah, huruf besar, atau cetak miring agar testee memberi perhatian khusus sehingga tidak terjadi penyamarataan.

Contoh yang kurang baik :

Suatu teori dapat dikatakan teori Behaviorisme apabila, kecuali :

- a. Respon merupakan akibat dari stimulus
- b. Prilaku yang terbentuk hasil dari proses conditioning
- c. Reward memperkuat tingkah laku yang terbentuk
- d. Internal motif sebagai dasar tujuan tingkah laku

8. Hindarkan penempatan kata yang berulang-ulang pada alternatif jawaban.

Contoh yang kurang baik :

Rata-rata waktu yang disediakan untuk menjawab 1 item tes pilihan ganda adalah.....

- a. Kira-kira 30 detik
- b. Kira-kira 60 detik
- c. Kira-kira 2 menit
- d. Kira-kira 2,5 menit

Bila ada kata yang berulang-ulang pada kemungkinan alternatif jawaban sebaiknya ditempatkan pada item soal.

9. Hindarkan kemungkinan adanya lebih dari 1 alternatif jawaban yang benar.

Contoh yang kurang baik :

Sesuai dengan tuntutan cara belajar siswa aktif, maka peranan guru akan berbeda dibandingkan dengan pendekatan konvensional. Peran guru dalam pendekatan CBSA sayogianya adalah sebagai.....

- a. Fasilitator
- b. Motivator
- c. Informator
- d. Evaluator

Dari contoh diatas dapat dikatakan option a dan b benar sehingga dapat membingungkan peserta.

10. Pengecoh harus berfungsi

Contoh yang kurang baik :

Yang manakah dari factor-faktor berikut ini yang mempengaruhi perkembangan bakat seseorang ?

- a. Pengalaman
- b. Suku bangsa
- c. Tempat tinggal
- d. Jarak sekolah

Option b,c,d kurang berfungsi karena ketiga option tersebut sangat tidak berhubungan langsung dengan perkembangan bakat seseorang.

11. Hindari kata-kata yang lemah pada batang soal

Kata-kata yang lemah pada batang soal seperti :keanyakan, kadang- kadang, sering kali, pada umumnya, dan lain-lain merupakan ungkapan yang tidak tegas dan jelas yang berindikasi tidak menentu, sehingga dapat mengakibatkan testee akan ragu-ragu dalam memilih jawaban yang benar.

Contoh yang kurang baik :

Pada umumnya metode ceramah sangat tepat digunakan untuk.....

- a. Menyampaikan informasi langka
- b. Membuka pelajaran
- c. Melatih keterampilan
- d. Mempromosikan satu produk

12. Hindari keterkaitan jawaban antar butir soal

Jawaban satu soal item tertentu jangan sampai memberi bantuan pada testee untuk menjawab soal item lainnya.

Contoh yang kurang baik :

1. Bila suatu kendaraan kecepatannya 100 km per jam, berapa lama waktu yang dibutuhkan untuk menempuh 75 km ?

- a. 30 menit
- b. 45 menit
- c. 60 menit
- d. 75 menit

2. Apabila selama waktu tempuh tersebut dihabiskan supir 6 batang rokok, berapa batang rokok yang dihabiskan supir selama 15 menit pertama ?

- a. 1 batang
- b. 2 batang
- c. 3 batang
- d. 4 batang

13. Jawaban yang benar disebar secara proportional

Dalam merakit soal hendaknya diusahakan agar pilihan jawaban yang benar (kunci jawaban) tersebar diantara a, b, c dan d secara proporsional dan ditentukan secara random serta jangan sampai terjadi ada semacam pola tertentu dalam penyelesaian jawaban yang benar.

Kebaikan Tes Pilihan Ganda

1. Dapat dihindarkan kesalahan yang sering terdapat dalam menjawab tes Esai akibat dari kesalahan pengertian-terhadap pertanyaan
2. Dalam memilih kemungkinan jawaban yang benar murid dibimbing untuk berfikir dalam membedakan kemungkinan-kemungkinan jawaban, apalagi kemungkinan jawaban yang benar itu hanyalah satu diantara pilihan lainnya.
3. Dimungkinkan untuk pengukuran bahan materi yang lebih luas dan mendalam, karena untuk satu butir soal dapat mencakup materi yang lebih banyak.
4. Seperti halnya tes objektif lainnya, kebaikan tes ini berupa :
 - a. Pelaksanaan tes memerlukan waktu yang relatif singkat.
 - b. Pemeriksaan jawaban dan pemberian skor lebih cepat dan objektif.
 - c. Kepercayaan terhadap pemakaian tes lebih tinggi dibandingkan dengan tes esai.
 - d. Cocok digunakan bagi peserta tes yang banyak.
 - e. Dapat mengevaluasi bahan pengetahuan yang mendalam, karena dimodifikasi dengan bentuk yang agak rumit.

Kelemahan Tes Pilihan Ganda

1. Cenderung untuk diterka-terka peserta jawabannya, khususnya bagi peserta yang kurang menguasai bahan atau peserta yang tingkat kemampuannya rendah.
2. Penyiapan tes memerlukan waktu yang banyak, serta menuntut kemampuan tertentu dalam merumuskannya.
3. Biayanya relatif banyak bila dibandingkan dengan tes esai, karena harus distensil/cetak berlembar-lembar.

B. Pengembangan Tes Benar-Salah

Tes Benar-Salah (True False Question) adalah salah satu bentuk tes objektif yang bentuknya merupakan serangkaian pernyataan yang harus ditentukan responden atau murid apakah pernyataan tersebut benar atau salah.

Tes ini bukanlah berbentuk pertanyaan tetapi pernyataan, karena yang diminta dari responden hanya reaksi "Benar" atau "Salah" saja. Oleh karenanya bentuk tes ini mempunyai tingkatan terkaan yang sangat tinggi disbanding bentuk tes lainnya. Hal ini menuntut agar penggunaan kata/kalimat harus benar-benar dipertimbangkan. Untuk mengevaluasi tujuan pengajaran yang berorientasi penanaman akan "pemahaman" lebih cocok dipergunakan bentuk tes ini dibandingkan dengan tujuan pengajaran yang menanamkan "pengetahuan". Untuk

tujuan pengajaran “aplikasi, analisa, dan evaluasi” yang bertaraf tinggi, bentuk tes ini juga kurang cocok dipergunakan.

Beberapa petunjuk dalam mengkontruksi soal Tes Benar-Salah :

1. Hindari penggunaan kata-kata yang bermakna “lemah atau kuat”

Kata-kata yang mempunyai konotasi “kuat/keras” seperti : selalu, tidak pernah, semuanya, tanpa kecuali, harus, dan sejenisnya akan mengakibatkan pernyataan yang sangat nyata yang akan mengarahkan pilihan jawaban responden kepada jawaban “salah”. Sebaliknya kata-kata yang mempunyai konotasi “lemah” seperti : kadang-kadang, pada umumnya, sebahagian, sebaiknya, memungkinkan, dan sejenisnya akan membimbing responden untuk cenderung memilih jawaban “benar”.

Contoh yang kurang baik :

“Teori psikoanalisis menganggap setiap bentuk perilaku selalu terkait dengan libido seksual”

“Manusia pada umumnya cemas akan kehidupan”

2. Hindari pernyataan yang meragukan

Kata-kata atau pernyataan yang kurang tegas kebenarannya atau kesalahannya dapat menimbulkan reaksi protes dalam diri responden, sehingga responden merasa tidak cukup hanya memberi jawaban benar atau salah. Pengertian yang bermakna seni, nilai filosofis yang cenderung berpatokan pada standar nilai instrinsik kurang tepat digunakan dengan bentuk tes ini.

Contoh yang kurang baik :

“Kegagalan merupakan awal dari keberhasilan”

“Hidup di kota lebih baik dari hidup di desa”

3. Hindari pernyataan yang ketepatannya kurang didasari “fakta, konsep, atau prinsip yang berlaku.

Soal atau pernyataan yang fakta atau konsepnya meragukan atau kurangnya bukti empiris yang dapat diukur sesuai dengan kenyataan akan dapat memancing reaksi mendatar/pesimis dari responden untuk menentukan jawaban. Prinsip yang terkandung dalam pernyataan juga harus prinsip yang berlaku/nyata dan mengandung satu pengertian.

Contoh yang kurang baik :

“Tangis bagi seorang bayi yang lahir adalah tanda kekurangan senangan lepas dari buaian rahim ibu yang nyaman”

“Penduduk kota lebih cepat peningkatannya dari penduduk didesa, karena kondisi kelahiran bayi yang lebih sehat dan usia hidup yang lebih lama”.

4. Hindari soal yang bahan kutipan mentah/langsung dari buku pelajaran

Butir pernyataan yang pembuatannya merupakan kutipan langsung dari buku bacaan/pelajaran tanpa ada perubahan bentuk atau pengembangannya akan

memudahkan responden dalam memberi alternatif jawaban dan cenderung sudah terpola dengan jawaban sebelumnya.

Contoh yang kurang baik :

Teori tabula rasa beranggapan bahwa anak ibarat kertas putih.

5. Sesuaikan penggunaan bahasa dengan kemampuan responden
Bahasa yang dipergunakan dalam merumuskan pernyataan soal hendaknya disesuaikan dengan tingkat kemampuan bahasa dari pada responden, sehingga tidak terjadi pemaknaan yang mungkin akan mengakibatkan kesalahan pengertian. Dengan kata lain kesulitan soal hendaknya jangan dipertautkan dengan kesulitan makna dari kata-kata.
Contoh yang kurang baik :
"Kiat sukses terletak pada motivasi"
Bila diberikan pada siswa SD sebaiknya gunakan :
"Rajin pangkal pandai, malas pangkal bodoh"
6. Hindarkan pertautan antar soal pernyataan
Setiap butir pernyataan soal harus berdiri sendiri sehingga butir pernyataan yang satu tidak berhubungan dengan butir pernyataan yang lain.
Contoh yang kurang baik :
 1. Carl. G. Jung adalah salah satu tokoh psikoanalisis
 2. Ia menyebutkan adanya jiwa kolektif
7. Hindarkan item tertentu membantu jawaban item lain
Sesuatu item jangan memberi isyarat atau membantu memberikan jawaban kepada item berikutnya.
Contoh yang kurang baik :
 1. Penyakit bolokokok disebabkan oleh kuman
 2. Penularan virus bolokokok di Indonesia berasal dari Imigran lembu Australia.
8. Hindarkan jawaban yang tersusun menurut pola tertentu
Bila soal/item memiliki jawaban yang tersusun atas bentuk atau pola-pola tertentu baik berupa jawaban yang benar maupun jawaban yang salah, maka responden akan cenderung mengikuti pola-pola tersebut dalam menjawab pertanyaan berikutnya.
Misalnya : Item 1 & 2 benar
Item 3, 4 & 5 salah
Item 6 & 7 benar
Item 8, 9 & 10 salah

Kebaikan Tes Benar-Salah

1. Dalam menjawab pertanyaan responden terbimbing untuk meningkatkan taraf berpikir secara kritis. Hal ini dimungkinkan karena responden harus menilai makna pernyataan secara hati-hati sehingga ia dapat membenarkan pernyataan atau menyatakan pernyataan salah.

2. Pokok masalah yang dapat digarap bisa mencakup lebih dari satu masalah. Hal ini memungkinkan dibuatnya soal-soal yang lebih luas yang permasalahannya lebih heterogen.
3. Dapat mengungkap pemahaman
Dalam menjawab tes benar-salah responden memerlukan pemahaman beberapa informasi tentang pokok masalah yang dipertanyakan. Hal ini akan tercapai apabila isi pernyataan bukan hanya meminta data singkat tetapi meminta pemahaman tentang fakta atau konsep.
4. Seperti halnya dengan bentuk tes obyektif lainnya tes ini mempunyai kebaikan sebagai berikut :
 - a. Memerlukan waktu yang ringkas untuk menjawabnya, sehingga dapat dibuat banyak soal dengan materi yang luas
 - b. Penyekoran yang lebih cepat dan obyektif
 - c. Kecocokan dan tingkat kepercayaan terhadap tes lebih tinggi
 - d. Cocok dilakukan bila peserta tes cukup banyak

Kelemahan tes Benar-Salah

1. Responden cenderung menerka jawaban bila suatu item tidak bisa dijawab
2. Merumuskan tesnya memerlukan waktu yang banyak
3. Sukar diterapkan untuk mengukur proses mental atau taraf berpikir yang bertaraf tinggi dan kompleks
4. Biayanya relatif lebih besar disbanding dengan tes esai

C. Pengembangan Tes Menjodohkan (Matching Items)

Tes menjodohkan adalah tes yang terdiri dari dua kolom dimana setiap batang butir soal yang terdapat dalam kolom pertama dapat diisi pasangannya/jawabannya dari kemungkinan-kemungkinan jawaban yang disajikan pada kolom kedua.

Bentuk tes menjodohkan dipergunakan untuk mengungkap kemampuan yang bersifat "ingatan" dan "pemahaman dangkal" dan kurang cocok untuk tujuan mengungkap kemampuan analisa, aplikasi dan evaluasi.

Beberapa petunjuk yang perlu dipedomani dalam menyusun tes ini :

1. Butir soal dan jawaban sesingkat mungkin
Butir soal yang terdapat pada kolom pertama bukan merupakan masalah-masalah yang diuraikan dengan kalimat-kalimat yang panjang. Demikian juga alternatif jawaban yang terdapat pada kolom dua merupakan data singkat misalnya : nama-nama orang atau binatang, tahun, dan lain-lain.
2. Harus memiliki isi masalah yang homogen
Keseluruhan butir soal pada satu kelompok tes menjodohkan bukan garapan masalah yang luas, tetapi masalah yang lebih seragam. Apabila dimaksudkan untuk menguji pengetahuan responden tentang nama pengarang dari beberapa judul karangan maka pada kolom pertama berisi "judul-judul karangan" sedang pada kolom kedua berisi "nama-nama pengarangnya".

3. Jumlah alternatif jawaban lebih banyak dari butir soal
Hendaknya dilebihkan 1 atau 2 alternatif jawaban pada kolom kedua dari jumlah alternatif jawaban yang harus ada.
4. Hindarkan jangan sejajar antara butir soal dengan jawabannya
Penempatan/urutan kemungkinan jawaban yang terdapat pada kolom kedua diatur sedemikian rupa agar tidak sejajar dengan masalah yang terdapat pada kolom pertama.
5. Idealnya satu kelompok berisi 5 sampai 7 butir soal
Bila materi yang akan diuji terlalu banyak atau melebihi 7 sebaiknya dibuat dua kelompok tes menjodohkan. Hal ini sebenarnya jarang terjadi mengingat butir soal dalam setiap tes menjodohkan harus mempunyai isi yang homogen.
6. Idealnya semua jawaban dimungkinkan untuk semua soal
Untuk tiap soal yang terdapat pada kolom pertama mempunyai kemungkinan untuk dijawab dengan daftar jawaban yang terdapat dalam kolom dua.

Kebaikan Tes Menjodohkan

Cocok untuk materi uji yang banyak tetapi mempunyai isi yang homogen dan mempunyai kesamaan jenisnya.

Misalnya bila tes yang digunakan untuk menilai kemampuan murid untuk menyebut ibu kota propinsi di Indonesia, maka akan diperoleh pengulangan-pengulangan kalimat pertanyaan bila disajikan dalam bentuk tes esai atau tes benar salah atau tes pilihan ganda.

