

Ambiyar

PENGUKURAN DAN TES DALAM PENDIDIKAN

Materi yang dimuat dalam buku ini lebih difokuskan aspek pengukuran dan tes yang disusun berdasarkan kebutuhan pendidik, instruktur, pengawas sekolah, dan mahasiswa. Buku ini disusun dalam rangka meningkatkan sumber daya manusia (SDM) di Indonesia. Terutama untuk mereka yang sudah bekerja sebagai pendidik di tingkat dasar sampai perguruan tinggi, instruktur dalam program pelatihan pengukuran dan tes, pengawas sekolah yang berkaitan dengan kompetensi evaluasi pendidikan, serta mereka yang sedang dalam pendidikan.

Upaya peningkatan kualitas SDM harus terus-menerus dilakukan dengan berbagai usaha, cara, dan pendekatan pendidikan berkelanjutan maupun pendidikan khusus sehingga dapat menghasilkan SDM yang mumpuni di era otonomi dan global ini.

Materi yang membahas tentang non tes masih sedikit dan terbatas pada alat ukur ceklis, skala bertingkat, dan anecdotal records dalam tes kinerja. Namun demikian aspek afektif dan materi tes yang berkaitan dengan faktor-faktor yang mempengaruhi tes, tes standar, dan penilaian belum dibahas di sini.

UNP PRESS
UNIVERSITAS NEGERI PADANG

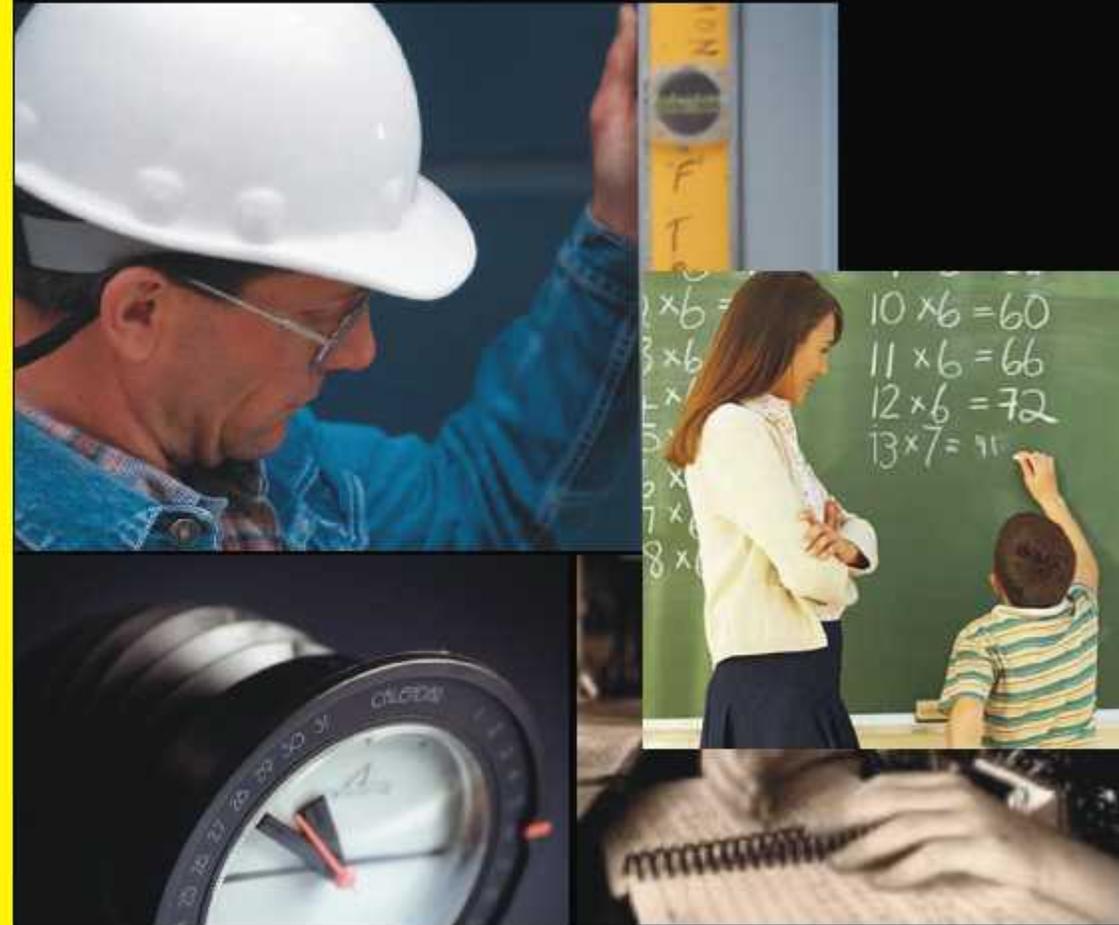
PENGUKURAN DAN TES DALAM PENDIDIKAN

AMBIYAR

PENERBIT UNP PRESS
UNIVERSITAS NEGERI PADANG
Jln. Prof. Hamka, Air Tawar Padang
Telp. (0751) 705253, 7055693, Fax (0751) 7155628



PENGUKURAN & TES DALAM PENDIDIKAN



Ambiyar

Penyunting & Desain
UNP PRESS

KATA PENGANTAR

Buku ini disusun dalam rangka meningkatkan sumber daya manusia (SDM) di Indonesia, baik bagi mereka yang sudah bekerja sebagai pendidik di tingkat dasar sampai perguruan tinggi, instruktur dalam program pelatihan pengukuran dan tes, pengawas sekolah yang berkaitan dengan kompetensi evaluasi pendidikan, dan mereka yang sedang dalam pendidikan. Upaya peningkatan kualitas SDM harus terusmenerus dilakukan dengan berbagai usaha, cara, dan pendekatan pendidikan berkelanjutan maupun pendidikan khusus sehingga dapat menghasilkan SDM yang mumpuni di era otonomi dan global ini.

Materi yang dimuat dalam buku ini lebih difokuskan pada aspek pengukuran dan tes. Materi yang membahas tentang non tes terbatas pada alat ukur ceklis, skala bertingkat, dan *anecdotal records* dalam tes kinerja.

Penulis menyadari akan segala kekurangan dan keterbatasan yang ada pada buku ini. Oleh karena itu diharapkan sumbang saran dari pembaca guna meningkatkan kualitasnya dan diharapkan kehadirannya dapat memberikan manfaat yang lebih banyak bagi pendidik, siswa, mahasiswa, instruktur, kepala sekolah, dan pengawas sekolah.

Ucapan terima kasih disampaikan kepada semua pihak yang telah membantu dan mendorong penyelesaian penulisan buku ini, terutama kepada Prof. Dr. Nizwardi Jalinus, M.Ed dan Prof. Dr. Syafri Anwar, M.Pd sebagai editor materi, Prof Dr. Syahrul, M.Pd dan Dr. H. Abdurahman, M.Pd sebagai editor bahasa, Drs. Ganefri, M.Pd, Ph.D, Drs. Syamsuarnis, M.Pd dan pengelola UNP Press Universitas Negeri Padang. Semoga buku ini menjadi referensi bagi para pembaca, terutama mahasiswa.

Padang, Juni 2011

Penulis.

DAFTAR ISI

	Halaman
KATA PENGANTAR	i
DAFTAR ISI	ii
BAB I. PENDAHULUAN	1
A. Pandangan Umum Pengukuran dan Tes	1
B. Pengertian Pengukuran dan Tes	3
C. Tujuan Tes	9
D. Tes dan Pembelajaran	11
E. Bentuk dan Pendekatan Dalam Tes	15
F. Manfaat Pengukuran dan Tes	18
G. Prinsip Pengukuran dan Tes	20
H. Rangkuman	22
I. Soal-Soal Latihan	24
BAB II. PENGEMBANGAN DAN PENYUSUNAN TES	27
A. Menentukan Tujuan Penilaian, SK, dan KD	28
B. Menentukan Jenis Alat Ukurnya	28
C. Aspek yang Diukur	29
D. Penyusunan Soal dan Kisi-Kisi Tes	35
E. Perakitan Butir Soal	39
F. Rangkuman	41
G. Soal-Soal Latihan	42
BAB III. TES OBYEKTIF	45
A. Tes Benar-Salah	45
B. Tes Menjodohkan	48
C. Tes Pilihan Ganda	50
D. Rangkuman	78
E. Soal-Soal Latihan	80
BAB IV. TES JAWABAN SINGKAT DAN ESSAI	86
A. Tes Jawaban Singkat	86
B. Tes Essai	90
C. Rangkuman	104
D. Soal-Soal Latihan	105
BAB V. TES KINERJA	107
A. Pengertian Tes Kinerja	107
B. Jenis Tes Kinerja	107
C. Tujuan Tes Kinerja	109
D. Kelebihan dan Kelemahan Tes Kinerja	110
E. Kunci Utama yang Diamati Dalam Tes Kinerja	111
F. Petunjuk Penyusunan Tes Kinerja	115
G. Pedoman Penggunaan Tes Kinerja	128
H. Rangkuman	130
I. Soal-Soal Latihan	131

BAB VI.	ANALISIS BUTIR SOAL	134
	A. Analisis Kesukaran Butir Soal	135
	B. Analisis Daya Pembeda Butir Soal	138
	C. Analisis Keberfungsian Distraktor	145
	D. Rangkuman	149
	E. Soal-Soal Latihan	149
BAB VII.	VALIDITAS DAN RELIABILITAS TES	153
	A. Validitas Tes	153
	B. Reliabilitas Tes	158
	C. Rangkuman	168
	D. Soal-Soal Latihan	169
BAB VIII.	INTERPRETASI HASIL TES	171
	A. Interpretasi Acuan Kriteria	172
	B. Interpretasi Acuan Norma	174
	C. Statistik untuk Mendeskripsikan Himpunan Skor	178
	D. Rangkuman.....	188
	E. Soal-Soal Latihan	189
BAB IX.	BIAS TES	191
	A. Pengertian Bias Tes	192
	B. Sifat-Sifat Penilaian Bias	192
	C. Performansi Bias	195
	D. Apa Yang Dilakukan Tentang Bias	197
	E. Pertimbangan Etis dan Prosedur Pemakaian Tes	199
	F. Pertimbangan Lain	202
	G. Rangkuman	204
	H. Soal-Soal Latihan	204
DAFTAR PUSTAKA		209
LAMPIRAN		211

BAB I

PENDAHULUAN

A. Pandangan Umum Pengukuran dan Tes

Sasaran utama pembelajaran, pada dasarnya adalah untuk mengembangkan domain kognitif, afektif, dan psikomotor peserta didik agar mampu menjalani kehidupan sebagai manusia yang paripurna. Untuk itu, kegiatan pembelajaran haruslah merupakan kegiatan yang terencana, terstruktur dan terukur, sehingga bisa diketahui apakah pembelajaran tersebut berhasil atau tidak. Oleh karena itu, kegiatan pengukuran pendidikan di sekolah juga harus mencakup tiga aspek yang menjadi sasaran utama pembelajaran. Aspek kognitif merupakan aspek pembelajaran yang berhubungan dengan kemampuan berpikir dan memahami konsep-konsep yang dipelajari. Aspek ini terentang dari jenjang yang paling rendah yaitu pengetahuan, pemahaman, aplikasi, analisis, sintesis dan sampai ke jenjang yang paling tinggi yaitu, evaluasi. Setiap tahapan dapat diukur tingkat pencapaiannya dengan menggunakan berbagai alat dan tehnik yang ada.