Kelemahan

1. Tingkat Probabilitas menjawab dengan benar tidak sama untuk setiap soal.
Bila murid telah pasti dapat menjawab item 1 dan 2 dari 6 butir soal yang ada, maka kemungkinan jawaban untuk item 3 dan 4 bukan lagi 6 kemungkinan, tetapi tinggal 4 kemungkinan
2. Gampang dikuasai oleh responden
Pengertian yang homogen sering ditafsirkan oleh penyusun tes dalam arti yang lebih sempit sehingga penyusun tes tanpa disadari membimbing responden untuk memberi jawaban yang benar walaupun responden kurang menguasai bahan yang ditanyakan.
Contoh :

Klasifikasi	Contoh Hewan
1. Hewan Ikan	Burung pipit
2. Hewan burung	Ikan mas

D. Pengembangan Tes Melengkapi

Tes melengkapi (Completion Test) yaitu tes yang mengharuskan peserta melengkapi atau mengisi titik-titik yang ada dalam butir soal dengan jawaban berupa kata, fakta atau konsep singkat.

Beberapa petunjuk dalam penyusunan butir soal :

1. Butir soal harus berupa pernyataan yang tidak lengkap.
Contoh : Pulau Terdapat di Danau Toba.
2. Rumuskan butir soal dengan jelas dan singkat
3. Hindarkan pengutipan langsung dari buku pelajaran
4. Jawaban yang disediakan harus berupa fakta, data atau konsep singkat, bukan merupakan kata-kata biasa.

Contoh salah :

Presiden adalah yang mengurus kehidupan berbangsa.

5. Jawaban bukan merupakan kalimat-kalimat yang panjang.

Hendaknya dibedakan tes melengkapi dengan tes uraian

Contoh yang salah :

Perbuatan baik berguna untuk dan

Law of effect dalam teori belajar adalah

Kebaikan tes melengkapi

1. Baik dilakukan pada murid sekolah dasar, karena kemampuan bahasanya masih sederhana.
2. Dapat mengurangi jumlah butir soal

Contoh :

Singamangaraja XII adalah pahlawan..... yang menentang wafat pada tahun dimakamkan sekarang di

Kelemahan Tes Melengkapi.

Tidak dapat mengukur kemampuan yang kompleks. Yang diukur terbatas pada kemampuan ingatan, sedangkan kemampuan proses berpikir yang lebih kompleks seperti problem solving sangat sukar dilakukan.

E. Pengembangan Tes Esai

Tes Esai (Essay Examination) adalah bentuk tes yang meminta jawaban-jawaban yang bersifat menguraikan, menerangkan, menunjukkan bagaimana sesuatu masalah terjadi atau menilai sesuatu hal/kejadian yang dilaksanakan secara tertulis baik penyajian soal ataupun jawaban.

Perumusan soal yang berkurang tepat akan mengakibatkan murid ragu dalam memberikan jawaban baik ketepatan yang diminta oleh soal, hal ini tentunya akan mengurangi sifat objektivitas penilaian.

Contoh :

Terangkanlah bagaimana terjadinya asimilasi !

Untuk jawaban dari pernyataan ini murid mungkin akan menjawab

- Murid A : Asimilasi terjadi pada siang hari
- Murid B : Proses asimilasi ialah pemasukan makanan pada tumbuh -- tumbuhan menjelaskan proses kimianya sekaligus.

Tentunya, kedua jawaban murid tersebut ada kebenarannya. Apakah guru

memberi nilai sama pada masing-masing murid ? Kalau diberikan nilai yang sama mungkin dipihak guru tidak setuju karena jawaban murid B lebih lengkap. Namun, bila di berikan nilai berbeda (lebih tinggi pada murid B) timbul pertanyaan berikut : apakah guru yakin bahwa murid A tidak memberi jawaban yang lengkap bila murid A ditanyakan secara langsung.

Starch dan Eliot melakukan eksperimen sebagai berikut : mereka mengirimkan pekerjaan seorang murid kepada 116 orang Guru ilmu pasti dalam berbagai sekolah menengah untuk memberi nilai (pekerjaan murid adalah sebagai jawaban dalam tes esai dalam mata pelajaran ilmu pasti). Nilai yang di pergunakan dari 0 s/d 100. Ternyata nilai yang di berikan guru – guru berbeda beda (nilai terendah 28 dan nilai tertinggi 92), dan dari nilai yang diberikan oleh 116 orang guru tersebut hanya 6 orang saja yang memderikan nilai yang sama. Penilaian ini menunjukkan bagai mana sukarnya mendapatkan penilaian yang objektif dan eksak dalam tes esai.

Contoh di atas adalah penilaian terhadap mata pelajaran ilmu pasti. Bagaimana pula sukarnya bila penilaian itu terhadap jawaban dalam bidang ilmu sosial. Semuanya menunjukkan kepada perlunya kehati-hatian kita dalam menyusun soal soal tes esai. Tes esai sudah lama di pergunakan namun diketahui pemakaian tes esai banyak sedikit digeser pemakaian tes objektif. Namun perlu diingat bahwa pemakaian tes esai harus /tidak mungkin terhapus seperlunya.

Tes Esai mempunyai kebaikan-kebaikan di samping kelemahannya. Untuk kriteria ujian yang menghendaki penilaian terhadap proses mental yang kompleks serta pemahaman yang mendalam terhadap sesuatu materi maka tes esai adalah lebih tepat digunakan.

Secara terperinci prinsip kontruksi tes esai dapat di sebut sebagai berikut :

1. Gunakan untuk mengevaluasi proses berpikir yang lebih kompleks.

Tes esai cocok dipergunakan apabila jawaban atas pertanyaan yang dimajukan memerlukan proses mental kompleks. Yang dimaksud proses mental kompleks dalam hal ini adalah bahwa dalam menjawab pertanyaan diperlukan hal-hal yang mencakup pengetahuan, pengertian, aplikasi, analisa, sintesa serta evaluasi. Semakin banyak ranah yang diminta, semakin tinggi proses mental yang diperlukan.

1. Gunakan pada mata pelajaran yang memetingkan kematangan proses berpikir.

Tes esai cocok di pergunakan bagi mata pelajaran yang bertujuan / mementingkan kematangan proses berpikir. Dengan perkataan lain bahwa mata pelajaran itu mempunyai pokok/sub pokok bahasan yang satu mempunyai hubungan yang erat atau merupakan dasar/penunjang kepada pokok/sub pokok bahasan yang diatasnya. Jadi, yang di pintingkan adalah proses berpikirnya bukan kemampuan akhir dari proses berpikirnya tersebut. Mata pelajaran seperti ini terdapat pada tingkatan sekolah yang lebih tinggi misalnya di SLTA atau di Perguruan Tinggi.

3. Gunakan untuk peserta yang kemampuan berbahasanya tinggi.

Tes esai cocok di pergunakan untuk apabila perbendaharaan dan kemampuan berbahasa dari yang di uji memadai/cukup tinggi. Hal tersebut perlu di perhatikan mengingat bahwa tes esai memerlukan jawaban yang bersifat menguraikan, menganalisa, membandingkan

perbendaharaan bahasa dan kemampuan berbahasa yang memadai/ cukup tinggi, hal ini tentunya belum banyak di punyai anak Sekolah Dasar.

4. Gunakan bila tenaga waktu cukup tersedia

Pemakaian tes esai harus mempertimbangkan ratio antar jumlah peserta ujian, waktu pemeriksaan yang tersedia dan jumlah pemeriksa jawaban.

Baik tidaknya suatu ujian di nilai dari objektivitasnya. Dengan demikian walaupun tes esai cocok di pergunakan namun efktivitas penilaian tersebut tidak akan sepenuhnya tercapai.

5. Penyusunan soal hendaknya mempedomani hal-hal berikut :

- a. Pertanyaan-pertanyaan direncanakan khusus meneliti hasil belajar yang diberikan. Untuk penyusunan pertanyaan itu hendaknya terlebih dahulu di periksa/disesuaikan dengan catatan pokok-pokok pikiran dan penegasan-penegasan sewaktu mengajar. Dalam hal ini untuk guru yang baik di mana selalu membuat satuan pelajaran (SP), catatan sewaktu mengajar dan satuan pelajaran adalah sumber informasi yang sangat di perlukan. Disamping itu sumber lain buku paket atau buku sumber yang dianjurkan di miliki oleh murid.
- b. Pertanyaan-pertanyaan hendaknya diarahkan pada dasar mayor learnings dan minor learnings. Pengertian tambahan boleh juga turut di tes tetapi bukan merupakan masalah pokok.
- c. Pertanyaan yang memerlukan proses mental kompleks dan pengertian-pengertian penting/sukar, harus dibedakan dari pertanyaan yang memerlukan kemampuan yang sederhana dan pengetahuan sempit. Perbedaan pertanyaan ini di hubungkan dengan bobot penilaian, dan waktu yang di sediakan untuk mengerjakan soal.
- d. Rumuskan pertanyaan secara jelas dan terbatas mengenai arti dan pentingnya persoalan.
Untuk hal ini hendaklah dipilih kata tanya yang tepat dan rumusan isi persolan yang ditanyakan, misalnya :
Mengapa terjadi gempa ?
Bagai mana semboyan bangsa Indonesia”
”Apakah yang terjadi jika murid dalam kelas saling curiga mencurigai?”
Pertanyaan di atas itu meragukan, sebaiknya adalah :
”Sebutkan 2 macam penyebab terjadinya gempa !”
”Tuliskan bunyi semboyan bangsa Indonesia !”
” Sebutkan akibat dari jika murid-murid dalam kelas saling curiga mencurigai terhadap ketertiban dan keamanan kelas”
- e. Hindari kata-kata yang bersifat sangat subjektif
misalnya : selalu, mungkin, adakalanya, kadang-kadang, biasanya, beberapa dan lain-lain.
Contoh yang salah :
”Sebutkan beberapa orang,Pahlawan Nasional !”
Contoh yang benar :
”Sebutkan sedikitnya 3 orang pahlawan Nasional !”

- f. Rumuskan kalimat pertanyaan itu sependek mungkin dan dengan kata-kata yang mudah di mengerti
 Contoh yang salah :
 "Sebutkanlah akibat dari apabila penduduk desa banyak menebangi pohon-pohonan di hulu sungai sehingga hampir mengalami kegundulan pada musim penghujan!"
- g. Sesuaikan kalimat dengan kemampuan bahasa yang di tes.
 Tester dituntut untuk dapat mengidentifikasi kemampuan bahasa lisan dan tulisan dari murid yang di tes, terutama pada murid-murid Sekolah dasar. Kemampuan bahasa lisan murid Sekolah Dasar tidak banyak berbeda dengan kemampuan bahasa tertulisnya. Dengan demikian kalimat pertanyaan itu harus disesuaikan dengan bahasa anak-anak sehari-hari.

Kebaikan dan kelemahan tes esai

Secara tidak langsung membahas tentang prinsip konstruksi tes esai diatas sebenarnya telah disinggung tentang kebaikan dan kelemahan tes esai itu. Berikut ini akan di uraikan beberapa kebaikan dan kelemahan tes esai itu. Berikut ini akan di uraikan beberapa kebaikan dan kelemahan tes esai di bandingkan dengan bentuk tes yang lain terutama dengan tes objektif, yaitu :

Kebaikan Tes Esai

1. Murid mempunyai kebebasan untuk memberikan jawaban terhadap pertanyaan berdasarkan hasil pemikirannya. Hal ini dapat di mungkinkan karena tujuan tes esai ini pada umumnya ditujukan untuk melihat kemampuan analisa, aplikasi dan evaluasi dari siswa terhadap suatu masalah. Dengan demikian apabila tujuan tes sekedar untuk meneliti ingatan anak terhadap sesuatu data, maka pemakaian tes kurang tepat.
2. Murid terlatih mengeluarkan pendapat dengan menggunakan bahasa sendiri. Kebaikan kedua ini adalah merupakan kegunaan ganda (nilai tambah) dari penggunaan tes esai. Disamping menilai pencapaian tujuan pengajaran juga murid terlatih dan terbiasa mengeluarkan pendapat dengan menggunakan pendapat dengan bahasa sendiri, yang ini tentunya juga meningkatkan kemampuan mentalnya.
3. Mempersiapkan tes esai dan menyelenggarakan lebih mudah dan tidak memakan waktu yang banyak dan dana yang banyak. Waktu pelaksanaan tes biasanya sekitar 90 menit s/d 180 menit. Karena tes esai memerlukan uraian, perbandingan dan sebagainya maka soal yang di majukanpun tidak banyak (mungkin sekitar 7 soal) menyusun sebanyak itu sudah barang tentu jauh lebih sedikit waktunya bila harus menyusun soal objektif sebanyak 80 soal untuk waktu sekitar 100 menit. Dalam penyampaian soalpun cukup didiktekan saja (biayanya sedikit/tidak ada) sedang untuk tes objektif di perlukan biaya pencetakannya.
4. Ruang yang di perlukan untuk pelaksanaan tes esai tidak terlalu menuntut ruangan yang luas, karena saling bekerjasama antara murid tidak semudah tes objektif.

Kelemahan Tes Esai

1. Karena soal yang di susun hanya sedikit (beberapa pertanyaan saja) maka tes esai ini kurang dapat di lakukan sebagai alat pengukur hasil belajar murid terhadap ruang lingkup yang lebih luas.

2. Karena sukarnya merumuskan kalimat pertanyaan yang jelas, tegas dan terbatas serta besarnya jarak kemampuan pengetahuan dan kemampuan berbahasa antara guru (pembuat tes) dan murid (penjawab tes) maka murid sering tidak dapat menangkap maksud dari pertanyaan atau salah memberi jawaban atas pertanyaan atau salah memberi jawaban atas pertanyaan yang di sajikan.
3. Penilaiannya sering subjektif. Hal yang diakibatkan oleh sukarnya memberi patokan nilai bagi jawaban yang berbeda-beda, banyaknya pekerjaan yang akan di nilai, serta beragamnya bentuk / corak tulisan dari murid-murid.
4. Pemberian nilai hanya dapat dilakukan oleh penyusun soal mengingat guru tersebutlah yang memahami benar akan maksud pertanyaan tersebut tentang ketetapan, dan kedalaman jawaban yang di inginkan.

Kebaikan dan kelemahan dari tes esai yang disebutkan diatas itu akan bisa bertambah apabila pemakaian tes esai di maksudkan untuk murid Sekolah Dasar. Untuk mengatasi rendahnya proses berpikir murid/anak dan kurangnya kemampuan bahasa tulis murid/anak, maka bagi murid/anak Sekolah Dasar dapat dipakai bentuk tes mengisi titik (Completion Test), untuk menggantikan tes esai.

Memang sebenarnya bagi beberapa orang Completion Test digolongkan pada bentuk objektif, karena data isian hanya berupa data singkatsaja, namun karena adanya kebebasan memberi jawaban maka completion test bagi beberapa ahli digolongkan pada bentuk tes esai.

F. Pengembangan Tes Lisan

Pada dasarnya tes lisan sama dengan tes uraian. Perbedaan keduanya hanya terletak pada penyajiannya. Dalam penyajian tes lisan dilakukan dalam komunikasi langsung antara penguji dengan yang diuji. Tes lisan pada umumnya dilakukan untuk mengevaluasi hasil belajar berupa kemampuan untuk mengemukakan pendapat-pendapat atau gagasan-gagasan secara lisan.