Aspek afektif adalah bagian dari sikap, emosi, dan kepribadian yang meliputi kemampuan untuk menerima, merespon, menilai, mengorganisasi, dan internalisasi nilai-nilai dalam karakter. Berbeda dengan aspek kognitif, aspek afektif ini memiliki karakteristik yang tidak gampang untuk terlihat dan terukur, akan tetapi dapat diukur dengan menggunakan alat ukur yang tepat dan representatif, sehingga perkembangan sikap dan emosi serta perilaku siswa dapat diketahui dan dikembangkan untuk mencapai kematangan dan kedewasaan. Adapun aspek psikomotor merupakan bagian dari keterampilan yang harus dimiliki peserta didik untuk dapat menjalankan perannya dalam berbagai bidang yang akan digelutinya. Aspek ini terdiri atas enam kategori, dari yang paling sederhana, yaitu gerakan meniru kemudian gerakan menyusun (*manipulating*), presisi (*precision*), artikulasi, gerakan terampil, dan naturalisasi.

Sehubungan dengan ketiga aspek sasaran belajar di atas, maka guru dari jenjang pendidikan dasar, menengah sampai pendidikan tinggi perlu memahami mengapa mereka perlu melaksanakan pengukuran terhadap setiap aspek tersebut. Pengabaian terhadap satu aspek secara terusmenerus akan mengakibatkan

ketimpangan perkembangan peserta didik dan menghalanginya untuk menjadi manusia yang paripurna. Di samping itu, guru juga perlu mengetahui dalam setiap pembelajaran yang dilaksanakan hal-hal apa saja yang harus diukurnya, karena tidak mungkin semua aspek akan mendapat porsi yang sama pada setiap pembelajaran. Dalam hal ini diperlukan kepiawaian guru untuk mengidentifikasi hal-hal yang menjadi prioritas untuk dilaksanakan di kelas dan diukur ketercapainnya.

Guru juga harus mengetahui kapan pengukuran sebaiknya dilaksanakan. Meskipun secara objektif kegiatan pengukuran pendidikan bersifat kontinu atau berkesinambungan. Selain itu guru perlu memiliki perencanaan yang matang dalam melaksanakan pengukuran. Apakah akan dilaksanakan dalam bentuk ujian harian (blok) atau ujian pada akhir semester. Mungkin juga guru bisa melaksanakan pre-test dan postes setiap pembelajaran. Setiap keputusan yang diambil guru akan berimplikasi pada perencanaan pengukuran yang akan dilakukannya, baik menyangkut waktu, alat, dan analisis terhadap hasil yang diperoleh.

Hal berikutnya yang perlu diketahui guru adalah menyangkut alat ukur yang akan digunakan. Sesuai dengan tujuan yang hendak dicapai dalam kegiatan pembelajaran, guru bisa menggunakan alat ukur yang bervariasi, mulai dari alat ukur berbentuk tes seperti tes esai dan tes objektif, maupun alat ukur nontes seperti skala sikap, kinerja, produk, portofolio atau proyek. Berbagai alat ukur yang ada akan memberi guru berbagai alternatif dalam mengungkapkan pencapaian siswa sesuai dengan tujuan pembelajaran dan indikator yang telah ditetapkan.

Pada saat guru telah mampu melaksanakan pengukuran dengan menggunakan alat ukur yang tepat, maka hal selanjutnya yang perlu diketahui guru adalah menginterpretasikan hasil pengukuran tersebut, sehingga mampu memberi makna terhadap pendidikan. Secara umum guru dapat menggunakan pendekatan acuan norma (PAN) dalam menginterpretasikan atau menggunakan pendekatan acuan kriteria (PAK) yang telah memiliki standar atau kriteria tertentu. Kapan guru harus menggunakan masing-masing pendekatan ini diperlukan pemahaman yang komprehensif akan tujuan pengukuran yang hendak dilakukan.

Oleh karena pendidikan adalah kegiatan terencana yang harus bisa dipertanggungjawabkan kepada berbagai pihak, maka guru harus menguasai bagaimana melaporkan pencapaian siswa dalam pembelajaran, baik kepada siswa maupun kepada orang tua, lembaga pendidikan, dan masyarakat. Untuk itu,

kegiatan pelaporan akan sangat terkait dengan pengukuran yang dilakukan. Semakin tepat pengukuran yang dilakukan, semakin tepat pula hasil yang diperoleh, dan pada akhirnya semakin tepat juga laporan yang akan diterima siswa dan masyarakat. Oleh karena itu kegiatan pelaporan merupakan muara dari berbagai kegiatan pengukuran yang dilakukan guru di sekolah.

Pencapaian hasil pendidikan dan pembelajaran tidak hanya diperlukan oleh siswa dan masyarakat, tetapi juga diperlukan oleh negara dan pemerintah guna mengetahui secara nasional berapa tingkat keberhasilan yang diperoleh oleh sekolah di seluruh wilayah RI. Oleh karena itu, pengukuran hasil belajar dilaksanakan dalam bentuk ujian nasional (UN). Memang masih menjadi perdebatan pro dan kontra tentang pelaksanaan UN, seperti penyamaan soal secara nasional untuk kondisi dan situasi berbagai aspek yang berbeda untuk setiap daerah di Indonesia, atau mata ujian yang tidak meliputi setiap mata pelajaran di sekolah sehingga menyebabkan deskriminasi terhadap bidang keilmuan. Apapun permasalahannya kegiatan pengukuran pendidikan merupakan hal yang tidak bisa diabaikan oleh semua pihak.

B. Pengertian Pengukuran dan Tes

1. Pengertian Pengukuran

Pengukuran dapat didefinisikan sebagai proses untuk memperoleh informasi tentang atribut atau karakteristik dari obyek, orang, atau kejadian dapat yang ditentukan atau dibedakan (Oriondo, *et.al.* 1998:20). Linn dan Grounlund (1996:6) mengatakan pengukuran adalah: *the process of obtaining a numerical description of the degree to which an individual possesses a particular characteristic*. Sementara itu Wiersma dan Jurs (1990: 7) mengatakan, *“measurement is often stated as the assignment of numerals to objects or events according to rules, and a numeral is a symbol, such as, 1, 2, 3, and so on”*. Dari beberapa pendapat itu dapat disimpulkan bahwa pengukuran adalah suatu proses untuk memperoleh informasi tentang atribut atau karakteristik dari obyek, orang atau kejadian agar dapat ditentukan atau dibedakan yang bersifat kuantitatif. Atribut atau karakteristik dari orang yang dapat diukur berkaitan dengan kemampuannya, sikapnya, keterampilannya, kreativitasnya, dan sebagainya.

Sementara kejadian yang berkaitan dengan gempa, atribut yang dapat diukur antara lain kekuatan gempa, dan kedalaman pusat gempa. Begitu pula obyek yang berkaitan dengan pariwisata, atribut yang dapat diukur antara lain keamanan, kebersihan, keindahan, dan biaya. Bagi individu tertentu keamanan merupakan yang utama diperhatikan untuk datang ke obyek pariwisata, sedangkan bagi individu yang lain perhitungan biaya yang diutamakan

Ada atribut yang mudah untuk diidentifikasi atau ditentukan, misalnya mengukur panjang, temperatur, dan waktu, Dengan observasi, seorang dapat dengan mudah menerima atau menceritakan, apakah sesuatu itu panas atau dingin, panjang atau pendek? Alat seperti meteran, termometer, dan jam dapat menentukan atribut lebih tepat, sebab alat tersebut secara spesifik dibuat untuk maksud demikian.

Ada atribut atau variabel yang tidak mudah diobservasi. Atribut-atribut ini biasanya berkaitan dengan karakteristik dari seseorang, misalnya bakat, emosi, intelegensi, prestasi belajar, dan sebagainya. Misalnya bakat wiraswasta memang tidak mudah untuk mengukurnya atau mengobservasinya. Ada tes bakat dibuat oleh Flanagan yang diberi nama dengan *Flanagan Aptitude Clasification Test* (FACT), yang terdiri dari 11 komponen tes bakat. Di dalam manual FACT dari 11 komponen tes bakat yang tersedia, belum ada untuk tes bakat wiraswasta. Penulis menggunakan 4 macam tes bakat, yaitu kode (*code*), ingatan (*memory*), pemahaman (*judgement and comprehension*), dan ungkapan (*expression*) untuk tes bakat wiraswasta. Begitu pula halnya dengan tes bakat tentang sepak bola. Negara dengan penduduk besar seperti Indonesia tentu tidak akan menemui kesulitan mencari 11 sampai 20 orang pemain sepak bola berbakat, jika ada tes bakat yang secara spesifik untuk maksud tersebut. Hal ini merupakan tantangan dan sekaligus peluang bagi ahli psikometri untuk membuat instrumen tes bakat yang spesifik untuk sepak bola dan bakat wiraswasta.

Pengukuran aspek psikologis ini tidak dapat dilakukan dengan observasi, tetapi dapat secara mudah dikuantifikasi. Bagaimanapun, kuantifikasi adalah perlu untuk membuat penentuan atau pembedaan dari atribut menjadi kurang ambigu (bermakna ganda) dan subyektif. Kuantifikasi mengurangi ambigu, seperti ketika berat diungkapkan dalam besaran 50 kg dengan deskripsi berat atau ringan, yang sebelumnya memiliki interpretasi berbeda dari setiap orang ke orang lain.

Dalam pendidikan, kuantifikasi adalah perlu, sebab kuantifikasi memungkinkan informasi dibawa dan diinterpretasikan dengan menjadi sedikit ambigu dan subyektif dari deskripsi kualitatif seperti intelegensi atau prestasi. Tes pendidikan atau inventori psikologis mencoba untuk memperoleh informasi tentang atribut individu dengan cara yang lebih sistematis.

Ada tiga langkah umum yang dilibatkan dalam pengukuran. Pertama, identifikasi dan definisi kualitas dari atribut yang diukur. Kedua, menentukan susunan operasi/kerja dimana atribut dapat dibuat manifestasi dan dapat dipahami (*perceivable*). Ketiga, memantapkan prosedur atau definisi untuk translasi observasi kedalam besaran kuantitatif.

Tindakan kuantifikasi pengukuran dapat mencakup aspek yang luas, mulai dari menghitung hadirin pada tempat bermain sampai menghitung kecepatan (*timing speeds*) partikel atom. Pengukuran dapat menjadi tepat atau estimasi kasar, derajat ketepatan bervariasi dengan tujuan berbeda. Dengan demikian, kualitas pengukuran meliputi skor pengukuran yang bervariasi. Perbedaan ini mempengaruhi arti (*meaning*) terhadap angka-angka yang diperoleh dan dapat pula mempengaruhi bermacam-macam hal, seperti perhitungan. Ini menghasilkan perbedaan level dari pengukuran yang secara umum dinamakan skala.

Apabila mengacu pada kualitas pengukuran, maka secara esensial bermaksud untuk memperoleh sejumlah informasi yang terdapat pada suatu skor yang dihasilkan oleh pengukuran. Mengukur hasil belajar siswa jarang setepat mengukur karakteristik fisik seperti tinggi dan berat. Hasil belajar siswa lebih sulit didefinisikan dan unit pengukuran biasanya tidak unit fisik (*physical units*). Pengukuran pada siswa bervariasi kualitasnya, sehingga butuh skala pengukuran berbeda. Terminologi yang menguraikan tingkatan pengukuran pada skala ini adalah nominal, ordinal, interval, dan rasio.