Beberapa hal yang perlu di perhatikan dalam pembuatan Tes antara lain :

1. Rencana terlebih dahulu pokok-pokok yang akan di tanyakan
Hal ini di maksudkan agar pelaksanaan ujian benar-benar tertuju pada pokok persoalan yang akan diuji di samping itu berguna agar setiap yang diuji mendapat pokok uji yang sama, termasuk rencana waktu ujian.
2. Sampaikan pertanyaan dengan cara yang baik
Ujian lisan bukan melisankan pertanyaan tetapi berdialog yang bertujuan meneliti sampai dimana pemahaman yang teruji pada suatu pokok ingin di teliti.
3. Ciptakan " Raport " sebelum memulai ujian.
Pada umumnya setiap yang diuji sebelum ujian sudah di hinggapi rasa was-was dan takut. Untuk penguji hendaknya terlebih dahulu menghilangkan rasa was-was dan takut itu, misalnya terlebih dahulu berbicara hal-hal yang bersifat umum dan pribadi.
4. Dahulukan pertanyaan yang mudah di perkirakan dapat di jawab.
Hal ini dimaksudkan agar si teruji mempunyai rasa percaya diri merupakan motivasi untuk dialog tanya jawab yang selanjutnya.

5. Hindarkan kata-kata yang bersifat langsung menyalahkan jawaban. Hal ini dimaksud agar si teruji tidak merasa tertekan harus di ingat bahwa tujuan pokok ujian lisan adalah meneliti pemahaman mendalam seperti kongnitip alikasi atau evaluasi.
6. Sebaiknya penilaian di berikan segera setela ujian di perbaiki. Adalah kurang tepat apabila si penguji secara terang-terangan memberikan nilai ujian selama ujian berlangsung. Nilai yang rendah akan menimbulkan rasa tertekan si teruji.
7. Formasi tempat duduk sebaiknya tidak berhadapan langsung. Hal ini di maksudkan untuk menghindari perasaan " teradili " yang dapat menimbulkan rasa enggan atau takut bila posisi duduk antara penguji dengan teruji berhadapan secara vertikal.

Keunggulan Tes Lisan

1. Bila di maksudkan untuk mengukur proses berpikir yang kompleks tes lisan sangat cocok di pakai. Misalnya seperti dalam ujian dalam mempertahankan suatu karya ilmiah.
2. Memungkinkan untuk mengecek jawaban tes tertulis urain, sering jawaban si siswa kurang dapat di pahami si pemeriksa jawaban ujian. Hal ini dapat di atasi dalam tes lisan misalnya dengan makna jawaban siswa dapat lebih objektif dan adil.
3. Tidak ada kesempatan untuk tidak jujur. Menyontek atau minta bantuan pada teman seperti sering terjadi pada tes tertulis, sehingga penilaian semakin objektif.

Kelemahan Tes Lisan

1. Ketidakadilan mudah terjadi. Sikap yang teruji banyak mempengaruhi sikap penguji. Bila yang diuji menunjukkan ketidak sopanan yang menguji cenderung mempersulit pertanyaan dan memperbanyak waktu dan sebagainya.
2. Penguji sering menyimpang dari persoalan/pokok yang diuji. Atas pertanyaan yang di berikan sering tidak dijawab sesuai dengan masalah yang dimaksudkan dan mungkin akan membawa si penguji untuk meneruskan pertanyaan yang sesuai dengan jawaban yang melantur itu.
3. Waktu pelaksanaan ujian sering tidak terstandar.
Ada peserta ujian yang menggunakan waktu terlalu banyak atau terlalu sedikit. Bagi si teruji yang agak lancar penguji cenderung memperdalam pertanyaan yang sebenarnya tidak sesuai dengan rencana, demikian juga pada si teruji yang kurang menguasai materi yang di uji, penguji cenderung mempersingkat waktu karena si penguji memperkirakan pertanyaan yang berikut tidak mungkin di jawab, pada hal belum tentu demikian halnya.
4. Memerlukan banyak formasi penyampaian.
Bagi pesertai yang banyak seharusnya si penguji harus mempersiapkan beberapa format pertanyaan untuk menghindari pertanyaan yang sama untuk setiap orang. Hal ini menuntut persiapan dan waktu yang banyak bagi si penguji.
5. Peluang subjektivitas dalam penilaian lebih terbuka.

Sipenguji cenderung membuat penilaian pada akhir ujian sehingga sipenguji mungkin lupa, atau bila siteruji dapat menarik perhatian atau simpati si penguji, penilaian si teruji cenderung baik tidak sesuai dengan tingkat kebenaran jawaban.

G. Pengembangan Tes Tindakan

Tes tindakan dimaksudkan untuk mengukur keterampilan siswa dalam melakukan sesuatu kegiatan dalam tes tindakan, persoalan disajikan dalam bentuk tugas yang harus di kerjakan oleh yang diuji adalah hasil belajar psikomotor materi ujian di tujukan pada proses pelaksanaan kegiatan hasil akhir (produk) dari kegiatan. Kecenderungan pengujian adalah penilaian hasil akhir (produk) pada hal proses kegiatanpun sangat penting untuk di nilai.

Berikut ini berupa petunjuk dalam pelaksanaan tes tindakan

1. Uraian kegiatan yang akan di peraktekan/dilakukan
Tugas yang akan di lakukan/di uraikan responden dalam bentuk kegiatan yang terperinci. Hendaknya ujian kegiatan itu di bedakan antara yang pokok dengan yang pendamping. Kegiatan yang pokok menentukan hal yang penting diamati, hal ini dimaksudkan agar penilaian yang lebih representatif.
2. Petunjuk melakukan kegiatan hendaknya di buat.
Petunjuk yang di buat itu haruslah tersusun dengan sistematis baik langkah-langkahnya maupun materi yang di tuntut.
3. Tentukan perilaku yang akan diukur dalam pedoman pengamatan.
Unsur kegiatan dan penilaian yang di ukur di susun secara logis. Untuk memudahkan pengamatan, hendaknya di bedakan antara urutan langkah-langkah dengan urutan pentingnya unsur kegiatan.
Adalah sangat baik apabila pedoman pengamatan yang sekaligus dengan pedoman penilaian ini diketahui oleh siteruji agar dia dapat merencanakan kegiatannya.
4. Rencanakan alat perlengkapan
Alat perlengkapan hendaknya tersedia sebelum tes dilakukan. Alat perlengkapan dimaksud mungkin harus disediakan oleh penguji atau siteruji. Bila alat perlengkapan harus disediakan oleh siteruji, hendaknya penjelasan tentang jenis alat, mutu dan lain-lain harus disampaikan pada siteruji. Untuk menghindarkan keragaman alat yang memungkinkan berpengaruh pada pelaksanaan tindakan yang akhirnya menyulitkan penilaian yang obyektif, sebaiknya alat perlengkapan disediakan oleh penguji.

Keunggulan tes tindakan

1. Sangat tepat dilakukan untuk mengukur perilaku psikomotor.
2. Kesempatan menyontek tidak ada

Kelemahan tes tindakan

1. Penilaian sangat sulit
Dalam mengamati tindakan memerlukan kemampuan yang tinggi selain mengingat. Tindakan yang dilakukan, dan criteria yang dituntut dalam tindakan sangat kompleks. Juga jumlah penilaian/tindakan yang diamati cukup banyak.

2. Biaya relatif besar

Karena kegiatan yang diperlukan memerlukan alat perlengkapan sudah barang tentu memerlukan biaya yang cukup besar.

3. Waktu relatif lama

Tes tindakan sukar dilakukan pada peserta yang banyak secara serempak, oleh karena itu tes tindakan sebaiknya dilakukan secara individual yang hal ini mengakibatkan tuntutan waktu yang lama.

Lembar Pengamatan Tes Tindakan						
Nama :						
No. Peserta :						
Bidang Studi :						
Tugas :						
No.Urut	Aspek yang dinilai	Nilai				
		1	2	3	4	5
1. 2. dst						
Catatan :						
Nilai :						
Penguji Mulia Siregar						

RANGKUMAN

1. Setiap bentuk butir soal mempunyai kebaikan dan kelemahan. Oleh karena itu diperlukan pertimbangan-pertimbangan dalam memilih bentuk butir soal yang akan dipakai dalam setiap perangkat tes yang akan disusun. Dalam hubungan ini tujuan/criteria penilaian, materi yang akan dinilai, status peserta yang akan diuji dan waktu yang tersedia harus benar-benar dikaji secara benar dan tepat.
2. Dalam perumusan butir soal perlu memperhatikan hal yang berhubungan dengan kaidah penyusunan kalimat, pemakaian kata, dan kemampuan bahasa yang akan diuji. Rumusan bentuk soal hendaknya singkat, jelas dan mudah dimengerti.
3. Khusus pengembangan achievement tes erat hubungannya dengan tujuan umum pengajaran (TUP) dan tujuan khusus pengajaran. Dalam penyusunan butir soal tes harus bertitik tolak dari tujuan tersebut.

BAB IV

ADMINISTRASI TES

A. Penyiapan Perangkat Tes

Beberapa langkah penyiapan perangkat tes adalah sbb:

1. Menetapkan tujuan

Menetapkan tujuan tes berhubungan dengan maksud penggunaan tes. Tujuan penggunaan tes akan memberi corak terhadap penyusunan butir soal.

Mis : Tes yang dimaksudkan untuk EBTA, itemnya harus disusun secara menyeluruh mengungkap materi yang diajarkan mulai dari yang mudah sampai kepada yang sukar serta berorientasi pada kemampuan intelektual tertentu.

2. Analisa Materi

Pemahaman yang mendalam akan cakupan isi dan luas materi yang akan diukur sangat perlu diperhitungkan. Untuk tes prestasi pendalaman akan kurikulum, buku-buku pelajaran digunakan sebagai petunjuk untuk merumuskan tes dengan baik yang akan dituangkan dalam bentuk kisi-kisi. Studi literature sebagai sumber untuk menentukan defenisi operasional dari variable-variabel psikologis yang akan diukur sesuai dengan tujuan yang dimaksud sangat mendukung.

3. Menentukan Kisi-kisi

Kisi-kisi atau blue print atau table of specification ,lay out, plan, atau frame work adalah suatu daftar berbentuk matriks yang memuat komponen-komponen sebagai berikut :

a. Pokok bahasan

Pokok bahasan dan subpokok bahasan merupakan ruang lingkup butir soal yang disusun. Pokok bahasan dan sub pokok bahasan perlu ditetapkan pembobotannya yang ditetapkan dengan kategori : penting, sedang dan kurang penting.

b. Aspek perilaku

Aspek perilaku berupa keadaan intelektual, afektional, psikomotor yang dimiliki subyek uji baik sebagai hasil belajar maupun berupa potensi yang diperinci kedalam kategori atau jenjang kemampuan menurut taxonomi Bloom yang akan diukur dengan tes yang dibuat.

Mis : Intelektual mencakup :

1. Ingatan
2. Pemahaman
3. Penerapan
4. Analisis
5. Sintesis
6. Evaluasi

5. Bentuk soal

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

6. 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.

7. 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 itu baru ditentukan jumlah butir soalnya. Namun demikian perlu mempertimbangkan daya konsentrasi subyek uji, serta perimbangan waktu yang diperlukan untuk pengerjaan soal yang mudah dan yang sukar.

B. Format-format Perencanaan

1. Format spesifikasi butir soal

Format ini menggambarkan :

- a. alokasi butir soal sehubungan dengan tujuan tes (aspek yang akan dites)
- b. alokasi butir soal sehubungan dengan isi (materi yang akan di tes)

Tabel 4.1. Spesifikasi Butir Soal
Bidang Studi : Psikologi Umum

Alokasi butir soal berdasarkan Tujuan	1. Pengetahuan30 % 2. Pengertian50 % 3. Penerapan20 %
Alokasi butir soal berdasarkan Materi/isi	1. Aliran Psikologis50 % 2. Fungsi Psikologis50 %

Penentuan besarnya persentase alokasi soal berdasarkan tujuan dan materi isi disesuaikan dengan aspek perilaku yang diukur.

Dari format spesifikasi butir soal/item dilanjutkan dengan perencanaan alokasi butir soal ditinjau dari isi/pokok bahasan dan tujuan tes berdasarkan aspek perilaku yang diukur. Dengan kata lain format perencanaan yang akan dibuat menggambarkan persentase alokasi butir soal pada setiap tujuan tes dan setiap isi pokok bahasan.

Tabel 4.2 : Perencanaan Tes Hasil Belajar Psikologi Umum

Materi/Sub pokok bahasan Aspek-aspek	Tujuan Tes		
	Pengetahuan	Pemahaman	Aplikasi
Aliran Psikologi	15 %	25 %	10 %
Fungsi Psikogis	15 %	25 %	10 %

Keterangan : Penentuan dari perencanaan persentase alokasi item ini didasarkan dari spesifikasi item .

Setelah perencanaan diatas selesai dilanjutkan dengan lay out tes, yang menggambarkan jumlah butir soal untuk setiap bentuk butir soal dari setiap persentase yang terdapat dalam perencanaan tes yang telah dibuat.

Tabel 4.3 : Lay Out Tes Psikologi Umum

Materi/Sub pokok bahasan Aspek-aspek	Tujuan Tes		
	Pengetahuan	Pemahaman	Aplikasi
Aliran Psikologis	B-S = 3 P-G = 4	B-S = 3 P-G = 4	B-S= 3 P-G=4
Fungsi-fungsi Psikologis	B-S = 3 P-G = 3	B-S = 2 P-G = 3 Es = 2	B-S= 2 P-G= 2 Es = 2

Keterangan :

Tes B-S = 16 item

Tes P-G = 20 item

Tes Esei = 4 item

Total = 40 item

Selanjutnya akan diteruskan dengan pembuatan format kisi-kisi tes.

2. Format kisi-kisi

Format kisi-kisi adalah format yang dibuat terakhir dari seluruh langkah kegiatan perencanaan tes/butir soal. Selain penentuan yang lebih terperinci dari lay out, pada kisi-kisi tes diperkirakan tingkat kesulitan (DP) butir soal yang akan ditulis. Kisi-kisi tes yang baik mempunyai ciri-ciri sebagai berikut :

- a. Pada setiap item terlihat tujuan tes
- b. Pada setiap item terlihat isi pokok bahasan yang dites
- c. Pada setiap item terlihat tingkat kesukarannya (DP)
- d. Pada setiap item jelas bentuk tesnya

Tabel 4.4 : Kisi-kisi Tes Psikologi Umum

Materi/Sub pokok Bahasan/Aspek	DP	Tujuan Tes					
		Pengetahuan		Pemahaman		Aplikasi	
		B-S	P-G	B-S	P-G	P-G	Es
Aliran Psikogi	MD						
	SK						
	SK						
Fungsi Psikologis	MD						
	SD						
	SK						

MD = Mudah SD = Sedang SK = Sukar

Menjadi catatan bahwa pengkategorian DP diatas ini menjadi 3 tingkatan hanya semata-mata untuk perencanaan penulisan item, tetapi belum menjadi indeks kesukaran (p) ataupun indeks daya beda (d) yang dihasilkan dari perhitungan yang benar. Perhitungan analisis butir untuk menentukan daya beda dapat dilakukan setelah didapatkan data hasil uji coba.

Berpedoman pada kisi-kisi tes tersebut ditulislah butir soal yang bersangkutan. Untuk penulisannya dapat dipedomani prinsip-prinsip sesuai dengan bentuk tes dimaksud.

Tabel 4.5 : Format Penulisan Soal
Mata Kuliah : Psikologi Umum

Pokok Bahasan/ Aspek	TIK	Aspek Perilaku	Bentuk Soal	Tkt Kesukaran	Rumusan soal No :
					1. 2. 3. 4. 5.

C. Petunjuk Tes

Petunjuk tes dimaksudkan untuk membimbing peserta tes bagaimana mereka melaksanakan/mengerjakan/menjawab tes (pedoman pengerjaan tes). Petunjuk ini terdiri dari 2 macam, yaitu :

Petunjuk Umum dan Petunjuk Khusus

Petunjuk umum, berisi :

- a. Pengisian identitas peserta
- b. Tempat memberi jawaban
- c. Penjelasan system penilaian
- d. Waktu pengerjaan tes

Petunjuk Khusus, berisi :

- a. Tanda yang dipergunakan dalam memberi jawaban
- b. Jawaban pilihan yang diminta

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.

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.

Penjelasan tentang system 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 biasanya pada petunjuk disebutkan apakah ada keterkaitan untuk menjawab/mengerjakan soal secara berurutan atau boleh memilih jawaban yang paling diketahui peserta.

Waktu pelaksanaan tes, menghusus pada jumlah waktu yang disediakan biasanya dituliskan dalam jumlah menit.

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

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 biasanya 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 an 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 diatas, sesuai dengan kebutuhan, petunjuk tes objektif sering ditambahkan dengan :

- jangan mencoret perangkat soal.

- 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 berapa 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 misalnya :
 - menimbulkan rasa bosan/jemu
 - Kelelahan yang tinggi sehingga peserta condong "asal" memberi jawaban
 - Memberi kesempatan untuk saling kerjasama.
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 seajar kelas dengan peserta yang akan diuji.
4. Sering terjadi seorang penguji menambah atau mengurangi waktu pelaksanaan tes dari yang direncanakan sebelumnya. Hal ini sangat tidak baik apabila peserta telah mengetahui sebelum pelaksanaan tes dimulai. Walaupun misalnya sebahagian besar peserta telah selesai namun waktu masih ada, tidak baik apabila penguji mengakhiri ujian sebelum waktu yang ditentukan, semasih ada peserta yang masih bekerja, dan menambahkan waktu ujian apabila peserta belum ada yang selesai walaupun waktu telah berakhir. Adalah lebih baik mempertimbangkan keadaan tersebut dalam membuat skor.

E. 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 disekolah sekolah (pada pelaksanaan tes sumatif), pegawai sering diturut sertakan sebagai pengawas. Perlu diingat bahwa ketentuan penilaian turut ditentukan oleh baik tidaknya penyajian tes, terutama dalam tes objektif.

Beberapa hal yang berhubungan dengan perangkat soal menyangkut :

1. Penulisan butir soal

Bagi tes objektif terutama tes pilihan ganda, sebaiknya alternatif jawaban tersusun dibawah walaupun mungkin bisa dibuat dua atau lebih alternatif jawaban dalam satu baris.

2. Penyajian perangkat soal sebaiknya dilakukan dengan memperbanyak perangkat soal sesuai dengan jumlah peserta tes. Hal ini terutama bagi tes objektif.
3. Kondisi tulisan perangkat soal harus jelas dan terang. Untuk itu karena kemungkinan adanya kesalahan pengetikan dan kurang baiknya cara memperbanyak perangkat soal, sebaiknya penguji memeriksa lebih dahulu perangkat soal yang diperbanyak untuk memperbaiki/membuat ralat.
4. Pada tes esai walaupun sedikit soalnya adalah baik bila diperbanyak dibandingkan dengan apabila soal didiktekan.

Berikut ini diuraikan beberapa factor yang berhubungan dengan pelaksanaan tes.

1. Faktor fisik meliputi hal-hal yang berhubungan dengan fisik peserta dan pengawas, sedang factor 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/sebaik mungkin. Pada pagi hari adalah waktu yang cocok dalam melaksanakan ujian mengingat peserta masih diharapkan merasa "senang/tenang".
Bila dalam satu hari itu ada beberapa mata ujian yang lain diberi waktuyang "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 ia memulai/memasuki ruangan ujian. Untuk itu hendaknya sebelum peserta memulai ujian, pengawas terlebih dahulu menciptakan suasana tenang pada diri peserta dengan memberikan bimbingan secara kelompok.
"Suasana bersahabat" perlu diciptakan oleh pengawas ujian apalagi pengawas adalah orang yang asing bagi peserta.
 - c. Pengawas hendaknya tidak menunjukkan sikap-sikap yang kurang baik selama dan sebelum ujian berlangsung. Seperti yang disebut diatas (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 pengawas ujian.
 - d. Rasa bersaing diantara peserta akan mengurangi keinginan untuk bekerjasama. Hal ini perlu ditekankan pengawas sebelum memulai ujian, dengan kata lain dalam bimbingannya sebelum ujian dimulai hendaknya ditimbulkan 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. Ujian yang "tiba-tiba" di sekolah sering juga dilaksanakan, sebaiknya pelaksanaan tes yang seperti itu ditujukan bukan dalam rangka pembuatan nilai murid, tetapi lebih

ditekankan pada tujuan agar murid selalu mempersiapkan diri dari rumahnya akan pelajaran yang akan diterimanya disekolah, sehingga hasil yang dicapainya lebih baik.

- f. Bangku dan meja dimana peserta akan duduk mengerjakan ujian hendaknya memadai dilihat dari kemungkinan untuk menulis dan keamanan dalam duduk. Dalam hubungan dengan meja tulis yang dipergunakan terutama dalam tes esai harus sebaik mungkin.
- g. Kondisi ruangan ujian hendaknya "sehat". Pertukaran udara ruangan hendaknya memenuhi syarat kesehatan.

Luasnya ruangan ujian harus disesuaikan dengan jumlah peserta ujian yang memungkinkan peserta sukar untuk bekerjasama (saling contoh-mencontoh) hal ini terutama dalam tes objektif.

Pemakaian ruangan yang sangat luas misalnya memungkinkan memuat peserta 80 orang lebih, malah sampai ratusan orang kurang baik. Apalagi penggunaan ruang terbuka, merupakan kondisi ruangan yang kurang baik.

Besarnya ruangan sebaiknya cocok untuk menampung peserta paling banyak 20 orang dengan dua orang pengawas. Bila melebihi kondisi tersebut, ketenangan dan objektivitas ujian sukar dicapai, walaupun untuk peserta yang banyak diikuti oleh penambahan pengawas. Kesulitan yang paling sering timbul adalah ketepatan waktu tes, pengumpulan lembar jawaban dan "pemerataan" pengawasan.

- h. Benda-benda/bahan-bahan yang dapat membantu peserta ujian untuk menjawab pertanyaan harus dihindarkan dari ruangan ujian. Biasanya bila ujian dilaksanakan diruang kelas, benda seperti peta dan lain yang sering dipajang diruangan harus disingkirkan.
- i. Suasana luar ruangan hendaknya turut dipertimbangkan, misalnya kebisingan suara yang sangat mengganggu kedalam ruangan ujian.
- j. Selain factor luar, factor perangkat tes yang disajikan dapat tergantung ketenangan ujian. Tulisan yang kurang jelas dan kalimat yang meragukan peserta akan peserta banyak untuk meminta penjelasan pada waktu tes sedang berjalan, dan hal ini tentunya akan mengganggu peserta yang sedang asik menjawab/memikirkan jawaban soal. Untuk itu sebelum perangkat tes dibagikan hendaknya terlebih dahulu dikoreksi apakah ada tulisan yang kurang jelas atau pertanyaan yang sangat meragukan, perbaikan perangkat tes dilakukan sebelum ujian dimulai.

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" misalnya tes seleksi penerimaan, tes sumatif akhir semester.

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

- a. Kehadiran peserta tes

Kehadiran peserta tes biasanya diharapkan hadir di lokasi pelaksanaan kira-kira 10 menit sebelum tes dimulai. Hal ini dimaksudkan untuk mengurangi

keterlambatan peserta tes, dan memberikan ketenangan pada peserta sebelum memasuki ruangan serta waktu yang dipergunakan menyampaikan beberapa pengumuman dan lain sebagainya.

b. Penyediaan bahan-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 menghindari penipuan peserta.

Tanda peserta sebaiknya memulai beberapa item identitas, misalnya nama, no. ujian dan mungkin tanggal lahir serta alamat. Pemuatan beberapa item identitas ini menghindari nama yang serupa. Disamping itu bagi tes penerimaan atau 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 kedalam ruangan ujian.
- Larangan saling pinjam-meminjam alat selama ujian berlangsung.
- Larangan berkomunikasi antara peserta selama ujian berlangsung.

e. Tanda-tanda Waktu Pelaksanaan Tes

Peserta 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 sesuatu secara serentak, untuk hal ini dipergunakan sesuatu bunyi, misalnya bunyi lonceng.

Pada umumnya tanda waktu itu meliputi :

- Aba-aba memasuki ruangan
- Aba-aba memulai ujian
- Aba-aba kesiapan mengakhiri ujian (biasanya lebih kurang 5 menit) sebelum ujian diakhiri.
- Aba-aba akhir waktu pelaksanaan tes.

f. Sanksi-sanksi

Peserta hendaknya mengetahui sanksi-sanksi atas pelanggaran tata tertib tes, misalnya :

- Sanksi atas keterlambatan hadir
- Sanksi atas tidak membawa tanda peserta
- Sanksi bila bekerja sama dan lain-lain.

g. Hal-hal yang Berhubungan dengan Pengawasan

Selain tata tertib bagi murid diperlukan ketentuan-ketentuan bagi pengawas, misalnya :

- Tanda pengenal pengawas, hal ini menghindari pengawas yang bukan orangnya, atau menghindari turut campurnya orang lain dalam pelaksanaan ujian.
- Kehadiran pengawas pada lokasi ujian
Sebaiknya pengawas sudah hadir di lokasi ujian lebih kurang 20 menit sebelum ujian dimulai yang dimaksudkan untuk perencanaan penggantian pengawas yang berhalangan hadir dan jugsan untuk dipergunakan dalam

menyampaikan hal-hal tertentu sesuai dengan ketentuan pengawas dan juga dalam membagi/menerima bahan-bahan ujian.

-Ketentuan ruang gerak pengawas, misalnya pengawas tidak boleh memasuki ruang ujian lain, pengawas tidak diperkenankan terlalu banyak bergerak didalam ruang ujian selama ujian berlangsung.

h. Beberapa ketentuan atau Petunjuk Tambahan :

Ketentuan/petunjuk tambahan maksudnya adalah ketentuan/petunjuk pelaksanaan tes yang dilaksanakan untuk memperjelas petunjuk yang disampaikan secara tertulis, diantaranya adalah :

-Walaupun pada perangkat tes telah dicantumkan cara bekerja hendaknya petunjuk ini disampaikan juga secara lisan sebelum ujian dimulai. Semakin baik apabila berupa contoh.

-Pengalaman menunjukkan bahwa dalam menuliskan identitas peserta sering lupa walaupun tempat atau petunjuk untuk itu telah disediakan pada lembar jawaban. Untuk itu adalah sangat tepat bila sebelum tes dimulai murid/peserta tes diajak serentak mengisi identitas yang diperlukan, setelah itu barulah tes dimulai dalam pekerjaan yang sebenarnya.

-Pemberitahuan waktu yang tinggal (sekitar 5 menit lagi menjelang selesai) harus dilakukan secara hati-hati apalagi dengan kata-kata "ancaman" seperti awas waktu tinggal 5 menit lagi. Selain dapat mengganggu konsentrasi anak juga berakibat kepada ketenangan ujian. Untuk ini sebaiknya sebelum ujian dimulai pengawas memberitahukan jumlah waktu yang disediakan dan jumlah item perangkat ujian. Dengan cara ini peserta tes akan lebih hati-hati dalam pemakaian waktu.

-dalam membagi perangkat soal dan mengumpulkan pekerjaan peserta hendaknya dilakukan sendiri oleh pengawas. Kurang baik kalau dilakukan oleh peserta tes.

-untuk menjaga ketenangan ujian pengawas hendaknya tidak melayani peserta yang meminta penjelasan secara kelompok, untuk itu hendaknya dalam petunjuk permulaan (sebelum ujian dimulai) pengawas memberitahukan cara penyampaian masalah yang dihadapi peserta sewaktu ujian berlangsung, misalnya dengan mengangkat tangan, dan untuk hal yang demikian pengawas mendekati peserta serta menyampaikan jawaban secara individual.

-teguran dari pengawas sewaktu ujian berlangsung kepada seseorang peserta yang melanggar ketentuan hendaknya jangan sampai mengganggu peserta lainnya. Untuk itu sebelum ujian dimulai hendaknya pengawas memberitahukan peserta yang melanggar ketentuan tidak langsung ditegur tetapi dicatat dalam formulir pelaksanaan ujian.

3. Pemeriksaan jawaban tes.

Sesuai dengan bentuk tes, pemeriksaan jawaban secara umum terbagi atas/bentuknya antara 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 system "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. Setiap tingkat kesukaran soal diberikan bobot nilai misalnya :

Tabel 4.6 Tingkat Kesukaran dan Bobot Nilai Butir Soal

Tingkat kesukaran butir soal	Bobot Nilai	Bobot Nilai
Sukar	3	6
Sedang	2	5
Mudah	1	4

Pengutaraan dua macam bobot nilai diatas maksudnya dapat dipilih satu diantaranya atau bagi guru yang berpengalaman dapat membuat bobot nilai yang bervariasi lain, misalnya karena melihat luasnya bahan yang diuji dengan butir soal yang dirasa sukar. Namun, perlu banyak diingat bahwa penentuan bobot nilai jangan terlalu banyak perbedaan antara butir soal sukar dibandingkan dengan butir soal sedang maupun dengan butir soal mudah. Untuk ini hendaknya dalam menulis/mengkonstruksi tes esai perbandingan tingkat kesukaran soal antara sukar, sedang dan mudah tidak terlalu berbeda banyak.

- c. 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 jumlahbutir soal (5).
- d. Rentangan yang dipakai dalam menskor tes secara keseluruhan atau menskor setiap soal hendaknya lah 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. Biasanya rentangan skor yang dipergunakan untuk setiap butir soal adalah 10 angka untuk keseluruhan tes adalah 100.

- e. Rentangan tingkat kebenaran jawaban dibuat sesederhana mungkin sesuai dengan kesederhanaan rentangan skor. Yang umum digunakan adalah 2 atau 4 rentangan tingkat kebenaran jawaban, misalnya :
 X sangat benar, benar, kurang benar, dan salah atau
 X jawaban tepat, kurang tepat, dan salah.
 Dalam hubungannya dengan rentangan skor dapat dibuat sebagai berikut

Tabel 4.7 Rentangan Kebenaran Jawaban dan Skor

Rentangan kebenaran jawaban	Rentangan Skor	
	Jenis 1	Jenis 2
Jawaban tepat	8 – 10	9 - 10
Jawaban kurang tepat	4 – 7	3 - 8
Jawaban tidak tepat	1 – 3	1 - 2

Pemberian rentangan skor untuk setiap rentangan kebenaran jawaban dimaksudkan adanya variasi ketetapan jawaban. Sebaiknya rentangan kebenaran jawaban yang berada pada tingkat menengah diberi rentangan skor yang lebih banyak. Pemberian skor bagi jawaban yang tidak tepat dimaksudkan ialah menghindarkan skor yang ekstrim rendah dan sebagai penghargaan atas usaha karena kemungkinan kesalahan jawaban akibat kesalahan pengertian akan makna kalimat butir soal.

- f. Sebelum menskor sesuatu kertas jawaban periksa lebih dahulu buku sumber dan rencana pengajaran untuk melihat jawaban yang benar dan mungkin akan dijumpai. Hal ini dimaksudkan untuk melengkapi kunci jawaban yang telah dibuat sebelumnya.
- g. Perlu ditekankan bahwa factor-faktor seperti tulisan, kebersihan pekerjaan atau tinggi rendahnya skor kecuali bagi mata pelajaran yang tertentu, misalnya bagi mata pelajaran menggambar.
- h. Setiap jawaban harus dibaca secara keseluruhan untuk mendapatkan inti jawaban dan kualitas jawaban, barulah kemudian mengklasifikasikannya pada rentangan kebenaran jawaban dan seterusnya memberi skor.
- i. 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 menghindarkan 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 objektivitas 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 jumlah pemeriksa.