Istilah nominal, ordinal, interval, dan rasio benar-benar membentuk suatu hirarki. Skala pengukuran nominal adalah kurang canggih (*the least sophisticated*) dan berisi informasi yang sedikit. Skala ordinal, interval, dan rasio meningkat berturut-turut kecanggihannya. Pengaturan suatu hirarki disebabkan informasi yang terdapat di tingkat yang lebih rendah ada juga terdapat di tingkat yang lebih tinggi, bersama dengan data tambahan. Sebagai contoh, skala pengukuran interval berisi semua informasi yang dimiliki skala nominal dan ordinal, ditambah beberapa

masuk ke pelengkap (*supplementary*). Bagaimanapun skala rasio dari atribut yang sama akan berisi informasi yang lebih dari skala interval. Ide ini akan menjadi lebih jelas ketika setiap skala pengukuran diuraikan di bawah ini.

a. Skala Nominal

Skala nominal adalah skala yang bersifat kategorial dengan menggolongkan obyek atau peristiwa. Angka-angka yang ada tidak menyiratkan hitungan apapun, tetapi kategori harus satu sama lain menyeluruh dan eksklusif. Misalnya, orang secara nominal mendesain posisi baseball dengan memberi angka numerik 1 untuk pelempar (*pitcher*), angka 2 untuk penangkap (*catcher*), angka 3 untuk penjaga garis pertama (*first baseman*), angka 4 untuk penjaga garis kedua (*second baseman*), dan sebagainya. Pemberian angka ini sewenang-wenang dan tidak ada perhitungan apapun. Sebagai contoh, 1 tambah 2 tidak sama dengan 3, sebab seorang *pitcher* tambah seorang *catcher* tidak sama dengan seorang *baseman*. Contoh lain adalah agama. Untuk agama Hindu diberi angka numerik 1, agama Kristen diberi angka numerik 2, agama Islam diberi angka numerik 3, agama Budha diberi angka numerik 4, dan agama Protestan diberi angka numerik 5. Apabila angka 1 ditambah 3 tidak sama dengan 4, sebab agama Hindu ditambah agama Islam tidak sama dengan agama Budha.

b. Skala Ordinal

Skala ordinal adalah angka yang dirancang berdasarkan urutan (*rank order*). Sebagai contoh skala pengukuran adalah urutan individu dalam kelas menurut skor tes yang diperolehnya. Skor tes siswa dapat diurut dari pertama, kedua, ketiga, dan seterusnya sampai skor terendah. Skala ini memberi informasi lebih dari skala pengukuran nominal, tetapi skala ini masih memiliki keterbatasan. Unit dari pengukuran ordinal hampir bisa dipastikan berbeda. Banyak poin-poin yang memisahkan kesatu dan kedua para siswa mungkin tidak sama jarak urutannya, begitu juga jumlah poin yang memisahkan siswa keenam dan kelima. Unit pengukuran ini dapat dianalogikan dengan sebuah mistar (*ruler*) yang beberapa sentimeter lebih panjang dibanding dengan yang lain.

c. Skala Interval

Untuk dapat menambahkan dan mengurangi skor, maka dapat digunakan skala interval, kadang-kadang dinamakan dengan interval yang sama atau unit

pengukuran sama. Skala pengukuran ini berisi sifat-sifat nominal dan ordinal serta ditandai oleh unit yang sama antara poin-poin skor. Misalnya, termometer dan kalender tahunan. Perbedaan antara temperatur 10° dan 20° adalah sama dengan antara 47° dan 57° . Demikian juga dengan perbedaan panjang waktu antara 1946 dan 1948 sama dengan antara 1973 dan 1975.. Ukuran ini menjelaskan sifat-sifat fisis, bahwa interval adalah memiliki jarak yang sama. Sebagai contoh, satu tahun adalah waktunya untuk mengira bumi mengorbit matahari. Keuntungan dari unit pengukuran sama adalah secara langsung penjumlahan dan perbedaan bisa dipikirkan secara numerik dan logis. Bagaimanapun titik nol dalam pengukuran interval sesungguhnya suatu keputusan yang tidak mutlak, sebagai contoh, 0 (nol) derajat tidak berarti bahwa tidak ada temperatur.

d. Skala Rasio

Skala rasio merupakan jenis pengukuran yang paling canggih, karena meliputi semua sifat skala pengukuran terdahulu, tetapi pada skala rasio titik nol tidak sewenang-wenang, skor nol menunjukkan adanya apa yang diukur. Sebagai contoh, jika kekayaan seseorang sama dengan nol, maka ia tidak mempunyai kekayaan sama sekali. Ini tidak menguji ilmu sosial, dimana kesalahan setiap item (yakni; memperoleh skor nol), mungkin tidak menunjukkan ketidaklengkapan pengetahuan ilmu sosial. Pengukuran rasio jarang dicapai dalam penilaian pendidikan, baik dalam aspek kognitif maupun afektif. Skala pengukuran rasio, memungkinkan rasio perbandingan, seperti Ann adalah 1,5 kali lebih tinggi dari saudara kecilnya Mary. Begitu pula jarang mengatakan bahwa intelegensi seseorang 1,5 kali lebih besar dari intelegensi orang lain. IQ sebesar 120 mungkin secara numerik 1,5 kali IQ 80, tetapi seseorang dengan IQ 120 tidak 1,5 kali cerdas dari IQ 80.

Dalam merencanakan tes untuk butir soal dari domain spesifik yang mendekati pengukuran rasio haruslah hati-hati. Sebagai contoh, perhatikanlah suatu perkalian pasangan angka-angka kurang dari 10. Ada 45 kombinasi untuk pasangan angka ini. Bagaimanapun, guru mungkin secara acak memilih 5 atau 10 butir untuk diberikan kepada siswa tertentu. Kemudian, proporsi butir soal yang dapat dijawab siswa secara benar bisa digunakan mengestimasi banyak butir soal yang telah dikuasai siswa dari 45 butir. Jika siswa menjawab 4 dari 5 butir secara benar, maka sah mengestimasi bahwa siswa akan mendapatkan 36 butir secara benar dari 45 butir, jika 45 butir telah diadministrasikan. Ini mungkin disebabkan

susunan (set) materi telah didefinisikan secara jelas dalam tujuan pembelajaran, butir soal telah diacak, dan sampel yang representatif dari materi.

Kebanyakan pengukuran dalam bidang pendidikan lebih baik dibanding pengukuran nominal atau ordinal, dan sedikit pengukuran interval. Pengukuran dalam bidang pendidikan umumnya antara skala ordinal dan interval. Studi empiris menunjukkan bahwa perhitungan pada skala ini pantas, dan skor yang diperoleh merupakan informasi untuk banyak keputusan tentang siswa dan pengajaran.

2. Pengertian Tes

Tes secara harfiah berasal dari bahasa Prancis kuno "testum" artinya piring untuk menyisihkan logam-logam mulia. Tes adalah serangkaian pertanyaan atau latihan atau alat lain yang digunakan untuk mengukur keterampilan, pengetahuan, kecerdasan, kemampuan, atau bakat yang dimiliki oleh seseorang atau kelompok.

Berdasarkan definisi tersebut, dapat dijelaskan bahwa tes merupakan alat ukur yang berbentuk pertanyaan atau latihan, dipergunakan untuk mengukur kemampuan yang ada pada seseorang atau sekelompok orang. Sebagai alat ukur dalam bentuk pertanyaan, maka tes harus dapat memberikan informasi mengenai pengetahuan dan kemampuan obyek yang diukur. Sedangkan sebagai alat ukur berupa latihan, maka tes harus dapat mengungkap keterampilan dan bakat seseorang atau sekelompok orang.

Tes merupakan alat ukur yang standar dan obyektif sehingga dapat digunakan secara meluas untuk mengukur dan membandingkan keadaan psikis atau tingkah laku individu. Dengan demikian, berarti sudah dapat dipastikan tes akan mampu memberikan informasi yang tepat dan obyektif tentang obyek yang hendak diukur baik berupa psikis maupun tingkah lakunya, sekaligus dapat membandingkan antara seseorang dengan orang lain. Jadi dapat disimpulkan bahwa tes adalah suatu cara atau alat untuk mengadakan penilaian yang berbentuk suatu tugas atau serangkaian tugas yang harus dikerjakan oleh siswa atau sekelompok siswa sehingga menghasilkan nilai tentang tingkah laku atau prestasi siswa tersebut. Prestasi atau tingkah laku tersebut dapat menunjukkan tingkat pencapaian tujuan instruksional pembelajaran atau tingkat penguasaan

terhadap seperangkat materi yang telah diberikan dalam proses pembelajaran, dan dapat pula menunjukkan kedudukan siswa yang bersangkutan dalam kelompoknya.

Dalam kaitan dengan rumusan tersebut, sebagai alat evaluasi hasil belajar, tes minimal mempunyai dua fungsi, yaitu: *pertama*, untuk mengukur tingkat penguasaan terhadap seperangkat materi atau tingkat pencapaian terhadap seperangkat tujuan tertentu, dan *kedua*, untuk menentukan kedudukan atau peringkat siswa dalam kelompok, tentang penguasaan materi atau pencapaian tujuan pembelajaran tertentu. Fungsi pertama lebih dititikberatkan untuk mengukur keberhasilan program pembelajaran, sedangkan fungsi kedua lebih dititikberatkan untuk mengukur keberhasilan belajar masing-masing individu peserta tes.

C. Tujuan Tes

Dilihat dari segi tujuannya dalam bidang pendidikan, tes dapat dibagi menjadi:

1. Tes Kecepatan (*Speed Test*)

Tes ini bertujuan untuk mengevaluasi peserta tes dalam hal kecepatan berpikir atau keterampilan, baik yang bersifat spontanitas maupun hafalan dan pemahaman dalam mata pelajaran yang telah dipelajarinya. Waktu yang disediakan untuk menjawab atau menyelesaikan seluruh materi tes ini relatif singkat dibandingkan dengan tes lainnya, sebab yang lebih diutamakan adalah waktu yang minimal dan dapat mengerjakan tes itu sebanyak-banyaknya dengan baik dan benar, cepat dan tepat penyelesaiannya. Tes yang termasuk kategori tes kecepatan misalnya tes intelegensi, dan tes keterampilan bongkar pasang suatu alat.

2. Tes Kemampuan (*Power Test*)

Tes ini bertujuan untuk mengevaluasi peserta tes dalam mengungkapkan kemampuannya (dalam bidang tertentu) dengan tidak dibatasi secara ketat oleh waktu yang disediakan. Kemampuan yang dievaluasi bisa berupa kognitif maupun psikomotorik. Soal-soal biasanya relatif sukar menyangkut berbagai konsep dan pemecahan masalah dan menuntut peserta tes untuk mencurahkan segala kemampuannya baik analisis, sintesis dan evaluasi.

3. Tes Hasil Belajar (*Achievement Test*)

Tes ini dimaksudkan untuk mengevaluasi hal yang telah diperoleh dalam suatu kegiatan. Tes hasil belajar, baik itu tes harian (formatif) maupun tes akhir semester (sumatif) bertujuan untuk mengevaluasi hasil belajar setelah mengikuti kegiatan pembelajaran dalam suatu kurun waktu tertentu.

4. Tes Kemajuan Belajar (*Gains/Achievement Test*)

Tes kemajuan belajar disebut juga dengan tes perolehan adalah tes untuk mengetahui kondisi awal peserta tes sebelum pembelajaran dan kondisi akhir peserta tes setelah pembelajaran. Untuk mengetahui kondisi awal peserta tes digunakan *pre-tes* dan kondisi akhir peserta tes digunakan *post-tes*.

5. Tes Diagnostik (*Diagnostic Test*)

Tes diagnostik adalah tes yang dilaksanakan untuk mendiagnosis atau mengidentifikasi kesukaran-kesukaran dalam belajar, mendeteksi faktor-faktor yang menyebabkan terjadinya kesukaran belajar, dan menetapkan cara mengatasi kesukaran atau kesulitan belajar tersebut. Tes ini dilaksanakan untuk bimbingan belajar, pengajaran remedial (*remedial teaching*), menemukan kasus-kasus, dan lain-lain. Soal-soalnya disusun sedemikian rupa agar dapat ditemukan jenis kesulitan belajar yang dihadapi oleh para siswa.

6. Tes Formatif

Tes formatif adalah penggunaan tes hasil belajar untuk mengetahui sejauh mana kemajuan belajar yang telah dicapai oleh siswa dalam suatu program pembelajaran tertentu. Tes ini berorientasi kepada proses belajar mengajar untuk memperbaiki program pengajaran dan strategi pelaksanaannya.

7. Tes Sumatif

Istilah sumatif berasal dari kata "sum" yang berarti jumlah. Dengan demikian tes sumatif berarti tes yang ditujukan untuk mengetahui kompetensi siswa dalam sekumpulan materi pelajaran (pokok bahasan) yang telah dipelajari. Tes ini dilaksanakan pada akhir program, yakni akhir caturwulan, akhir semester, dan akhir tahun. Tes ini berorientasi kepada produk, bukan kepada proses.

8. Tes Penempatan

Tes penempatan adalah tes yang ditujukan untuk mengetahui kemampuan dan keterampilan prasyarat yang diperlukan bagi suatu program belajar dan penguasaan belajar seperti yang diprogramkan sebelum memulai kegiatan belajar untuk program itu. Dengan kata lain, tes ini berorientasi kepada kesiapan siswa untuk menghadapi program baru dan kecocokan program belajar dengan kemampuan siswa.

9. Tes Seleksi

Tes seleksi adalah tes yang ditujukan untuk menseleksi kemampuan dan keterampilan seseorang yang diperlukan bagi suatu lembaga pendidikan atau tempat kerja tertentu.

D. Tes dan Pembelajaran

1. Hubungan antara Tes Hasil Belajar dengan Proses Pengajaran

Tes hasil belajar merupakan tata cara sistematis yang dilakukan untuk mengukur sampel representatif dari materi belajar (*learning task*) siswa. Tujuan dilakukannya tes hasil belajar adalah untuk mengukur seperangkat hasil belajar yang diharapkan. Tes hasil belajar tidak hanya dilaksanakan pada akhir proses pengajaran, juga pada awal pembelajaran dan proses pembelajaran.

Tabel 1.1 Hubungan antara Tes Hasil Belajar dengan Proses Pengajaran

Pembelajaran	Hasil Belajar
Pengajaran sangat efektif bilamana 1. Pembelajaran diarahkan pada sasaran belajar yang sudah ditetapkan secara jelas 2. Metode dan bahan ajar kongruen dengan hasil belajar yang ingin dicapai	Tes sangat efektif bilamana 1. Tes ditujukan untuk mengukur capaian belajar yang sudah ditentukan dengan jelas. 2. Sifat dan fungsi butir soal tes kongruen dengan capaian yang ingin diukur

3. Pembelajaran didisain sesuai dengan karakter dan kebutuhan siswa.	3. Butir tes dirancang untuk memenuhi level bacaan dan karakteristik lainnya dari siswa
4. Keputusan pengajaran didasarkan pada informasi yang valid, dan relevan dengan proses pengajaran	4. Hasil tes memberikan informasi yang berarti, dan relevan dengan proses pengajaran
5. Secara periodik siswa diberitahu mengenai proses belajar mengajarnya	5. Ketentuan dibuat untuk memberikan umpan-balik tentang hasil tes kepada siswa
6. Perbaikan yang disediakan untuk siswa yang tidak mencapai hasil belajar yang ditetapkan	6. Kelemahan belajar diungkap melalui butir tes tersendiri atau butir tes kelompok
7. Melakukan peninjauan terhadap efektivitas pengajaran secara periodik, lalu melakukan modifikasi terhadap hasil belajar yang diharapkan dan pengajaran yang sesuai dengan kebutuhan	7. Hasil tes memberikan informasi berguna untuk mengevaluasi tujuan, metoda dan bahan pengajaran yang sesuai

Sumber : Gronlund, 1993: 2

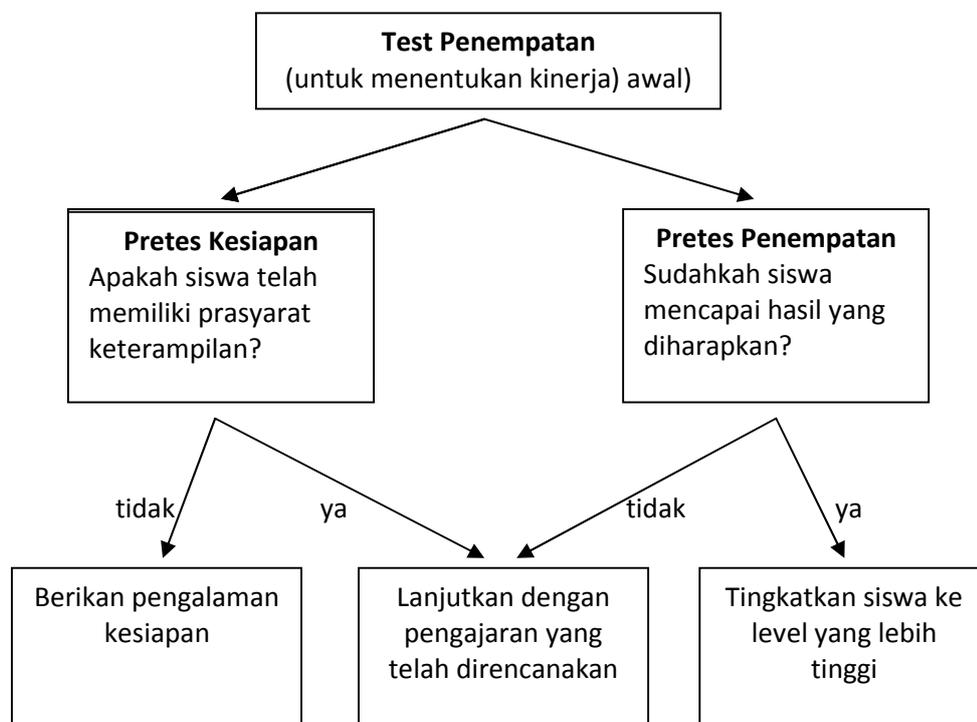
2. Peran Tes dalam Proses Pengajaran

Selama proses belajar-mengajar guru dituntut untuk membuat berbagai keputusan, melalui tes guru dapat meningkatkan efektifitas dalam mengambil keputusan yang obyektif. Berdasarkan waktu keputusan yang harus dibuat oleh seorang guru selama proses belajar mengajar dibagi menjadi tiga, yaitu: (a) pada awal pembelajaran (tes penempatan), (b) selama proses pengajaran (tes formatif dan diagnostik), dan (c) tes akhir proses pembelajaran (tes sumatif). Berikut ini dijelaskan peran tes pada awal, selama, dan akhir proses pembelajaran.

a. Pada Awal Pembelajaran (Tes Penempatan)

Tes untuk maksud ini dilakukan pada pertemuan awal atau pada setiap awal pokok bahasan. Materi tes penempatan mencakup seluruh materi prasyarat

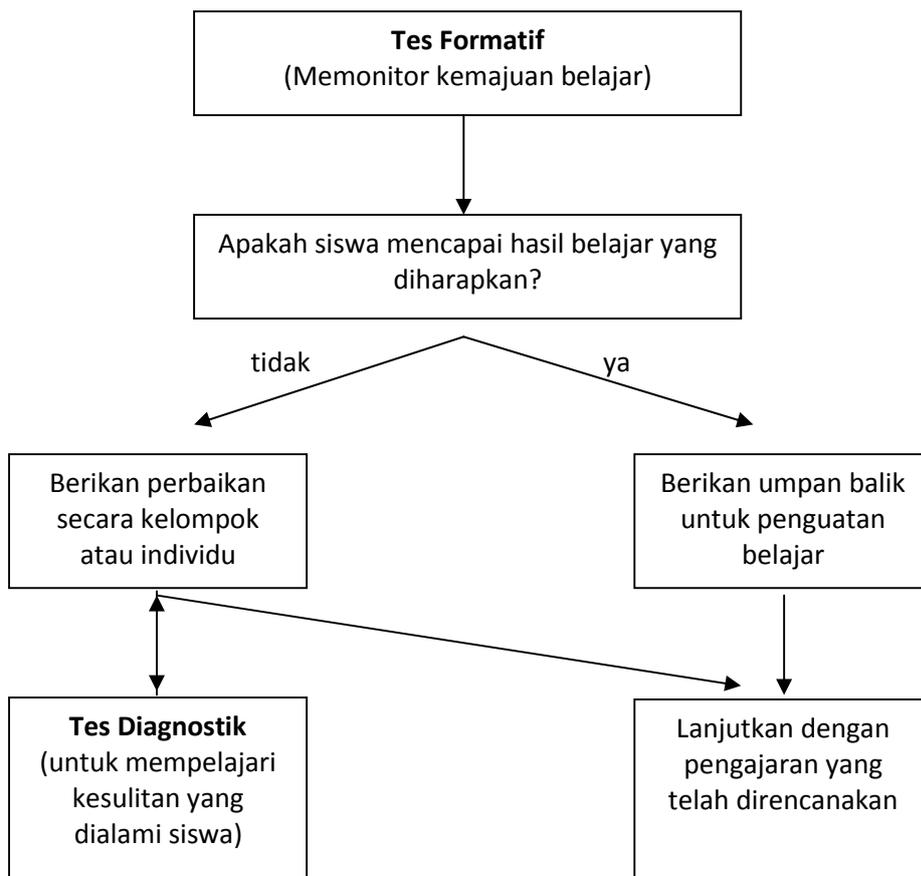
yang dibutuhkan dalam rangka ketercapaian rencana pembelajaran. Dengan adanya tes ini guru dapat mengambil keputusan terhadap hasil tes seperti pada bagan berikut ini



Gambar 1.1 Hubungan Tes Penempatan dengan Pengajaran (Gronlund: 1993:4)

b. Selama Proses Pengajaran (Tes Formatif dan Tes Diagnosa)