- j. 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).
- k. Waktu yang digunakan 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 (lelah, kurang sehat, kalut).

Salah satu contoh penggunaan bobot nilai :

Guru A menguji muridnya 5 soal esai tes, dengan tingkatan kesukaran butir soal yang dibuat adalah :

- Butir soal nomor 1 mudah
- Butir soal nomor 2 mudah
- Butir soal nomor 3 sedang
- Butir soal nomor 4 sedang
- Butir soal nomor 5 sukar

Setelah pekerjaan murid diperiksa, untuk murid nama Misro adalah sebagai berikut :

- Butir soal nomor 1 skor 10
- Butir soal nomor 2 skor 8
- Butir soal nomor 3 skor 8
- Butir soal nomor 4 skor 7
- Butir soal nomor 5 skor 5

Tingkat skor berdasarkan bobot adalah sebagai berikut :

- Tingkat kesukaran mudah, bobot skor 3
- Tingkat kesukaran sedang, bobot skor 4
- Tingkat kesukaran sukar, bobot skor 5

Untuk mendapatkan skor Misro dalam tes tersebut dihitung melalui table berikut :

Tabel 4.8. Proses Perhitungan Skor Akhir

N0. Soal	Tkt Kesukaran	Skor	Bobot Skor	Skor x Bobot	Skor Tes
1.	Mudah	10	3	30	139 ----- = 7,3 19
2.	Mudah	8	3	24	
3.	Sedang	8	4	32	
4.	Sedang	7	4	28	
5.	Sukar	5	5	25	
Jumlah		38	19	139	

(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 yang benar. 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 butir soal 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 yang khusus. Disamping itu sedikit berlawanan/kesalahan dalam mengikuti petunjuk akan mengakibatkan kerugian besar pada peserta tes.
- c. Menerka jawaban adalah salah satu kelemahan yang terdapat dalam tes objektif. Untuk itu dalam mencari skor setiap peserta dalam tes objektif dipergunakan rumus yang bertujuan mengurangi pengaruh terkaan terhadap skor tes. (Sebaiknya hal ini diberitahukan pada petunjuk umum tes, sehingga menjauhkan peserta dari penerkaan jawaban pertanyaan).

Rumus yang dimaksud adalah :

$$Sc = R - \frac{W}{Op - 1}$$

Dimana :

- Sc = Skor (skor Akhir)
- R = Right (jumlah jawaban betul)
- 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 "system Weight" sesuai dengan tingkat kesukarannya bentuk soal dan banyak pertanyaan untuk setiap bentuk soal.

Rangkuman

1. Dalam merencanakan perangkat tes hasil belajar yang pertama harus dilakukan adalah menentukan tujuan tes yang seterusnya dilakukan analisis kurikulum, analisis buku pelajaran serta meninjau kembali program pengajaran (Program tahunan, semesteran/catur wulan dan satuan -- satuan pelajaran yang disusun).
2. Sebelum menyusun butir soal harus dilakukan perencanaan-perencanaan butir soal, kisi-kisi butir soal 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 fasilitas 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 dalam pengawasan penyajian tes harus ditapkan dan diketahui/dipahami oleh peserta ujian dan pengawas ujian (terutama hal ini dimaksudkan pada ujian-ujian yang dilaksanakan secara bersama dan serentak seperti ujian EBTA 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 V
TEKNIK PENGOLAHAN DAN PENGGUNAAN
HASIL PENGUKURAN

A. Penyajian Skor

Setelah diperoleh skor-skor dari suatu hasil pengukuran, langkah seterusnya adalah melakukan pengolahan terhadap hasil-hasil tersebut. Untuk memudahkan pembacaan skor-skor tes, perlu dilakukan pengorganisasian data skor melalui daftar distribusi frekuensi dan pembuatan grafik.

1. Daftar distribusi frekuensi

Daftar distribusi frekuensi terdiri dari 2 macam yaitu daftar distribusi tunggal dan daftar distribusi bergolong/kelompok. Untuk lebih jelasnya dapat dilihat dalam table berikut :

Tabel 5.1. Skor Mentah Responden Dalam Urutan Acak

25	33	35	37	55	27	40	33	39	28
34	29	44	36	22	51	29	21	28	29
33	42	15	36	41	20	25	38	47	32
15	27	27	23	46	10	16	34	18	14
46	21	19	26	19	17	24	21	27	16

Penyajian data skor table 4.1 sukar dibaca karena tidak ada keterpolaannya. Berapakah skor tertinggi atau skor terendah tidak, skor berapakah yang paling banyak frekuensinya, apakah penyebaran skor normal atau merata tidak mudah terbaca dan terjawab dengan segera, padahal keadaan seperti itu erat hubungannya dengan penafsiran terhadap hasil yang diperoleh.

Untuk memberi gambaran yang lebih terorganisir serta menghindarkan pemborosan waktu alat-alat tulis, maka penyajian data skor perlu dikelompokkan, hal ini terutama apabila dat skor sangat banyak. Pengelompokan yang dimaksud dalam table berikut ini .

Tabel 5.2 Tabel Frekuensi Tunggal Skor Mentah Responden

Skor	Turus	Frekuensi
55	I	1
51	I	1
47	I	1
46	II	2
44	I	1
42	II	2
41	I	1
40	I	1
39	I	1
38	I	1
37	I	1
36	II	2
35	I	1
34	II	2
33	III	4
32	I	1
29	III	3
28	II	2
27	III	4
26	I	1
25	II	2
24	I	1
22	I	1
21	III	3
20	I	1
19	II	2
18	I	1
17	I	1
16	II	2
15	II	2
14	I	1
10	I	1

Tabel 5.3 Tabel frekuensi Bergolong Skor Mentah Responden

Internal Skor	Turus	Frekuensi
55 – 59	I	1
50 – 54	I	1
45 – 49	III	3
40 – 44	IIII	4
35 – 39	IIII I	6
30 – 34	IIII II	7
25 – 29	IIII IIII II	12
20 – 24	IIII I	6
15 – 14	IIII III	8
10 – 14	II	2

Penyajian skor pada table 4.3 lebih singkat dibanding dengan table 4.2. Pada kedua teknik penyajian ini sudah tergambar penyebaran skor apakah distribusi normal atau tidak. Teknik penyajian dengan pengelompokan pada table distribusi bergolong ini sering dipakai dalam rangka pengolahan skor-skor mentah menjadi yang lebih berarti misalnya menentukan skor mentah menjadi nilai bentuk huruf atau dalam menafsirkan hal-hal yang berhubungan dengan keadaan tes atau kelompok siswa.

Sehubungan dengan pengelompokan distribusi skor ada hal-hal yang perlu diketahui yaitu :

- a. Hal skor nyata pada suatu klas interval dari satu klas interval dengan klas interval yang di bawah atau di atasnya berbeda satu point (angka), padahal mungkin saja skor yang diperoleh tidak angka bulat. Untuk itu pedoman yang dipergunakan adalah sebagai berikut : Angka yang tertinggi dari suatu klas interval adalah setengah angka diatas skor tertinggi dari klas interval dimaksud (upper real class limits) sedang angka yang terendah dari suatu klas interval adalah setengah angka dibawah skor terendah dari klas interval yang dimaksud (lower real class limits) dapat dilihat dari contoh table berikut :

Tabel 5.4 Batas Nyata Dari Suatu Klas Interval

Kelas Interval	Batas Nyata
50 – 54	49,5 – 54,5
10 – 19	9,5 – 19,5
2,5 – 2,9	2,45 – 2,95

Batas nyata dari suatu klas interval dapat dipergunakan untuk menentukan besar unit interval dari suatu klas interval yaitu : besar unit

interval = URCL - LRCL (Batas Nyata Atas dikurang Batas Nyata Bawah).

Contoh pada table di atas dapat dihitung besar unit interval dari kelas interval 50 - 54 adalah $54,5 - 49,5 = 5$ unit.

- b. Hal besarnya unit interval dan banyaknya kelas interval . Besarnya unit interval tergantung pada jauhnya jarak skor tertinggi dengan skor terendah. Banyaknya kelas interval juga harus seimbang dengan banyaknya skor mentah (peserta suatu ujian). Banyaknya kelas interval, yang umum berkisar sepuluh kelas interval dan bagi data yang cukup besar frekuensinya dan cukup jauh jarak skor tertinggi dan terendah dapat dibuat sampai dengan dua puluh kelas interval. Besarnya unit interval sebaiknya adalah ganjil agar mudah menentukan angka tengah dan kelas interval yang dimaksud, hal ini dimaksudkan memudahkan untuk melakukan perhitungan statistik selanjutnya.
 - c. Hal Turus (Tally Marks)
Tally Mark atau turus sebenarnya tidak mempunyai fungsi langsung pada perhitungan (penafsiran) selanjutnya. Tally Marks berfungsi dalam memudahkan dan meningkatkan ketelitian kerja dalam perhitungan frekuensi dari setiap skor mentah terutama dalam distribusi yang meliputi cukup banyak peserta/skor mentah.
2. Grafik dari suatu distribusi frekuensi untuk menggambarkan suatu distribusi frekuensi dapat digunakan teknik diagram batang (Histogram) dan teknik diagram garis (Poligon). Kedua teknik ini memerlukan table yang terdiri dari garis-garis vertical dan garis horizontal. Garis vertical untuk penempatan frekuensi dan garis horizontal untuk penempatan skor atau kelompok skor.

B. Ukuran Tendensi Sentral

1. Mode (Modus)

Mode adalah suatu skor yang "sering timbul". Hal ini berarti mengunjuk kepada skor yang paling tinggi frekuensinya. Mode kurang dapat menunjukkan mutu dari suatu kelas, terutama pada kelompok/kelas yang tidak normal (tidak sebanding antara jumlah responden yang tergolong mampu, sedang dan rendah). Dan sering pula kita jumpai bahwa didalam suatu distribusi ada dua atau lebih skor yang mempunyai frekuensi sama. Dalam keadaan demikian analisa mengenai mutu kelompok tidak dapat dilakukan karena kalau kita katakan misalnya 34 menjadi mode, berarti kelas tersebut mempunyai mutu yang baik (mengingat jumlah soal hanya 40), tetapi kalau kita sebut mode adalah 19 maka dapat kita sebut mutu kelas itu rendah. Dengan satu teknik menghasilkan dua gambaran yang berbeda menunjukkan bahwa teknik ini tidak valid. Teknik mode dapat kita pergunakan apabila distribusi skor menunjukkan kenormalan.

2. Median

Median disebut juga dengan angka tengah (skor yang terletak ditengah). Median ini ialah suatu skor/nilai atau titik yang membagi sederetan skor yang tersusun berdasarkan besar kecilnya atas dua bagian.

Kadang pendistribusian skor dilakukan dalam data kelompok, karena mengingat banyaknya pengikut tes. Untuk itu lokasi/letak Median ditentukan melalui rumus :

$$\text{Median} = \frac{N + 1}{2}$$

Cara menghitung Median dengan bentuk data distribusi frekuensi bergolong/kelompok yang menggunakan kelas interval Yaitu :

$$\text{Median} = Bb + \frac{\frac{1}{2} N - Cfb}{Fd} \times i$$

Bb = Batas Nyata bawah

N = Jumlah responden

Cfb = Jumlah frekuensi kumulatif di bawah klas interval dimana terletak nilai Median

Fd = Frekuensi yang terdapat dalam kelas interval dimana terletak nilai Median

i = Jumlah unit interval

3. Mean

Mean merupakan rata-rata klas dimana setiap skor itu turut menentukan besarnya mean. Oleh karena itu dibandingkan dengan Mode dan Median, penentuan mutu klas berdasarkan Mean adalah yang paling baik. Oleh karena rata-rata hitung tidak hanya berdasarkan keadaan-keadaan setempat/skor-skor setempat, melainkan berdasarkan pada skor kelompok secara keseluruhan.

Ada 3 macam rumus menentukan nilai Mean, yaitu :

Rumus 1.

$$\text{Mean} = \frac{\sum X}{N}$$

X = Skor

N = Jumlah responden

Rumus 2.

$$\text{Mean} = \frac{\sum fX}{N}$$

X = Skor

N = Jumlah subyek

Rumus 3. Disebut rumus Mean Terkaan

$$\text{Mean} = \text{MT} + \frac{\sum fx}{N} \times i$$

MT = Mean terkaan (tanpa perhitungan)

f = Frekuensi (banyak subyek yang memiliki nilai tertentu)

x = deviasi nilai individu dari mean

N = Jumlah keseluruhan subyek

i = interval

Rumus mean terkaan digunakan pada data yang tersusun dalam bentuk distribusi frekuensi bergolong.

Hasil perhitungan mean berada kurang lebih sedikit pada pertengahan skor maximum atau pertengahan nilai tertinggi. Dengan keadaan ini maka kita sebut mutu hasil pengukuran sedang. Jika mean berada disekitar daerah skor maksimum, maka kita sebut hasil pengukuran bermutu baik/tinggi, sedang apabila mean berada disekitar daerah skor minimum/nilai terendah, maka kita simpulkan mutu hasil pengukuran kurang.

C. Ukuran Variabilitas

1. Range

Range dapat diartikan adalah jarak antara skor tertinggi dengan skor terendah dari suatu deretan skor. Melihat keterangan ini, maka range dapat dipergunakan dalam menganalisa luasnya penyebaran suatu distribusi hasil pengukuran. Luas tidaknya suatu penyebaran distribusi hasil pengukuran menunjukkan heterogen atau homogenya kelompok subyek.

Dalam penilaian hasil belajar informasi ini diperlukan untuk memperbaiki proses pengajaran yang lebih baik. Kalau pada mulanya metode, cara dan bahan dalam proses pengajaran disamakan, karena mengingat kelompok/kelas itu merupakan kelas yang sejajar (misal :sama-sama kelas 1), tetapi dengan penggambaran heterogenitas kemampuan kita harus membedakan metode, cara dan juga mungkin bahan untuk masing-masing kelompok.

2. Semi Antar Kuartil

Sering dan mungkin dalam suatu kelompok terdapat skor/hasil individu yang ekstrim baik ekstrim yang rendah maupun ekstrim tinggi. Skor yang ekstrim ini akan mempengaruhi hasil perhitungan range dan mean. Akibatnya kesimpulan kita yang berbentuk secara keseluruhan itu akan menyimpang, walaupun proses perhitungan Mean dan Range sudah benar.

Agar kesimpulan-kesimpulan tidak menyimpang, sebaiknya skor yang ekstrim itu dikesampingkan (tidak diikutsertakan) dalam perhitungan. Agar terdapat ketepatan dalam menentukan skor gunakanlah teknik Semi antar kuartil.

3. Standar Deviasi

Teknik analisa Standar Deviasi adalah lebih baik dari Range dan Semi antar kuartil, karena teknik standar deviasi berdasarkan pada skor-skor keseluruhan, dan perhitungan-perhitungan penyimpangan didasarkan pada salah satu ukuran tendensi sentral yaitu Mean dan Median.