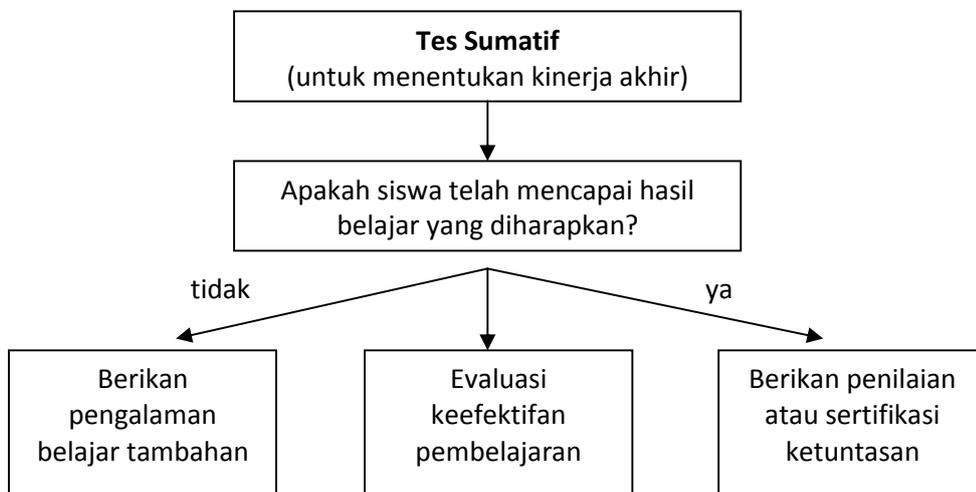
Tes formatif adalah tes yang dilaksanakan pada akhir setiap pokok bahasan dengan tujuan memantau proses atau perkembangan belajar siswa. Tes formatif umumnya mencakup materi satu atau beberapa pokok. Manfaat tes formatif adalah 1) mengukur semua hasil belajar yang diharapkan dalam satu pokok bahasan, 2) hasilnya dapat dimanfaatkan untuk meningkatkan pengajaran. Tes diagnostik adalah tes yang dilakukan untuk mengetahui kesulitan belajar siswa dalam satu pokok bahasan atau materi. Butir soal tes diagnosa dibuat dengan materi yang diasumsikan sulit bagi siswa. Dengan begitu penyebab kesalahan dalam pembelajaran dapat diidentifikasi.



Gambar 1.2. Hubungan Tes Formatif dengan Pengajaran (Gronlund: 1993:5)

3. Akhir Proses Pengajaran (Tes Sumatif)

Tes sumatif adalah tes yang dilaksanakan setelah sekumpulan program pelajaran selesai diberikan (berakhir). Tes ini bertujuan untuk mengukur tingkat penguasaan siswa terhadap seluruh materi pembelajaran.



Gambar 1.3. Hubungan Tes Sumatif dengan Pengajaran (Gronlund: 1993:7)

4. Peranan Lain Tes Dalam Membantu Proses Belajar Mengajar

a. Membantu Memotivasi Siswa

Tes dapat memberikan pengaruh langsung kepada siswa berupa: (1) Memberikan tujuan jangka pendek dari pembelajaran, (2) Memperjelas materi pelajaran yang harus dipelajari siswa, dan (c) Umpan-balik yang terpusat pada proses pembelajaran siswa

b. Membantu Mengingat dan Transfer Pembelajaran

Oleh karena tes didisain untuk mengukur hasil belajar siswa, maka secara tidak langsung tes telah membantu siswa dalam mengulang pelajaran yang telah diberikan.

c. Membantu Evaluasi Diri Siswa

Dengan dilaksanakannya tes, maka siswa dapat mengetahui kelebihan dan kekurangannya dalam belajar, sehingga siswa dapat memperbaikinya di masa yang akan datang.

d. Membantu Mengevaluasi Efektifitas Pengajaran

Hasil penilaian dari tes yang telah dirancang dengan baik dapat memberikan informasi terhadap efektifitas proses pengajaran yang telah berlangsung, seperti kesesuaian capaian belajar, media pengajaran serta metode pengajaran. Dengan begitu pasca pelaksanaan tes proses pengajaran dapat disesuaikan berdasarkan hasil tes siswa.

E. Bentuk dan Pendekatan Dalam Tes

1. Bentuk Tes

Dilihat dari jawaban siswa yang dituntut dalam menjawab atau memecahkan persoalan yang dihadapinya, maka tes hasil belajar dapat dibagi menjadi 3 jenis : 1) tes lisan, 2) tes tertulis, dan 3) tes kinerja atau perbuatan. Penggunaan setiap jenis tes tersebut seyogyanya disesuaikan dengan domain (ranah) perilaku yang hendak diukur. Misalnya tertulis atau tes lisan dapat digunakan untuk mengukur kawasan kognitif, sedangkan domain psikomotor cocok atau tepat diukur dengan tes kinerja, dan kawasan

afektif biasanya diukur dengan non tes, seperti angket (kuesioner) untuk mengukur sikap siswa terhadap pembelajaran.

2. Pendekatan Dalam Tes

Ada dua pendekatan dalam tes, yaitu tes dengan pendekatan acuan norma dan tes dengan pendekatan acuan kriteria.

a. Tes acuan norma

Tes acuan norma adalah tes yang kriteria penilaiannya ditetapkan setelah pelaksanaan tes. Dari analisis skor yang diperoleh siswa ditetapkan kriteria keberhasilan siswa. Dengan demikian kriteria bersifat relatif.

b. Tes acuan kriteria

Pada test acuan normal kriteria keberhasilan siswa ditetapkan terlebih dahulu sebelum tes diadakan. Dengan demikian ada standar minimal dimana siswa harus menguasai suatu materi tertentu atau suatu keterampilan yang diharapkan.

Tabel 1.2. Perbandingan Antara Tes Acuan Norma dengan Tes Acuan Kriteria pada Tes Hasil Belajar

	Tes Acuan Norma	Tes Acuan Kriteria
Prinsip Penggunaan	Tes survai	Tes ketuntasan
Penekan Utama	Mengukur perbedaan individu dalam pencapaian hasil belajar	Mendeskripsikan tugas-tugas yang dapat dikerjakan siswa
Interpretasi Hasil	Membandingkan kinerja hasil tiap siswa	Membandingkan kinerja dengan domain hasil belajar yang telah ditentukan secara jelas
Cakupan Materi	Mencakup materi yang lebih luas	Terfokus pada sekumpulan materi pelajaran (<i>learning task</i>) yang terbatas

Sifat Perencanaan Tes	Menggunakan tabel spesifikasi	Menggunakan domain spesifikasi detil yang disukai.
Prosedur Pemilihan Butir Soal	Butir tes yang dipilih dapat memberikan perbedaan hasil maksimum antar siswa. Butir soal yang mudah umumnya dihilangkan dari tes.	Menyertakan semua butir soal yang dianggap dapat mendeskripsikan kinerja siswa.
Standar kinerja	Tingkat kinerja ditentukan berdasarkan posisi relatif siswa dalam suatu kelompok	Tingkat kinerja ditentukan secara absolut

Sumber : **Gronlund, 1993: 13**

c. Menyusun Tes Acuan Norma dan Tes Acuan Kriteria

Secara umum dalam penyusunan tes acuan kriteria dan tes acuan normal terdapat beberapa persamaan di antaranya: (1) Domain pencapaian hasil belajar telah ditentukan dengan jelas melalui kumpulan tugas pembelajaran (*learning task*) yang relevan dengan pengajaran, (2) Tes direncanakan untuk mengukur sampel yang representatif dari populasi materi sesuai dengan domain pengajaran yang ingin dicapai, (3) Menggunakan butir soal tes yang bervariasi, (4) Menyertakan aturan menjawab soal, (5) Melakukan kontrol terhadap faktor yang dapat berkontribusi terhadap hasil tes, (6) Memperhatikan pada faktor-faktor yang dapat memberikan penafsiran ganda terhadap hasil tes.

Sementara itu ada perbedaan penyusunan dari tes acuan norma dan tes acuan kriteria. Pada tes acuan norma (1) terfokus pada rentangan (*range*) yang luas dari materi pelajaran, (2) memiliki lebih sedikit butir soal, dan (3) soal mudah sehingga memberikan rentangan yang lebih lebar dalam penskoran, sedangkan pada tes acuan kriteria (1) terfokus pada domain materi pelajaran yang dibatasi, (2) memiliki jumlah butir soal tes yang lebih banyak, dan (3) soal menuntut deskripsi rinci dari kinerja siswa.

F. Manfaat Pengukuran dan Tes

Pengukuran dan tes yang dilaksanakan oleh guru akan bermanfaat bagi guru, siswa, dan sekolah.

1. Manfaat Bagi guru

a. Mengetahui tingkat penguasaan siswa.

Dengan adanya tes yang dilaksanakan oleh guru dapat diketahui atau dideteksi siswa yang telah dan belum mencapai kriteria ketuntasan minimal (KKM). Bagi siswa yang belum mencapai KKM diadakan pembelajaran remedial, sedangkan bagi siswa yang telah melampaui KKM diberikan pembelajaran pengayaan.

b. Mengetahui ketepatan materi yang diberikan

Apabila hampir semua mendapat hasil yang tidak memuaskan setelah dilaksanakan tes yang menanyakan suatu materi pokok tertentu, berarti materi pokok tersebut belum dikuasai siswa. Dengan arti kata siswa mengalami kesulitan pada materi tersebut. Guru dapat meneliti lebih jauh keadaan materi yang belum dikuasai tersebut dan mengupayakan perbaikannya. Sebaliknya, apabila hasil dari tes yang dilaksanakan guru menunjukkan hampir semua siswa menguasai materi tersebut, maka materi tersebut tidak perlu diulangi lagi. Lebih lanjut, dengan diadakan tes oleh guru dapat diketahui cakupan dan jenis materi yang telah dikuasai siswa.

c. Mengetahui ketepatan metode yang digunakan

Hasil tes dapat menunjukkan tepat atau tidaknya metode pembelajaran yang dipergunakan oleh guru dalam menyajikan suatu materi tertentu. Apabila hasil tes mengecewakan, terdapat kemungkinan metode pembelajaran yang diterapkan untuk penyajian materi tertentu belum atau tidak sesuai. Guru berkewajiban mencari metode pembelajaran yang mengaktifkan siswa, inovatif, efektif dan menyenangkan (PAIKEM) bagi siswa.

2. Manfaat Bagi siswa

a. Mengetahui kemampuannya

Dengan diadakannya tes dan diperoleh hasilnya, siswa dapat mengukur dan mengetahui tingkat pencapaiannya (kemampuannya). Apa hasil yang telah dicapainya telah memuaskan atau tidak memuaskan. Tentunya siswa yang belum berhasil akan lebih giat lagi menyiapkan diri menghadapi tes pada periode berikutnya agar memperoleh nilai yang baik.

b. Memotivasi belajar

Apabila hasil tes memuaskan siswa, maka siswa tersebut terdorong untuk mengulangi atau bahkan memperbaiki hasilnya. Dengan demikian siswa merasa termotivasi untuk belajar lebih giat lagi agar dapat mempertahankan tingkat pencapaiannya dan meningkatkannya.

c. Umpan Balik

Umpan balik berupa informasi yang diberikan kepada siswa mengenai kemajuan belajarnya perlu diberikan sesegera mungkin. Umpan balik dapat bersifat korektif maupun konfirmatif. Dengan sesegera mungkin memberikan umpan balik dapat dihindari kekeliruan belajar yang berlarut-larut yang dialami peserta didik.

d. Meningkatkan kebiasaan belajar

Apabila hasil tes yang diperoleh siswa kurang atau belum memuaskannya, maka siswa harus merubah kebiasaan belajarnya. Selama ini menggunakan kebiasaan belajar 2 x 5, mungkin dirubah menjadi 5 x 2. Dengan belajar yang teratur, siswa akan lebih siap menghadapi tes yang diberikan guru.