Standar deviasi dapat diartikan merupakan suatu jarak yang terletak diatas dan dibawah Mean. Dalam menentukan keadaan dari pada kelompok melalui standar deviasi dari Mean ini berhubungan erat dengan kurve normal. Artinya kita harus lebih dahulu mengetahui pendistribusian suatu skor yang menunjukkan keadaan (Normal Curve). Dalam normal curve penentuan persentase itu umumnya dibagi 6 yaitu :

- a. Daerah $-3,0$ SD s/d $-2,0$ SD
- b. Daerah $-2,0$ SD s/d $-1,0$ SD
- c. Daerah $-1,0$ SD s/d $0,0$ SD
- d. Daerah $0,0$ SD s/d $+1,0$ SD
- e. Daerah $+1,0$ SD s/d $+2,0$ SD
- f. Daerah $+2,0$ SD s/d $+3,0$ SD

Dalam hubungan dengan analisa keadaan kelompok/kelas apakah homogen atau heterogen, pembagian ini disederhanakan menjadi 3 bagian, yaitu :

- a. Daerah sekitar Mean ($-1,0$ SD s/d $+1,0$ SD frekuensinya $68,26$ %)
- b. Daerah dibawah $-1,0$ SD frekuensinya $15,87$ %
- c. Daerah diatas $+1,0$ SD frekuensinya $15,87$ %

Ada 3 macam rumus Standar Deviasi yang disesuaikan dengan cara pendistribusian skor. (Buka kembali Statistik 1).

- a. Rumus dalam data single ialah :

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

SD = Standar Deviasi

x = Deviasi

N = Jumlah subyek

Tabel 5.5 Perhitungan Standar Deviasi Dalam Data Tunggal

Nama	Skor (X)	Deviasi x	Deviasi kuadrat x ²
Sumur	34	12,97	168,22
Ramsal	34	12,97	168,22
Sahat	31	9,97	99,40
Hamonangan	30	8,97	80,46
Masotance	29	7,97	63,52
Oriana	29	7,97	63,52
Riko	28	6,97	48,58
Mual	27,5	6,97	48,86
Bernard	25	3,97	15,76
Oereka	25	3,97	15,76
Tertiana	24	2,97	8,82
Lestari	23	1,97	3,88
Kitty	23	1,97	3,88
Gondaria	22	0,97	0,94
Jakobus	22	0,97	0,94
Henri	21,5	0,47	0,22
Norma	21	-0,03	0,01
Octaria	20,5	-0,53	0,28
Courad	20	-1,03	1,06
Paulus	19	-2,03	4,12
Kadir	19	-2,03	4,12
Ulina	18,5	-2,53	6,40
Yanti	16	-5,03	25,30
Veronika	15	-6,03	36,36
Ida	14	-7,03	49,42
Waluyo	12	-9,03	81,54
Dameria	10	-11,03	121,66
Ariani	8	-13,03	169,78
Bonar	6	-15,03	225,90
Fendy	4	-17,03	290,02
N = 30	Σ 631		Σ 1.799,95

$$SD = \frac{\sqrt{\Sigma x^2}}{N}$$

$$SD = \frac{\sqrt{\Sigma 1799,95}}{30}$$

$$SD = \sqrt{59,99}$$

$$SD = 7,75$$

- b. Rumus menghitung Standar Deviasi dalam bentuk data yang dikelompokkan yaitu data kelompok yang tidak memakai interval yaitu :

$$SD = \frac{\sqrt{\sum fx^2}}{N}$$

SD = Standar Deviasi
 F = Frekuensi
 x = Deviasi
 N = Jumlah subyek

- c. Rumus menghitung Standar Deviasi dalam bentuk data dikelompokkan dan memakai interval yaitu :

$$SD = \sqrt{\left\{ \frac{\sum fx^2}{N} \right\} - \left\{ \frac{fx}{N} \right\}^2}$$

3. Normal Curve

Kurve normal dipergunakan untuk mengetahui bagaimana distribusi skor itu. Distribusi yang paling normal adalah apabila sesuai dengan distribusi yang digambarkan oleh normal curve. Istilah normal curve disebut juga bell shaped curve atau normal probability kurve. Kurve normal berarti bagian kiri dan kanan simetris dengan biasanya letak Mean, Mode dan Median berhimpit menjadi satu titik.

D. Pembuatan Nilai/Konversi Skor Menjadi Nilai

Apa yang diuraikan diatas barulah merupakan analisa responden dalam rangka kedudukannya dalam kelompok. Di Sekolah analisa/interpretasi skor kepada nilai adalah sangat penting. Seorang guru sering membuat nilai langsung dari hasil tes. Hal ini terutama bila ia melaksanakan tes esai. Tanggapan ini adalah keliru, karena nilai yang pertama itu barulah merupakan skor mentah. Setiap nilai skor diperoleh pada pertama dilaksanakan pemeriksaan pekerjaan murid belumlah merupakan nilai akhir. Apalagi nilai yang akan diambil adalah merupakan evaluasi terhadap hasil belajar dalam rangka penerimaan suatu sertifikat (penilaian hasil belajar dalam rangka mengetahui apakah seseorang itu telah dapat dianggap berhasil dalam suatu tingkatan program studi).

Dalam interpretasi skor test ini masih banyak terdapat ketidak seragaman. Pada suatu pihak kita lihat adanya praktek penilaian yang tidak bersistem sama sekali. Tapi pada pihak lain ada suatu system yang dianut. Sebahagian besar memuat nilai itu dalam bentuk angka yakni dari 0 s/d 100, atau dari 0 s/d 10. Sebahagian lagi membuat nilai itu dalam bentuk huruf, hal ini terdapat pada banyak sekolah di Amerika Serikat misalnya A, B, C, D, dan E.

Bentuk interpretasi yang dibuat juga masih belum seragam. Dari sekian banyak variasi cara interpretasi umumnya kita temui beberapa yang umum. Setiap cara itu menunjukkan dasar pandangan tertentu. Baiknya cara itu dapat dicek dengan membandingkannya dengan tujuan kita mengadakan penilaian. Beberapa cara yang umum dapat dilihat pada uraian berikut ini.

1. Perbandingan skor yang diperoleh dengan skor maksimum yang mungkin dicapai dalam test tersebut.

Dalam cara ini dipergunakan rumus nilai :

$$\text{Nilai} = \frac{\text{Skor mentah}}{\text{Jumlah item}} \times 100 \text{ (nilai 0 s/d 100)}$$

Atau

$$\text{Nilai} = \frac{\text{Skor mentah}}{\text{Jumlah item}} \times 10 \text{ (nilai 0 s/d 10)}$$

Contoh :

Seorang anak yang dalam mengerjakan suatu test objektif berbentuk benar salah (jumlah item atau skor maksimum 40) terdapat sebagai berikut :

Benar dikerjakan = 30 item

Salah dikerjakan = 5 item

Tidak dikerjakan = 5 item

Skor mentah anak = 30 - 5 = 25

Nilai anak = $\frac{25}{40} \times 100 = 62,5$ (nilai dari 0 s/d 100)

atau = $\frac{25}{40} \times 10 = 6,25$ (nilai dari 0 s/d 10)

Cara ini sebenarnya kurang dapat diterima, karena bisa terjadi bahwa anak-anak secara keseluruhan skor mentah sangat rendah sekali. Akibatnya nilai anak akan rendah pula semuanya. Keberatan lain ialah bahwa test yang dibuat oleh guru belum tentu valid dan reliable. Karena hal ini terjadi berarti guru sudah menganggap ia sudah benar-benar berbobot dalam memberi pelajaran. Artinya metode, cara dan bahan yang diberikannya benar-benar telah cocok. Masalah ini tentunya harus mendapat analisa yang lebih mendalam lagi. Teknik ini termasuk penilaian secara patokan (PAP) dan sangat cocok bila system pembelajaran tuntas (Mastery Learning).

2. Passing grade

Hal ini dapat menghindarkan masalah yang dimajukan diatas. Dengan metode ini memungkinkan pendistribusian nilai berdasarkan curve probability normal. Passing grade berdasarkan kepada pembuatan nilai pada Mean. Jadi kita lebih dahulu menghitung Mean dan selanjutnya menentukan nilai yang

terletak pada Mean. Pengambilan Mean sebagai patokan yang pertama untuk menentukan nilai mengingat Mean adalah pengukuran berdasarkan keseluruhan distribusi skor. Nilai pada Mean biasanya 5 atau 5,5 bagi nilai dari 0 s/d 10 atau 50 dan 55 bagi nilai dari 0 s/d 100.

Setelah kita menentukan nilai pada Mean maka berarti skor yang sama dengan Mean tersebut diberi nilai yang ditetapkan tadi, yaitu 5 atau 5,5 ; 5,5 atau 55 selanjutnya untuk memberi nilai pada setiap murid dipergunakan rumus berikut :

$$N_x = N_M + \frac{X - X_M}{X_M} \times N_M$$

N_x = Nilai sesuatu skor (yang akan dicari nilai)

N_M = Nilai pada Mean

X = Sesuatu skor (yang akan dicari nilai)

X_M = Skor yang terletak pada Mean

Contoh :

Berdasarkan pada perhitungan sebelumnya bahwa Mean dalam distribusi itu = 21, 03 dibulatkan 21, jadi nilai untuk skor 21 adalah : (umpamanya ditentukan 50). Jadi nilai Norma = 50, dan skor norma = 21. Misalkan kita mau mencari nilai untuk Bonda.

Skor Bonda = 29

Nilai pada Mean = 50

Skor pada Mean = 21

$$\text{Jadi nilai Bonda} = 50 + \frac{29 - 21}{21} \times 50 = 69,05$$

Demikianlah kita hitung satu persatu sampai semua skor anak ditukar menjadi nilai.

Catatan : Untuk menghindarkan jangan sampai ada nilai dibawah nol atau diatas seratus, maka skor yang nampaknya sangat ekstrim (tinggi maupun rendah), skor ini disisihkan dari perhitungan dalam mencari Mean dan untuk skor mentah yang ekstrim tinggi diberi nilai tertinggi (100 atau 10) sedang skor yang ekstrim rendah diberi nilai terendah (0).

3. Standar skor

Untuk ini diperlukan lebih dahulu perhitungan Mean dan Standar Deviasi.

Rumus Standar skor :

$$Z (\text{Skor}) = \frac{X - M}{SD}$$

Z = Standar Skor
 X = Skor Mentah
 M = Mean
 SD = Standar Deviasi

Rumus nilai : $Z(\text{skor}) = 10 \times Z + 50$
 10 dan 50 = nilai tetap
 Z = Standard skor (hasil perhitungan melalui rumus standar skor).

Contoh : Berdasarkan nilai terdahulu.

Mean dalam distribusi itu 21,03

SD dalam distribusi itu 7,75

Perhitungan nilai untuk lestari kita lakukan sebagai berikut :

Skor Lestari : 23

Jadi : $23 - 21,03$

Standar skor Lestari = $\frac{\quad}{7,75} = 0,25$

Nilai Lestari = $10 \times 0,25 + 50 = 52,5$

Demikianlah seterusnya setiap skor mentah ditukar ; pertama ke standar skor, kemudian kepada nilai melalui rumus-rumus tersebut.

4. Skala Sigma

Interpretasi melalui skala sigma adalah teknik interpretasi melalui Mean dan Standar Deviasi. Cara ini termasuk teknik interpretasi yang dianjurkan dalam kurikulum sekolah Dasar sampai Sekolah Lanjutan Tingkat Atas. Skala nilai yang dipergunakan ialah dari 1 s/d 10. Setiap nilai ditentukan melalui Standar Deviasi. Melalui penentuan Standar Deviasi dapat ditentukan pula berapa skor yang dituntut.

Berikut ini dapat dilihat table konversi yang dipakai untuk teknik skala sigma ini.

Tabel 5.6 Tabel Konversi Skor Kedalam Nilai Berskala 1 – 10

Skala Sigma	Skala Nilai 1 – 10	Skala Skor
+2,25 SD	10	Mean +2,25 x SD
+1,75 SD	9	Mean +1,75 x SD
+1,25 SD	8	Mean +1,25 x SD
+0,75 SD	7	Mean +0,75 x SD
+0,25 SD	6	Mean +0,25 x SD
-0,25 SD	5	Mean -0,25 x SD
-0,75 SD	4	Mean -0,75 x SD
-1,25 SD	3	Mean -1,25 x SD
-1,75 SD	2	Mean -1,75 x SD
-2,25 SD	1	Mean -2,25 x SD

Langkah perhitungan :

1. Hitung Mean dari distribusi Skor
2. Hitung Standar Deviasi
3. Tentukan besar skor untuk setiap skala sigma/skala nilai
4. Transformasikan setiap skor anak kepada nilai melalui skala skor yang dihitung dalam langkah ketiga.

Catatan : Mungkin Skor tidak sama benardengan skala skor yang dihitung. Untuk ini diberi penambahan/pengurangan nilai yang sebanding dengan kelebihan/kekurangan dari skala skor yang terdekat.

Berikut ini diberi contoh mentransformasikan skor kepada nilai melalui skala sigma ini. Data diambil dari hasil yang disebut dimuka.

$$\begin{aligned}\text{Mean Distribusi} &= 21,03 \\ \text{Standar Deviasi} &= 7,75\end{aligned}$$

Dari data ini dapat kita hitung Skala skor yang dituntut oleh setiap nilai, sebagai berikut :

Tabel 5.7 Perhitungan Skala Skor Untuk Setiap Nilai 1 – 10

Skala Nilai 1 – 10	Perhitungan Skala Skor
10	$21,03 + 2,25 \times 7,75 = 38,47$
9	$21,03 + 1,75 \times 7,75 = 34,59$
8.	$21,03 + 1,25 \times 7,75 = 30,72$
7.	$21,03 + 0,75 \times 7,75 = 26,84$
6.	$21,03 + 0,25 \times 7,75 = 22,97$
5.	$21,03 - 0,25 \times 7,75 = 19,09$
4.	$21,03 - 0,75 \times 7,75 = 15,22$
3.	$21,03 - 1,25 \times 7,75 = 11,34$
2.	$21,03 - 1,75 \times 7,75 = 7,47$
1.	$21,03 - 2,25 \times 7,75 = 3,59$

Dari perhitungan pada table ini dapat dilihat bahwa untuk nilai 10 harus ada skor 38,47 (dibulatkan 38).

5. Deviasi Rata-Rata

Teknik ini dipergunakan bila skor akan ditransformasikan kepada nilai dalam bentuk huruf. Huruf biasa yang dipakai adalah : A, B, C, D, dan E. Dalam system Kredit di Perguruan Tinggi, nilai yang dipergunakan biasanya adalah nilai huruf.

Teknik deviasi rata-rata ini ada 2 macam yaitu :

a. Deviasi rata-rata dengan unit skala 1,5 (ditulis DR 1,5)

b. Deviasi rata-rata dengan unit skala 1,2 (ditulis DR 1,2)

Kedua variasi ini tidak banyak berbeda. Dengan memakai DR 1,5 berarti jumlah orang yang memperoleh nilai A dan E lebih banyak bila dibandingkan dengan bila memakai DR 1,2 sedang nilai C lebih sedikit. Untuk mendapatkan gambaran tentang perbedaan persentase antara DR 1,5 dengan DR 1,2 dan dalam hubungannya dengan distribusi normal dapat dilihat dari table berikut :

Tabel 5.8 Persentase Nilai A, B, C, D, dan E Melalui Perhitungan DR 1,5 dan DR 1,2 Bila Distribusi Dalam Keadaan Normal

Nilai	DR 1,5	DR 1,2
A	6,68 %	3,59 %
B	24,175 %	23,84 %
C	38,29 %	45,14 %
D	24,175 %	23,84 %
E	6,68 %	3,59 %

E. Laporan Hasil Penilaian

1. Pertimbangan dalam pelaporan hasil penelitian

Sehubungan dengan pelaporan hasil penilaian perlu memperhatikan hal-hal sebagai berikut :

a. Lambang nilai mudah dimengerti

Karena laporan nilai dimaksudkan untuk disampaikan pada pengguna (orang tua murid kalau disekolah), hendaknya lambing nilai yang dipergunakan mudah dimengerti dan untuk itu hendaknya dalam raport diberi penjelasan arti dari setiap lambing nilai.

b. Bentuk laporan hasil penilaian memungkinkan pengguna dapat terbimbing dalam mengenal kemampuan yang di tes yang selanjutnya dapat memilih, mempertimbangkan serta mengambil keputusan terhadap yang di test.

Di sekolah pada umumnya menyajikan angka tunggal dari suatu mata pelajaran kurang memberi pemahaman kepada orang tua tentang kemajuan anaknya. Adalah lebih baik apabila angka rata-rata kelas dicantumkan disamping nilai murid yang bersangkutan. Dengan menampilkan penilaian tersebut orang tua dapat lebih mengenal kemajuan anaknya.

c. Frekuensi dan penyampaian laporan penilaian dilakukan "sebaik mungkin"

Frekuensi laporan berhubungan dengan waktu, tingkat jabatan, tingkat sekolah. Semakin rendah tingkatan sekolah sebaiknya semakin banyak frekuensi laporan mengingat rentangan kemajuan belajar lebih besar serta sesuai dengan prinsip semakin "dini" bimbingan dilakukan semakin baik hasilnya.

Dalam penilaian prestasi kerja nilai-nilai angka perlu diterjemahkan dalam bentuk nilai kualitatif baik disampaikan secara tertulis maupun dalam bahasa

lisan, agar yang dinilai mendapatkan pemahaman yang cukup tentang kinerjanya selama periode waktu tertentu.

d. Perlu dipertimbangkan pemberian angka "rendah yang ekstrim"

Dalam lambing nilai 1 s/d 10 angka rendah yang ekstrim adalah 1, 2, dan 3. Bila dalam pengolahan nilai akhir mata pelajaran terdapat angka tersebut perlu dipertimbangkan untuk mencantulkannya dalam raport. Dalam hubungan ini perlu diingat bahwa laporan berfungsi sebagai alat motivasi. Dengan penyantuman angka rendah yang ekstrim dapat menurunkan motivasi, malah dengan penonjolan angka kurang dalam buku raport (dituliskan dengan tinta merah) mempunyai efek psikologis yang "rasanya" kurang baik.

2. Pengisian raport

Selain angka mata pelajaran dalam rapor dilaporkan aspek kepribadian seperti : kelakuan, kerajinan, dan absensi. Demikian juga halnya dengan saran-saran. Perlu kiranya diingat bahwa dalam rangka pengisian angka/nilai kelakuan, kerajinan dan absensi haruslah didasarkan pada pencatatan yang teratur.

Demikian juga dengan pemberian saran-saran, hendaknya jelas dan praktis. Kata-kata "lebih giat belajar " atau "Tingkatkan motivasi kerja" adalah lebih operasional dari sebutan "perbanyak waktu belajar di rumah" atau "jangan terlalu banyak melamun dalam bekerja" dan lain-lain.



RANGKUMAN

1. Pemahaman pendidik dan penguji terhadap statistik sederhana sangat perlu dalam rangka pengolahan penafsiran hasil pengukuran, pengujian, dan penilaian hasil belajar.
2. Perhitungan dengan ukuran tendensi ~~sentral~~ dan variabilitas dari suatu distribusi skor mentah suatu tes dapat memberi gambaran tentang mutu kelas, keadaan kelas dan kedudukan individu dalam kelompok yang selanjutnyadapat dipergunakan dalam perencanaan pengajaran dan perbaikan proses belajar mengajar.
3. Penentuan nilai dari suatu hasil pengukuran/test, harus dilakukan dengan suatu teknik yang disepakati sebelumnya sehingga objektivitas nilai dapat dicapai setinggi mungkin.
4. Hasil pengukuran dan testing selalu ditujukan untuk menganalisa keberhasilan peserta didik mengikuti suatu program pengajaran dan sekaligus juga menganalisa keberhasilan pendidik dalam melaksanakan program dimaksud yang akhirnya dapat dimuarakan dalam perbaikan dan pengembangan program pengajaran menuju mutu yang lebih tinggi.

BAB VI KAWASAN PENGUKURAN

Benyamin S. Bloom dkk, membagi kawasan belajar yang mereka sebut sebagai tujuan pendidikan menjadi tiga bagian yaitu kawasan kognitif, kawasan efektif dan kawasan psikomotor. Tes prestasi belajar, secara luas tentu mencakup ketiga kawasan tujuan pendidikan tersebut. Tes yang mengukur performansi maksimal termasuk tes inteligensi, tes bakat, tes prestasi belajar, tes profisiensi dan berbagai tes kemampuan lainnya lebih banyak berorientasi pada pengukuran kawasan kognitif. Oleh karena itu kita akan membatasi pembahasan kita secara khusus hanya pada kawasan kognitif saja dengan penekanan pada bentuk tes tertulis.

Dimensi kognitif berhubungan dengan segenap kemampuan intelektual. Dalam taksonomi Bloom dikenal 6 jenjang dimensi kognitif. Jenjang satu lebih tinggi dari jenjang yang lain, dan jenjang yang lebih tinggi akan dapat tercapai apabila jenjang yang rendah sudah dikuasai. Oleh sebab itu hubungan setiap jenjang bersifat hierarkis. Berdasarkan urutan dari terendah ke yang tinggi, keenam jenjang tersebut adalah :

1. Pengetahuan

Aspek ini mengacu pada kemampuan manusia dalam mengenal/mengingat semua jenis materi informasi yang diperoleh mulai dari yang sederhana sampai pada hal sukar. Yang penting disini mengingat keterangan dengan benar. Pada umumnya unsure pengetahuan ini menyangkut hal-hal perlu diingat seperti batasan, peristilahan, pasal, hukum, dalil, rumus, nama orang, nama tempat, dll. Penguasaan hal tersebut memerlukan hafalan dan ingatan. Tujuan dalam tingkatan pengetahuan ini termasuk kategori paling rendah dalam domain ini. Namun meskipun jenjang ini terendah tetapi sangat penting. Tanpa memiliki pengetahuan, seorang tidak mungkin mengembangkan kemampuan-kemampuan lain yang ada di atasnya. Informasi yang diterimanya dimasukkan ke dalam ingatan dan disimpan. Apabila informasi tersebut diperlukan, maka informasi tersebut dipanggil untuk digunakan. Proses pengolahan informasi boleh dikatakan tidak ada. Oleh karena itu, jenjang ini dianggap yang terendah dilihat dari proses berfikir.

2. Pemahaman

Aspek ini mengacu pada kemampuan memahami materi yang dipelajari. Pada umumnya unsur pemahaman ini menyangkut kemampuan menangkap makna suatu konsep yang ditandai antara lain dengan kemampuan menjelaskan arti suatu konsep dengan kata-kata sendiri. Informasi yang diterima disini tidak hanya disimpan begitu saja. Informasi tersebut diolah lebih lanjut menjadi sesuatu yang lebih tinggi kedudukannya. Kemampuan mengolah informasi itulah yang diharapkan untuk dikembangkan. Dalam tingkat pemahaman ini ada tiga subkategori kemampuan pokok yang dapat dibedakan. Ketiga kemampuan tersebut adalah kemampuan penerjemahan (misalnya dari lambang kearti), penafsiran, dan ekstrapolasi (menyimpulkan dari sesuatu yang telah diketahui).

3. Aplikasi/Penerapan

Aspek ini mengacu pada kemampuan menggunakan atau menerapkan pengetahuan yang sudah dimiliki pada situasi yang baru, yang menyangkut penggunaan aturan, prinsip dan sebagainya dalam memecahkan persoalan tertentu. Jadi dalam aplikasi harus ada konsep, teori, hukum, rumus, kemudian diterapkan atau digunakan dalam memecahkan suatu persoalan. Unsur kreativitas juga diperlukan dalam mengembangkan kemampuan aplikasi.

4. Analisis

Aspek ini mengacu pada kemampuan mengkaji atau menguraikan sesuatu kedalam komponen-komponen atau bagian-bagian yang lebih spesifik, serta mampu memahami hubungan diantara bagian yang satu dengan yang lain sehingga struktur dan aturannya dapat lebih dipahami. Disini orang harus melakukan pemisahan atas komponen-komponen dari suatu objek/informasi. Kemampuan analisis dapat pula melakukan klasifikasi lebih lanjut dari komponen-komponen yang sudah didaftarkan. Kemampuan analisis yang lebih tinggi adalah menentukan sifat keterhubungan antara berbagai komponen tersebut. Kemampuan analisis yang paling tinggi adalah kemampuan menemukan prinsip atau dasar, organisasi untuk keseluruhan informasi yang dikaji.

5. Sintesis

Aspek ini mengacu pada kemampuan memadukan berbagai konsep atau komponen sehingga membentuk suatu pola struktur atau bentuk baru. Aspek ini memerlukan tingkah laku yang kreatif. Kemampuan sintesis (membentuk) relatif lebih tinggi dari kemampuan analisis (menguraikan).

Kemampuan ini dituntut apabila kita menghadapi informasi yang berbeda-beda dan dari informasi yang berbeda-beda tersebut harus kita padukan sehingga menghasilkan sesuatu harmoni (menyelesaikan pertentangan yang ada) yang sifatnya orisinal.

6. Evaluasi

Aspek ini mengacu pada kemampuan memberikan pertimbangan atau penilaian terhadap gejala atau peristiwa berdasarkan norma-norma atau patokan-patokan tertentu. Kemampuan ini merupakan kemampuan tertinggi dari dimensi kognitif, sehingga memerlukan semua tingkatan kemampuan sebelumnya. Kemampuan evaluasi adalah kemampuan untuk memberikan pertimbangan mengenai nilai informasi yang ada. Untuk dapat memberikan nilai ini ia harus menggunakan kriteria.

Dari uraian tentang batasan setiap kategori tingkat pada kawasan kognitif, jelas bahwa setiap tingkatan memiliki kemampuan internal yang khusus dan yang satu lebih tinggi tingkatannya dari yang lainnya.

Ada beberapa bentuk kata kerja operasional yang dapat digunakan sebagai kata kunci dalam setiap item tes yang sesuai untuk mengungkap setiap jenjang kemampuan. Kata-kata kerja operasional tersebut seperti yang terdapat pada table dibawah ini.

Tabel 5.1 Kata Kerja Untuk Menunjukkan Tingkat Kemampuan Tertentu

Kategori Tkt Kemampuan	Kemampuan Internal	Kata-kata Kerja Operasional
PENGETAHUAN	Mengetahui... Mis : Istilah Fakta Aturan Urutan Metode Cara	Mengenal Menyebutkan Menunjukkan Memberi nama pada Mendesripsikan Memasangkan Menamakan Memilih Mendefenisikan
PEMAHAMAN	Menterjemahkan Menafsirkan Memperkirakan Menentukan Mis : Metode Langkah-langkah Prosedur Memahami Mis : Konsep, Kaidah Kaitan antara fakta Isi pokok Mengartikan/menginter pretasikan.... Mis : Table, Bagan, Grafik, Sket	Menjelaskan Menguraikan Meramalkan Mengklasifikasikan Mengikhtisarkan Mencontohkan Membaca Menyadari Membedakan Mengartikan Menerangkan Mengembangkan
PENERAPAN	Memecahkan masalah Membuat bagan & grafik Menggunakan..... Mis : Metode Prosedur Langkah-langkah Konsep Kaedah Prinsip Patokan	Mendemonstrasikan Menghitung Menghubungkan Mengoperasikan Membuktikan Menghasilkan Menyelesaikan Melengkapi Menyusun Menemukan

ANALISIS	<p>Mengenali kesalahan Membedakan..... Mis : Fakta dari interpretasi Data dari kesimpulan Menganalisis..... Mis : Bagian-bagian Hubungan antara Struktur dasar</p> <p>Menyusun urutan</p>	<p>Memisahkan Menemukan Memilih Menyisihkan Memilah-milah Membandingkan Mengestimasi Membagi Membuat diagram Menunjukkan hubungan antara Menguraikan</p>
SINTESA	<p>Menghasilkan.... Mis : Klasifikasi Karangan Kerangka teoritis Menyusun..... Mis : Rencana skema Program kerja</p>	<p>Menggabungkan Mengkombinasikan Merumuskan Mereka-reka Menciptakan Mendesain Merangkaikan Menghubungkan Menyimpulkan Merancangkan Membuat pola</p>
EVALUASI	<p>Menilai berdasarkan norma internal Mis : Hasil karya seni Mutu karangan Mutu pekerjaan Mutu ceramah Menilai berdasarkan norma eksternal.... Mis : Hasil karya seni Mutu karangan Mutu ceramah Program penataran Mempertimbangkan..... Mis : Baik buruknya Pro-kontra, untung rugi</p>	<p>Memperbandingkan Menyimpulkan Mengkritik Mengevaluasi Menimbang Memberi argumen Menafsirkan Membahas Memilih antara Memberi alasan Membedakan Mengklasifikasikan Mendukung Mematokkan Menyokong Menolak</p>

Kadang-kadang terdapat bagian pelajaran yang tidak dapat atau sukar untuk diungkap kemampuannya pada tingkat tertentu. Misalnya untuk mengungkap kemampuan siswa dalam pelajaran sastra Jawa, akan sangat sulit bila dilakukan pada tingkat analisis, karena itu harus ditulis item dalam tingkat penguasaan lainnya.

BAB VII PENILAIAN PORTOFOLIO

Bagi dunia pendidikan dan pengajaran di Indonesia, penilaian portofolio belum banyak dikenal / dikembangkan. Kurangnya perhatian akan penilaian portofolio mungkin disebabkan besarnya perhatian guru untuk mengajar pencapaian target kurikulum yang dirasakan sangat padat dan luas dibandingkan dengan waktu yang tersedia. Hal ini ialah banyak / kompleksnya masalah administrasi proses belajar dan pembelajaran yang dituntut.

Pembuatan portofolio erat hubungannya dengan kreatifitas siswa dan guru, sedang menurut hemat penulis proses belajar dan pembelajaran belum banyak diarahkan pada pengembangan kreatifitas siswa dan juga kurangnya bahan – bahan / sumber – sumber yang dapat mendorong tingginya kreatifitas siswa. Walaupun demikian banyaknya kekurangan dalam pengembangan penilaian portofolio, bukan berarti guru berhenti berbuat, namun dianjurkan perhatian guru hendaknya memulai pembuatan portofolio walaupun sesederhana mungkin, karena sangat membantu peningkatan hasil dari proses belajar dan pembelajaran.

1. Makna Penilaian Portofolio

Menurut Tierney Carter, dan Desai (1991), berdasarkan ciri-ciri portofolio yang dimajukannya dapat didefinisikan bahwa portofolio adalah koleksi atau kumpulan sistematik karya baik yang dikembangkan oleh siswa dan guru yang dapat berfungsi sebagai dasar untuk penelaan usaha, perbaikan, proses, dan pencapaian disamping untuk memenuhi tuntutan – tuntutan validitas yang umumnya dicapai oleh prosedur-prosedur pengujian yang lebih formal. Dari pengertian portofolio diatas data diambil beberapa hal penting, yaitu:

a. Penyusunan portofolio haruslah sistimatis.