3. Manfaat Bagi sekolah

a. Hasil belajar cerminan kualitas sekolah

Apabila hasil tes dari sebagian besar siswa memuaskan, maka sekolah dan guru telah berhasil. Sebaliknya sebagian besar dari siswa memperoleh hasil tes (ujian) yang belum memuaskan dan di bawah standar kompetensi kelulusan, maka sekolah dan guru dianggap kurang berhasil dalam membelajarkan siswa. Dengan demikian hasil belajar merupakan cerminan kualitas sekolah. Hal ini dapat

dicontohkan pada keberhasilan siswa dalam mengikuti ujian nasional (UN). Apabila sebagian besar siswa dari suatu sekolah dapat berhasil dalam menghadapi ujian nasional (UN) dengan presentase dan nilai yang tinggi, maka pengamat (pemerintah, orang tua, dan masyarakat) mengatakan sekolah tersebut berkualitas.

b. Membuat program sekolah

Untuk menuntaskan siswa dalam pembelajaran ujian dibuatlah suatu program oleh sekolah. Ada program pembelajaran remedial dan pengayaan. Bahkan untuk menghadapi ujian nasional diadakan program bimbingan belajar dengan menambah jam belajar di sekolah. agar siswanya lebih berhasil dalam.

c. Pemenuhan standar

Ada dua macam standar yang harus dipenuhi oleh siswa dalam pembelajaran, yakni standar isi yang disingkat dengan SI dan standar kompetensi lulusan yang disingkat dengan SKL. Dengan adanya tes yang dilaksanakan di sekolah, baik tes buatan guru maupun yang dibuat oleh pemerintah, maka hasil yang diperoleh siswa dari kedua ujian tersebut dapat diketahui apakah siswa telah memenuhi SI dan SKL.

G. Prinsip Pengukuran dan Tes

Prinsip dalam pengukuran dan tes hasil belajar yaitu:

1. Valid

Valid berarti tepat dan sesuai. Valid berhubungan dengan pemberian informasi persis (akurat) seperti keadaannya. Sesuai berhubungan dengan efektifitas alat untuk memerankan fungsinya sesuai dengan yang dimaksud dari alat ukur tersebut. Dengan demikian alat ukur (tes) hasil belajar yang disusun dan dikembangkan oleh guru dapat mengukur apa yang ingin diukur atau dapat mengukur secara jelas hasil belajar yang telah ditetapkan (sesuai dengan indikator).

2. Reliabel

Reliabel berarti stabil, konstan, tepat atau dapat dipercaya. Suatu tes akan dinyatakan reliabel apabila tes tersebut dikenakan kepada kelompok subyek yang sama, tetap memberikan hasil yang sama pula, walaupun saat pemberian tesnya berbeda. Apabila ada perbedaan hasilnya masih dalam rentangan *Standar Error Measurement (SEM)*. Tinggi rendahnya reliabilitas alat ukur dapat diketahui dengan menggunakan teknik statistik. Dengan demikian, alat ukur yang dibuat harus memiliki reliabilitas yang dapat diandalkan (tinggi).

3. Obyektif

Prinsip ini dimaksudkan, bahwa di dalam pengukuran menunjukkan aspek yang diukur sesuai dengan keadaan yang sebenarnya (obyektif), tanpa membeda-bedakan jenis kelamin, latar belakang budaya, dan ras serta berbagai hal yang memberikan kontribusi pada pembelajaran. Sebab ketidakobyektifan dalam pengukuran dapat menyebabkan menurunnya motivasi siswa dalam belajar.

4. Representatif

Alat ukur dapat mencerminkan butir-butir soal yang representatif, yang dijabarkan dari kompetensi dasar dan indikator yang telah disusun oleh guru. Untuk keperluan pembuatan butir-butir soal yang demikian, maka guru dapat membuat tabel kisi-kisi tes sebagai alat bantu yang berguna untuk menjaring butir-butir soal yang mewakili perilaku yang diharapkan.

5. Bentuk Soal Bervariasi

Ada beberapa bentuk alat ukur (tes) yang dapat digunakan untuk mengukur hasil belajar siswa. Guru diharapkan menggunakan alat ukur yang bervariasi untuk mengetahui kompetensi yang telah dimiliki. Guru tidak hanya menggunakan tes esai saja dalam mengukur kompetensi siswa, tetapi perlu divariasikan dengan tes obyektif.

6. Sesuai dengan Kegunaannya

Alat ukur harus didesain sesuai dengan kegunaannya untuk memperoleh hasil yang diinginkan. Alat ukur (tes) dapat digunakan untuk memperbaiki proses belajar mengajar (formatif), keperluan penentuan kemajuan belajar setelah menyelesaikan program pengajaran selama satu semester (sumatif), penempatan, seleksi, dan diagnostik.

7. Deskriminatif

Tes yang digunakan untuk mengukur kompetensi diharapkan dapat membedakan siswa yang mampu dan kurang mampu. Dengan demikian butir-butir soal yang disusun oleh guru mencerminkan indeks daya pembeda yang tinggi dan indeks kesukaran yang tidak sulit dan mudah.

8. Bermakna

Pengukuran dan tes yang dilakukan guru harus memberikan sumbangan positif pada pencapaian hasil belajar siswa. Oleh karena itu, pengukuran dan tes harus dinyatakan dan dapat dirasakan sebagai penghargaan untuk memotivasi siswa yang berhasil (*positive reinforcement*) dan sebagai pemicu semangat untuk meningkatkan hasil belajar bagi yang kurang berhasil (*negative reinforcement*), sehingga keberhasilan dan kegagalan siswa harus tetap diapresiasi.

G. Rangkuman

Pengukuran adalah suatu proses untuk memperoleh informasi tentang atribut atau karakteristik dari obyek, orang atau kejadian agar dapat ditentukan atau dibedakan yang bersifat kuantitatif. Ada empat skala pengukuran yaitu nominal, ordinal, interval, dan rasio. Tes adalah alat ukur berisi serangkaian pertanyaan atau latihan yang digunakan untuk mengukur keterampilan, pengetahuan, kecerdasan, kemampuan, atau bakat yang dimiliki oleh seseorang atau kelompok.

Dalam bidang pendidikan, tujuan tes dapat dibagi menjadi: (1) tes kecepatan (*speed test*), (2) tes Kemampuan (*power test*), (3) tes hasil belajar (*Achievement test*), (4) tes kemajuan belajar (*gains/achievement test*), (5), tes diagnostik (*diagnostic test*), (6) tes formatif, (7) tes sumatif, dan (8) tes penempatan. Dari bentuk tes yang digunakan dapat menjadi: (1) tes lisan, (2) tes tertulis, dan (3) tes kinerja atau perbuatan. Dari pendekatan yang digunakan, tes dibagi menjadi: (1) tes acuan norma, dan (2) tes acuan kriteria.

Pengukuran dan tes bermanfaat bagi guru, siswa, dan sekolah. Manfaat pengukuran dan tes bagi guru antara lain: (1) mengetahui tingkat penguasaan siswa, (2) mengetahui ketepatan materi yang diberikan, dan (3) mengetahui ketepatan metode yang

digunakan. Bagi siswa bermanfaat untuk (1) mengetahui kemampuannya, (2) memotivasi belajar, (3) umpan balik, dan (4) meningkatkan kebiasaan belajar. Sementara itu manfaat pengukuran dan tes bagi sekolah antara lain: (1) hasil belajar cerminan kualitas sekolah, (2) membuat program sekolah, dan (3) pemenuhan standar.

Prinsip dalam pengukuran dan tes antara lain: (1) valid, (2) reliabel, (3), obyektif, (4) representatif, (5) bentuk soal bervariasi, (6) sesuai dengan kegunaannya, (7) deskriptif, dan (8) bermakna.

H. Soal-Soal Latihan

Kerjakan dan jawablah soal di bawah ini.

1. Manakah definisi yang benar dari pengukuran?
 - A. Pengukuran mengacu pada standar tes.
 - B. Pengukuran mengacu pada tes yang dikonstruksi.
 - C. Pengukuran mengacu pada hasil tes.
 - D. Pengukuran mengacu pada prestasi
2. Di dalam mendesain suatu penilaian; manakah langkah yang mungkin terjadi dulu?
 - A. Memilih format tes yang benar.
 - B. Menentukan batasan dari setiap format tes.
 - C. Secara jelas menspesifikasikan apa yang ingin dinilai.
 - D. Mengetahui jika ada suatu standar tes untuk mengukur apa yang diinginkan.
3. Pemahaman metoda dan cara berbeda dalam menilai belajar menyiratkan pernyataan yang mana?
 - A. Beberapa jenis penilaian memiliki kesalahan dari yang lain.
 - B. Tidak ada prosedur penilaian tunggal yang dapat mengukur semua belajar.