Pengertian sistimatis dalam hal portofolio mengandung makna, bahwa karyanya baik itu tersusun dalam satu aspek mungkin berdasarkan mata pelajaran / bidang studi / mungkin dalam bidang kemampuan tertentu. Disamping itu makna sistimatis dalam pengertian portofolio, bermakna disusun sehubungan dengan waktu pembuatan karya- karya tersebut.

b. Pembuatan portofolio adalah kerja sama siswa dan guru. pembuatan portofolio adalah tertuju pada murid, artinya karya karya yang baik yang dimasukkan dalam portofolio adalah hasil baik dari murid. Dengan kata lain kreativitas muridlah yang dituntut dalam pembuatan karya baik itu. Seperti yang dikemukakan dimuka, proses belajar pembelajaran yang

untuk berprestasi baik dan semakin baik. Portofolio juga meningkatkan perbaikan praktek-praktek pengujian. Dalam hal ini kelemahan – kelemahan praktek pengujian seperti sukarnya menghilangkan ketidak jujuran siswa sewaktu ujian / sukarnya merumuskan butir soal yang benar benar valid, reliable, praktis dan objektif serta besarnya biaya / dana yang diperlukan dalam praktek- praktek pengujian dapat ditutupi dengan pelaksanaan portofolio.

Portofolio menawarkan sebuah kerangka yang dinamis dan berkelanjutan berdasarkan keadaan nyata dari siswa. Dengan pelaksanaan portofolio guru dan siswa lebih diperdayakan dalam meningkatkan karya-karya yang baik.

3. Bentuk Portofolio

Portofolio merupakan gambaran perilaku siswa dalam belajar jadi bukan sekedar hasil karya(produk). Bagi guru dan siswa portofolio lebih diutamakan untuk mengenal dan mengembangkan, bakat dan kemampuan yang beragam dan bervariasi , oleh karena itu guru dapat mengembangkan bakat dan kemampuan tersebut secara terarah. Agar siswa dapat lebih berkembang sesuai dengan potensinya , hendaknya guru dalam memberi saran / penilaian pada portofolio jangan bersifat kaku, walaupun dalam penilaian dituntut tolok ukur yang sama dan mendasar . dengan saran atau petunjuk yang kaku akan mengurangi peluang kepada para siswa untuk saling berbeda yang satu dengan yang lainnya. Dengan memberikan kebebasan sesuai dengan minat dan kemampuannya akan mengakrabkan diri siswa dengan portofolionya sendiri, untuk lebih tegas dikatakan bahwa pembuatan ortofolio harus menggabarkan sejumlah kemungkinan untuk siswa mengungkapkan kepribadian minat dan gaya belajar masing-masing oleh karena itu saran , petunjuk dan penilaian yang diberikan pada setiap karya siswa harus selalu mendorong untuk berkembang sesuai irama dan potensi masing-masing siswa secara individu. Ragam portofolio dapat dibuat sesuai kebutuhan dan bidang yang diminati namun, secara umum bagian – bagian fisik yang lazim di buat oleh sebuah portofolio adalah sebagai berikut:

- a. sebuah folder
folder adalah tempat menyimpan semua karya yang telah dipilih dan dikomentari, baik oleh siswa maupun guru.
- b. karya-karya terpilih
karya-karya terpilih, adalah semua karya yang dipilih secara cermat oleh siswa dari semua karya yang pernah dihasilkan siswa.
- c. lembaran komentar.
Komentar yang dimaksud adalah komentar dari guru atas karya siswa yang dipilih atau komentar siswa sendiri yang berfungsi sebagai evaluasi diri.

Karya – karya yang dipilih untuk dibuatkan dalam portofolio dapat berupa :

- 1). Karangan – karangan / buku cerita yang disusun oleh para ahli, yang diminati. Untuk hal ini guru harus hati – hati agar perimbangan jenis / macam karangan yang pernah dibaca tersebar sebaik mungkin dan jangan sampai mengurangi waktu untuk membaca buku pelajaran. Hal ini dimaksud karena siswa / I mempunyai kecenderungan yang besar untuk membaca cerita – cerita seperti komik – komik yang beredar banyak dan menarik perhatian.
- 2). Karangan karangan yang dibuat oleh para siswa sendiri dalam pelajaran membaca sudah barang tentu sering guru menyuruh murid membuat sebuah karangan. Dari karangan – karangan yang di buat hendaknya dipilih oleh murid dan guru . guru hendaknya dapat mengarahkan siswa tentang karangan yang pantas untuk dimasukkan dalam portofolio . dalam hal ini guru hendaknya secara bijaksana dan mampu menunjukkan karangan yang terpilih dilihat dari respon / komentar guru terhadap karangan . selain karangan yang dibuat atas respon guru , dalam portofolio juga sangat baik dimasukkan karangan yang dihasilkan murid sebagai hasil inisiatif sendiri untuk memberi makna yang lebih baik hendaknya murid dan guru secara konferensi adalah tindakan yang tepat dalam menunjukkan kekuatan dan kelemahan karangan tersebut.
- 3). Gambar – gambar sebagai hasil dari respon guru dalam memberi pelajaran / atau sebagai hasil kreatifitas siswa / murid. Untuk dapat memberi komentar yang jelas dan tepat hendaknya guru memberikan kesempatan kepada murid untuk menceritakan makna gambar yang dilihat dari pesan gambar yang direncanakan oleh murid. Hal ini juga mendorong peningkatan imajinasi murid dan wawasannya. Selain dari pada dimasukan dalam portofolio dapat juga gambar tersebut dipamerkan dalam kelas , sebagai tindakan reinforcement (hadiah) dalam memantapkan / peningkatan berprestasi.
- 4). Hasil – hasil ulangan / latihan tertentu . seorang murid yang pada suatu waktu mendapat nilai / skor yang tinggi, yang sebelumnya belum pernah diperolehnya adalah baik apabila lembaran kerjanya dimasukkan dalam portofolio sudah barang tentu lembaran kerja tersebut selain tercantum nilai dari guru perlu dibarengi dengan komentar – komentar guru misalnya” kamu telah berhasil dengan baik” atau” bapak bangga dengan nilai yang kamu peroleh”. Selain dari pada itu lembaran kerja sedikit kurang berhasil (belum sempurna) , mungkin baik juga dimasukan dalam portofolio . dalam hal seperti demikian ini, komentar guru hendaknya didahului tingkat keberhasilannya, untuk itu hendaknya dipilih kata / kalimat yang baik (tidak mencela) dan selanjutnya diberikan komentar tentang kelebihannya (kata /kalimat hendaknya tidak bersifat mencela).

Dari uraian diatas bahwa dalam membuat portofolio harus benar – benar dipertimbangkan hal berikut :

- a. Tujuan portofolio adalah respon meningkatkan aktifitas siswa dalam belajar.

- b. Portofolio dibuat atas inisiatif murid / siswa sedangkan guru berperan dalam mengarahkan , menunjukkan serta merangsang siswa agar berprestasi lebih baik .
- c. Portofolio jangan sampai menimbulkan persaingan yang tidak sehat . hal ini dimaksudkan bagi murid – murid yang kurang mampu dalam menyediakan bahan fisik dalam pembuatan portofolio , dan murid yang kurang mampu dalam berprestasi, guru harus memberikan kesan bahwa portofolio sifatnya individual.

4. Portofolio dan testing .

Berikut ini diberikan perbandingan antara portofolio dengan testing .

Portofolio	Testing
1.memuat srangkaian gambar / bacaan hasil karya siswa (hasil sesuai dengan apa yang dikerjakan) 2.siswa terlibat dalam menilai kemajuannya sendiri. 3.penilaian siswa secara individual sesuai kemampuan masing-masing 4.penilaian bersifat Kola Boratif 5.meningkatkan kreatifitas penilaian diri sendiri. 6.meningkatkan penilaian dan belajar dan pembelajaran.	1.menilai siswa dalam serangkaian tugas membaca sesuai atau mungkin tidak sesuai dengan apa yang dilakukan siswa. 2.guru sendiri yang memberi nilai Tanpa murid diikut sertakan. 3.menilai siswa dalam hal yang sama, tanpa melihat individual 4.penilaian tidak bersifat Kola Boratif 5.tidak meningkatkan penilaian diri sendiri 6.terdapat pemisahan yang nyata antara rabelajar , ujian dan pembelajaran.

5. Portofolio didalam kelas

Portofolio memberi kemungkinan pada anak untuk mengembangkan minat dan bakat sesuai dengan potensinya sendiri.

Hal ini berarti ragam nya isi portofolio bisa terjadi sebanyak siswa yang terdapat dalam kelas .

a. Portofolio di SD WESTERVILLE

Dis Sekolah ini dikembangkan portofolio tentang – pengalaman – pengalaman buku yang pernah dibaca, kegiatan – kegiatan menulis, kegiatan – kegiatan konferensi dan mengembangkan portofolio mereka dalam sejumlah kegiatan seperti:

- Menyortir proyek- proyek mereka untuk menentukan karya – karya yang mereka anggap istimewa

- Mengumpulkan komentar – komentar dari kawan – kawan dan guru – guru mengenai mengapa mereka anggap tulisan – tulisan tersebut istimewa
- Mendiskusikan karya – karya beserta komentar komentarnya dengan teman sekelas, para guru dan para orang tua
- Menentukan tujuan berikutnya bagi mereka sendiri dengan bimbingan guru.

Sebelum mengembangkan portofolio para siswa dinilai berdasarkan produk dan gurulah yang menentukan tingkat pencapaian dan tujuan belajar selanjutnya, setelah portofolio, guru dan siswa menilai upaya perbaikan proses dan prestasi siswa disamping kemampuan siswa sendiri dalam memulai kemampuan dan menetapkan tujuan – tujuan selanjutnya. Penilaian / evaluasi diri dilakukan secara langsung dan tidak langsung oleh guru. Penilaian adalah berupa angka yang diberikan atas diskusi dengan siswa yang mencakup aspek usaha perbaikan dan proses. Nilai yang diberikan adalah merupakan nilai akhir tidak ada pertimbangan perbandingan dengan murid lain.

b. Portofolio di SD UPPER ARLINGTON OHIO

Disekolah ini portofolio yang diperkenbang adalah berbasis literatur informal. setiap kuartal siswa memilih sejumlah karya dari bacaan (jurnal, kumpulan bacaan) dan tulisan mereka yang ingin dimasukkan kedalam portofolio. bersamaan dengan pengumpulan bahan, siswa mengembangkan tujuan selanjutnya. guru – guru, siswa – siswa dalam memilih bahan – bahan dan mendiskusikan pencapaian belajar penilaian bertujuan meningkatkan upaya dan perbaikan serta memperkembangkan tujuan tujuan berikutnya, tidak untuk sekedar memberikan angka.

6. Pengembangan portofolio

Dari pengalaman Tierney Carter dan Desai pada kelas yang mengembangkan penilaian portofolio, ditemukan hal-hal berikut :

- 1.) Siswa dari TK dan Perguruan Tinggi makin berdaya, terdorong dan apresiatif terhadap pengembangan dan mendiskusikan portofolio mereka.
- 2.) Para siswa memiliki kepemilikan portofolio, memiliki rasa lebih positif serta lebih berkembang terhadap kemajuan – kemajuan dan tujuan – tujuan mereka sebagai pembaca dan penulis secara bersinambungan.
- 3.) Penilaian menjadi lebih kolaboratif dari pada kompetitif
- 4.) Para orang tua turut memperhatikan secara langsung apa yang dicapai para siswa
- 5.) Kegiatan literasi dilakukan para siswa yang sejalan dengan kegiatan intra dan extra sekolah.
- 6.) Guru memperbolehkan wawasan yang lebih kaya dan jelas mengenai siswa
- 7.) Guru memiliki catatan / bukti – bukti penerimaan mengenai Apa yang benar – benar dilakukan siswa
- 8.) Para guru memiliki alat untuk mengembangkan parameter – parameter penilaian yang terpusat pada para siswa serta memiliki focus dalam membantu siswa mengevaluasi diri mereka sendiri

9.) para administrator memiliki alat untuk mengembangkan alat pemeriksaan kelas – kelas dan penampilan individu dalam kelas

Kegunaan portofolio juga dapat dilihat sebagai :

- a. Mengecek kesahihan hasil-hasil test lainnya (perbandingan)
- b. Sumber informasi yang lebih kaya mengenai prestasi, kemajuannya, dan perkembangan kesinambungannya dibandingkan dengan sumber informasi lain yang lebih formal.
- c. Membantu pemahaman banyak orang mengenai aspek-aspek belajar siswa secara positif sesuai kenyataan.

Tugas guru dalam pembuatan portofolio adalah membaca dan menganalisis folder-folder siswa dan membuat komentar-komentar harus diingat hal-hal berikut :

- Jangan membandingkan hasil karya siswa dengan kriteria lain
- Komentar portofolio sebagaimana adanya.
- Berikan waktu yang cukup untuk portofolio
- Berikan respon yang jujur, sesuai keadaan masing-masing anak
- Komentar dibuat sesingkat mungkin
- Komentar dapat berisi saran strategi mengatasi kesulitan pembuatan portofolio
- Menunjukkan gaya, keragaman untuk mengakui kekhasan pribadi murid
- Komentar sejauh mana siswa menggunakan literasi untuk melakukan refleksi terhadap isu-isu, dan untuk mewujudkan minat pribadinya.

7. Konferensi Portofolio

Konferensi portofolio yang dimaksud disini adalah pembicaraan/diskusi anatara siswa dan guru dalam membahas wawasan yang dikandung oleh bahan dalam portofolio. Dalam hubungan ini langkah umum adalah sbb:

1). Siswa menyajikan dan mendiskusikan portofolio dengan guru.

- Berikan kesempatan pada siswa untuk mengontrol portofolionya.
- Tanyakan kepada siswa apa yang mereka anggap merupakan kelebihan-kelebihan portofolionya siswa.
- Guru menunjuk aspek-aspek positif dari portofolio.

2). Mengahiri pertemuan dengan :

- Menciptakan tujuan-tujuan untuk siswa bekerja selanjutnya.
- Menentukan kegiatan portofolio selanjutnya.

8. Penilaian Portofolio

Portofolio yang telah dikembangkan diseleksi ulang, kemudian dikumpulkan menjadi satu untuk dinilai dan diperiksa dengan cermat. Dalam penilaian portofolio hendaknya meliputi hal berikut :

- a. Tolok ukur penilaian dikembangkan oleh seluruh anggota kelas (guru dan murid)

- b. Tolok ukur digunakan dengan cara membandingkan portofolio dengan masing-masing kategori.
- c. Hindarkan membandingkan portofolio antara siswa yang satu dengan yang lain.
- d. Tujuan utama perbandingan portofolio dengan tolok ukur adalah dalam rangka perkembangan siswa untuk digunakan landasan bagi pemberian bantuan dan dorongan perkembangan siswa.
- e. Pedoman analisis (tolok ukur) dilakukan secara individual bukan analisis komparatif diantara sekelompok siswa.

DAFTAR PUSTAKA

- Arikunto, Suharsimi. 1987. *Prosedur Penelitian*. Bina Aksara Jakarta.
- Azwar, Saifuddin. 1999. *Penyusunan Skala Psikologi*. Pustaka Pelajar Offset. Yogyakarta.
- , 1999. *Psikometri*. Pustaka Pelajar Offset. Yogyakarta.
- , 1997. *Reliabilitas Dan Validitas*. Pustaka Pelajar Offset. Yogyakarta.
- , 1996. *Tes Prestasi*. Pustaka Pelajar Offset. Yogyakarta.
- Hadi, Sutrisno. 1987. *Statistik*. Jilid 1. Andi Offset. Yogyakarta
- Hutagalung, Taher & Purba, Edward. 2000. *Evaluasi Pengajaran*. Fakultas Pendidikan. Universitas Negeri Medan.
- Suryobroto, Sumadi. 1984. *Pembimbing Ke Psikodiagnostik*. Edisi II. Andi Offset. Yogyakarta.