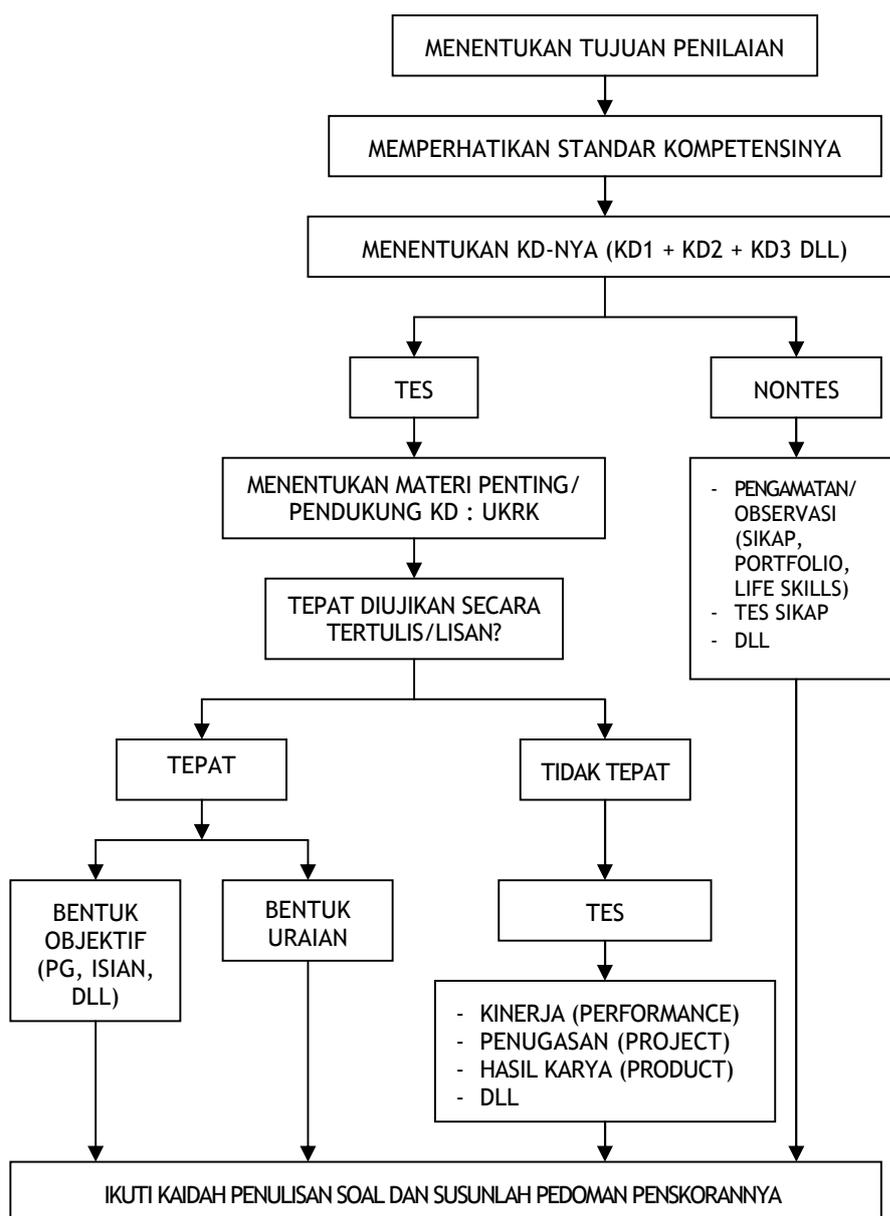
- C. Beberapa jenis penilaian menyebabkan beberapa kecemasan dari yang lain.
 - D. Guru tidak mungkin diberitahukan atas jenis penilaian tertentu.
4. Apakah yang dimaksud dengan kesalahan *sampling*?
- A. Hampir semua siswa boleh mengira.
 - B. Beberapa siswa boleh mengira lebih dari yang lain
 - C. Butir tes boleh berisi isyarat
 - D. Tes tidak dapat mengukur semua material yang diajarkan.
5. Seorang guru memutuskan untuk mengajar suatu unit materi ajar dan kemudian mentes apakah siswa telah menguasai materi yang diperkenalkan itu. Fungsi atau tujuan mana yang mungkin dilanggar oleh guru?
- A. Mengidentifikasi tujuan pembelajaran
 - B. Menilai kembali kebutuhan siswa
 - C. Menilai hasil belajar yang diharapkan
 - D. Penggunaan hasil
6. Apakah yang merupakan tujuan (*goal*) penilaian utama selama tahap " mengadakan pengajaran yang relevan"?
- A. Mengidentifikasi tujuan belajar
 - B. Meyakinkan tujuan belajar akhir telah dicapai
 - C. Memonitor materi ajar yang berkesinambungan
 - D. Memeriksa untuk melihat jika ada standard telah dipenuhi.
7. Apakah yang merupakan tujuan (*goal*) penilaian utama selama tahap " menilai hasil belajar yang diharapkan"?
- A. Mengidentifikasi tujuan belajar
 - B. Meyakinkan tujuan belajar akhir telah dicapai

- C. Memonitor material ajar yang berkesinambungan
- D. Memeriksa untuk melihat jika ada standard telah dipenuhi.
8. Dari empat tingkatan pengukuran, satu yang berisi kebanyakan informasi di dalam angka-angka adalah:
- A. interval.
- B. nominal.
- C. ordinal
- D. rasio.
9. Apa jenis tes yang merupakan contoh terbaik dari suatu penilaian secara eksternal?
- A. Acuan norma (*Norm-referenced*)
- B. Acuan kriteria (*Criterion-referenced*)
- C. Formatif
- D. Diagnostik
10. Untuk setiap situasi yang berikut, identifikasi apakah jenis pengujian yang hampir bisa dipastikan pengujian formatif (F) atau pengujian sumatif (S):
- _____ Ujian semester dalam suatu mata pelajaran bahasa Inggris.
- _____ Ujian setiap minggu dalam ejaan.
- _____ Pretes yang diberikan pada awal suatu mata pelajaran Matematika.
- _____ Tes prestasi standar yang diberikan kepada semua siswa dalam suatu sekolah pada minggu terakhir bulan Juni.
- _____ Tes yang mencakup empat tujuan sains yang diberikan pada tengah semester melalui suatu unit kurikulum.

BAB II

PENGEMBANGAN DAN PENYUSUNAN TES

Dalam pengembangan dan penyusunan tes, perlu ditetapkan terlebih dahulu tujuan penilaian dan kompetensi dasar yang hendak diukur. Adapun proses penentuannya secara lengkap dapat dilihat pada bagan berikut ini.



Keterangan: KD = Kompetensi Dasar
 KD1 + KD2 = Gabungan antar kompetensi dasar
 UKRK = Urgensi, Kontinuitas, Relevansi, Keterpakaian

Gambar 2.1 Prosedur Pengembangan Tes

Dari gambar 2.1 di atas terlihat bahwa langkah dalam pengembangan dan penyusunan tes adalah : menetapkan tujuan penilaian, SK, dan KD, menentukan alat ukur tes dan non tes, menyusun kisi-kisi tes.

A. Menentukan Tujuan Penilaian, SK, dan KD

Tujuan penilaian sangat penting karena setiap tujuan memiliki penekanan yang berbeda-beda. Misalnya untuk tujuan tes formatif, sumatif, diagnostik, dan seleksi. Tes formatif untuk mengetahui sejauh mana kemajuan belajar yang telah dicapai oleh siswa dalam suatu program pembelajaran tertentu dan untuk memperbaikinya. Tes sumatif ditujukan untuk mengetahui kompetensi siswa dalam sekumpulan materi pokok yang telah dipelajari. Tes ini dilaksanakan pada akhir program, yakni akhir semester, dan akhir tahun. Tes diagnostik ditujukan untuk mendiagnosis atau mengidentifikasi kesukaran-kesukaran dalam belajar, mendeteksi faktor-faktor yang menyebabkan terjadinya kesukaran belajar, dan menetapkan cara mengatasi kesukaran atau kesulitan belajar tersebut. Tes ini dilaksanakan untuk bimbingan belajar, pengajaran remedial (*remedial teaching*), menemukan kasus-kasus, dan lain-lain. Tes seleksi ditujukan untuk mengetahui kemampuan dan keterampilan prasyarat yang diperlukan bagi suatu program belajar dan penguasaan belajar. Untuk tujuan prestasi belajar, lingkup materi/kompetensi yang ditanyakan/diukur disesuaikan seperti untuk kuis/menanyakan materi yang lalu, pertanyaan lisan di kelas, ulangan harian, tugas individu/kelompok, ulangan semester, ulangan kenaikan kelas, laporan kerja praktik/laporan praktikum, ujian praktik. Selanjutnya Standar kompetensi (SK) dan Kompetensi Dasar (KD) merupakan acuan/target utama yang harus diukur dan dipenuhi.

B. Menentukan jenis alat ukurnya

Alat ukur yang dapat digunakan, yaitu tes atau non-tes. Untuk penggunaan tes diperlukan penentuan materi. Penentuan materi yang akan diujikan sangat penting karena di dalam satu tes tidak mungkin semua materi yang telah diajarkan dapat diujikan dalam waktu yang terbatas, misalnya satu atau dua jam. Oleh karena itu, setiap guru harus menentukan materi mana yang sangat penting dan penunjang, sehingga dalam waktu yang sangat terbatas, materi yang diujikan hanya menanyakan materi-materi yang sangat penting saja.

Penentuan materi penting dilakukan dengan memperhatikan kriteria berikut ini:

1. Urgensi, yaitu materi secara teoritis mutlak harus dikuasai oleh peserta didik,
2. Kontinuitas, yaitu materi lanjutan yang merupakan pendalaman dari satu atau lebih materi yang sudah dipelajari sebelumnya,
3. Relevansi, yaitu materi yang diperlukan untuk mempelajari atau memahami, mata pelajaran lain.
4. Keterpakaian, yaitu materi yang memiliki nilai terapan tinggi dalam kehidupan sehari-hari.

Langkah selanjutnya adalah menentukan jenis tes dengan menanyakan apakah materi tersebut tepat diujikan secara tertulis/lisan. Bila jawabannya tepat, maka materi yang bersangkutan tepat diujikan dengan bentuk soal apa, pilihan ganda atau uraian. Bila jawabannya tidak tepat, maka jenis tes yang tepat adalah tes perbuatan: kinerja (*performance*), penugasan (*project*), hasil karya (*product*), atau lainnya.

C. Aspek yang Dapat Diukur

Sebelum menyusun kisi-kisi tes, perlu ditentukan aspek (perilaku) yang dapat diukur. Dalam menentukan perilaku yang akan diukur, penulis soal dapat mengambil atau memperhatikan jenis perilaku yang telah dikembangkan oleh para ahli pendidikan, di antaranya seperti dikemukakan Benjamin S. Bloom. Ada tiga ranah yang dikembangkan oleh Bloom, yaitu ranah kognitif, afektif, dan psikomotor. Ketiga aspek dijelaskan berikut ini.

1. Kognitif

Aspek ini mencakup 6 (enam) tingkat berpikir, dimulai dari yang paling sederhana sampai dengan yang paling kompleks. Domain ini erat hubungannya dengan intelegensi, kemampuan berpikir, keterampilan memecahkan masalah. Enam tingkat berpikir tersebut adalah :

a. Pengetahuan (Knowledge)

Pengetahuan adalah aspek yang paling rendah dalam taksonomi Bloom yang didefinisikan sebagai kemampuan mengingat. Dalam jenjang kemampuan ini seseorang dituntut untuk dapat mengingat atau mengetahui adanya konsep, fakta, istilah, konvensi, kecenderungan (*trends*) dan urutan (*sequences*), klasifikasi dan kategori, kriteria, metodologi, prinsip dan generalisasi, teori, struktur, dan proses. Dalam kaitannya dengan rumusan tujuan pembelajaran, standar kompetensi,

kompetensi dasar, dan indikator menggunakan kata-kata operasional antara lain: menyebutkan, menunjukkan, mengenal, mengingat kembali, menyebutkan definisi, memilih, dan menyatakan.

b. Pemahaman (Comprehension)

Pemahaman didefinisikan sebagai kemampuan untuk memahami. Siswa dituntut untuk memahami translasi (merubah dari suatu bentuk ke bentuk yang lain), interpretasi (menjelaskan atau menyimpulkan materi), dan ekstrapolasi. Dalam kaitannya dengan rumusan tujuan pembelajaran, standar kompetensi, kompetensi dasar, dan indikator menggunakan kata-kata operasional antara lain : menjelaskan, membedakan, mencontohkan, meringkaskan, mendiskusikan, menceritakan, dan menggambarkan.

c. Aplikasi (Application)

Aplikasi didefinisikan sebagai kemampuan untuk menerapkan. Siswa dituntut menerapkan ide-ide umum, tata cara, metode-metode, prinsip-prinsip, serta teori-teori dalam situasi baru dan konkret. Apabila situasinya tidak baru, maka kemampuan yang diukur bukan lagi penerapan, tetapi ingatan. Suatu soal yang telah dipakai sebagai contoh di kelas mengenai penerapan suatu rumus, jangan lagi dipakai dalam tes. Jika soal yang sama itu disajikan, maka siswa menjawab hanya berdasarkan ingatan, bukan melalui penerapan. Harus diciptakan butir soal baru yang serupa tetapi tidak sama.

Pengukuran kemampuan ini umumnya menggunakan pendekatan pemecahan masalah (*problem solving*). Melalui pendekatan ini siswa dihadapkan dengan suatu masalah yang perlu dipecahkan atau diselesaikan dengan menggunakan pengetahuan yang telah dimilikinya. Dengan demikian, penguasaan aspek ini sudah tentu harus didasari aspek pemahaman yang mendalam tentang segala sesuatu yang berhubungan dengan masalah tersebut.

Dalam kaitannya dengan rumusan tujuan pembelajaran, standar kompetensi, kompetensi dasar, dan indikator menggunakan kata-kata operasional antara lain : menghitung, mendemonstrasikan, menerapkan, menoperasionalkan, memanipulasikan, memodifikasikan, menghubungkan, memecahkan masalah, mengilustrasikan, dan mempraktekkan.

d. Analisis (Analysis)

Analisis didefinisikan sebagai kemampuan menganalisis suatu informasi yang luas menjadi bagian-bagian kecil. Dengan demikian, situasi atau informasi tersebut menjadi lebih jelas. Misalnya; menggunakan suatu informasi/pengetahuan yang diperolehnya untuk memecahkan masalah. Kemampuan analisis diklasifikasikan atas tiga kelompok, yaitu : (1) analisis unsur, (2) analisis hubungan, dan (3) analisis prinsip-prinsip yang terorganisasi. Dalam analisis unsur diperlukan kemampuan merumuskan asumsi-asumsi dan mengidentifikasi unsur-unsur penting dan dapat membedakan antara fakta dan nilai. Dalam analisis hubungan menuntut kemampuan mengenal unsur-unsur dan pola hubungannya. Dalam analisis prinsip-prinsip yang terorganisasi menuntut kemampuan menganalisis pokok-pokok yang melandasi suatu organisasi, misalnya menentukan falsafah pengarang dari isi buku yang ditulisnya.

Dalam kaitannya dengan rumusan tujuan pembelajaran, standar kompetensi, kompetensi dasar, dan indikator menggunakan kata-kata operasional antara lain: membedakan, mendeteksi, mengidentifikasi, mengklasifikasikan, mendiskriminasikan, menyatakan, mengkategorikan, dan membandingkan.

e. Sintesis (Synthesis)

Sintesis didefinisikan sebagai kemampuan menggabungkan beberapa informasi menjadi suatu kesimpulan. Misalnya; memformulasikan hasil penelitian di laboratorium. Dengan demikian kemampuan yang diharapkan lebih menekankan kepada perilaku yang kreatif dengan tekanan utama pada memformulasikan pola-pola baru atau struktur yang baru.

Dalam kaitannya dengan rumusan tujuan pembelajaran, standar kompetensi, kompetensi dasar, dan indikator menggunakan kata-kata operasional seperti: mengembangkan, mengkombinasikan, mensintesis, mengorganisasikan, mengklasifikasikan, mem-formulasikan, dan memodifikasikan.

f. Evaluasi (Evaluation)

Evaluasi didefinisikan sebagai kemampuan mempertimbangkan mana yang baik dan mana yang buruk dan memutuskan untuk mengambil tindakan tertentu. Dalam kaitannya dengan rumusan tujuan pembelajaran, standar kompetensi, kompetensi dasar, dan indikator menggunakan kata-kata operasional

antara lain : memutuskan, mempertimbangkan, menyimpulkan, membandingkan, menstandarisasikan, dan merevisi.

2. Afektif

Domain kognitif dipergunakan dalam pengukuran penampilan maksimal (*what a person can do*), sedangkan dalam domain afektif dipergunakan dalam pengukuran penampilan khusus (*what a person does, do, or feel*). Penilaian afektif mencakup antara lain: sikap, tingkah laku, minat, emosi, motivasi, kerjasama, dan koordinasi dari setiap siswa. Penilaian afektif dilakukan melalui pengamatan dan interaksi langsung secara menerus. Pada umumnya dilakukan secara non-ujian (misalnya; untuk mengetahui siapa siswa yang bisa dipercaya, siapa siswa yang disiplin, siapa siswa yang rajin, siapa siswa yang berminat ke jurusan Ilmu Sosial atau Ilmu Alam, dan lain-lain. Setiap informasi yang diperoleh dikumpulkan dan disimpan sebagai referensi dalam penilaian berikutnya. Penilaian afektif dibagi atas penilaian afektif secara umum (budi pekerti) dan penilaian afektif per mata pelajaran.

Aspek penilaian afektif terdiri dari: (1) menerima (*receiving*), (2) menanggapi (*responding*), (3) menilai (*evaluating*), (4) mengorganisasi (*organization*), dan (5) membentuk watak (*characterization*). Pada aspek menerima termasuk kesadaran, keinginan untuk menerima stimulus, respon, kontrol, dan seleksi gejala atau rangsangan dari luar. Dalam kaitannya dengan rumusan tujuan pembelajaran, standar kompetensi, kompetensi dasar, dan indikator menggunakan kata-kata operasional antara lain: memilih, menguraikan, mengikuti, memberikan, memegang (*hold*), menempatkan, menamakan (*name*), menunjukkan, dan memakai. Pada aspek menanggapi termasuk reaksi yang diberikan: ketepatan reaksi, perasaan kepuasan, dan lain-lain. Dalam kaitannya dengan rumusan tujuan pembelajaran, kompetensi dasar, dan indikator menggunakan kata-kata operasional antara lain : menjawab, membantu, memenuhi (*comply*), menyesuaikan (*conform*), mendiskusikan, membantu, mempraktekkan, menampilkan, menceritakan, melaporkan, merespon, memilih, and menuliskan.

Pada aspek menilai termasuk kesadaran menerima norma, sistem nilai, dan lain-lain. Dalam kaitannya dengan rumusan tujuan pembelajaran, kompetensi dasar, dan indikator menggunakan kata-kata operasional antara lain : melengkapi, mendemonstrasikan, menguraikan, membedakan, menjelaskan, mengikuti, membentuk (*form*), memprakarsai, mengundang (*invite*), menggabungkan,

menjustifikasi, mengajukan, menyatakan, melaporkan, memilih, membagi (*share*), mempelajari, mengerjakan, dan menulis.

Pada aspek mengorganisasi termasuk pengembangan norma dan nilai dalam organisasi, sistem nilai, dan lain-lain. Dalam kaitannya dengan rumusan tujuan pembelajaran, kompetensi dasar, dan indikator menggunakan kata-kata operasional antara lain: menempelkan (*adhere*), merubah (*alter*), mengatur, mengkombinasikan, membandingkan, menyatukan (*integrate*), melengkapi, mempertahankan, menjelaskan, menggeneralisasikan, mengidentifikasi, , memodifikasi, mengorganisir, memerintah, menyiapkan, menghubungkan, dan mensintesis.

Pada aspek membentuk watak termasuk sistem nilai yang terbentuk mempengaruhi pola kepribadian dan tingkah laku. Dalam kaitannya dengan rumusan tujuan pembelajaran, kompetensi dasar, dan indikator menggunakan kata-kata operasional antara lain: menampilkan, mendiskriminasikan, mempengaruhi, mendengarkan, memodifikasi, mempraktekkan, mengajukan, menanyakan, merevisi, melayani, menyelesaikan, memakai, dan membuktikan (*verify*).

3. Psikomotor dan (Life Skill)

Tidak semua mata pelajaran dapat dinilai aspek psikomotornya (d disesuaikan dengan tuntutan kompetensi dasar yang harus dicapai siswa). Aspek psikomotor digunakan untuk pembelajaran yang banyak memerlukan praktek seperti : Pendidikan Agama, Pendidikan Seni, Pendidikan Jasmani, Praktik IPA, praktik Bahasa, dan praktik di bengkel atau workshop.

Aspek psikomotor terdiri dari: (1) meniru (*perception*), (2) menyusun (*manipulating*), (3) melakukan dengan prosedur (*precision*), (4) melakukan dengan baik dan tepat (*articulation*), dan (5) melakukan tindakan secara alami (*naturalization*). Pada tingkat persepsi (peniruan) berkenaan dengan perasaan untuk menirukan gerakan yang ditunjukkan. Ada tiga tingkatan untuk aspek ini, yaitu: (a) kesadaran/adanya rangsangan (*awarness of a stimulus*), (b) seleksi terhadap gerakan-gerakan yang relevan (*cue selection*), dan (c) melaksanakan dalam suatu penampilan (*to action in a performance*).

Dalam kaitannya dengan rumusan tujuan pembelajaran, standar kompetensi, kompetensi dasar, dan indikator menggunakan kata-kata operasional antara lain : pilihlah, bedakan, identifikasikan, hubungkan, pisahkan, dan sebagainya. Pada tingkat manipulasi berkenaan dengan kesiapan untuk melaksanakan suatu tindakan seperti yang diajarkan, dalam arti memilih yang diperlukan. Ada tiga jenis kesiapan

yaitu : (a) mental set (*mental readiness to act*), (b) physical set (*physical readiness to act*), dan (c) emosional set (*willingness to act*). Dalam kaitannya dengan rumusan tujuan pembelajaran, kompetensi dasar, dan indikator menggunakan kata-kata operasional antara lain: menampilkan, menunjukkan, membuat reaksi, mengawali, dan sebagainya.

Penilaian terhadap proses dan hasil belajar siswa di dalam kurikulum KTSP mencakup aspek-aspek kecakapan hidup (*life skill*). Aspek kecakapan hidup meliputi: (1) kesadaran diri, (2) kecakapan berpikir dan bernalar, (3) kecakapan akademik, (4) kecakapan vokasional, (5) kecakapan komunikasi, dan (6) kecakapan bekerja sama. Kesadaran diri terbagi atas: (a) kesadaran diri sebagai hamba Tuhan, makhluk social, dan makhluk lingkungan, dan (b) kesadaran akan potensi diri dan dorongan untuk mengembangkannya. Kecakapan berpikir dan bernalar terbagi atas: (a) kecakapan menggali informasi, (b) kecakapan mengolah informasi dan mengambil keputusan dengan cerdas, dan (c) kecakapan memecahkan masalah secara arif dan kreatif. Kecakapan akademik meliputi: (a) kecakapan merumuskan hipotesis, dan (b) kecakapan dalam merancang dan melaksanakan penelitian. Kecakapan vokasional meliputi: (a) vokasional dasar, dan (b) vokasional khusus. Kecakapan komunikasi meliputi: (a) kecakapan mendengarkan, (b) kecakapan berbicara, (c) kecakapan membaca, dan (d) kecakapan menuliskan pendapat atau gagasan. Kecakapan bekerja sama meliputi: (a) kecakapan sebagai teman kerja yang menyenangkan, dan (b) kecakapan sebagai pimpinan yang berempati.

4. Revisi taksonomi Bloom

Ada revisi taksonomi Bloom pada ranah kognitif yang dilakukan oleh Anderson *et.al* (2001). Pada ranah kognitif sebelumnya terdiri dari: (1) pengetahuan, (2) pemahaman, (3) aplikasi, (4) analisis, (5) sintesis, dan (6) evaluasi, akan tetapi setelah diadakan revisi, maka ranah kognitif dari Bloom terdiri dari: (1) *remember* (mengingat), (2) *understanding* (pemahaman), (3) *apply* (aplikasi), (4) *analyze* (analisa), (5) *evaluate* (evaluasi), dan (6) *create* (kreasi). Pada aspek *remember* terbagi atas: (a) *recognizing*, dan (b) *recalling*. Pada aspek *understanding* terbagi atas : (a) *interpreting*, (b) *exemplifying*, (c) *classifying*, (d) *summarizing*, (e) *inferring*, (f) *comparing*, (g) *explaining*. Pada aspek aplikasi terbagi atas : (a) *executing*, dan (b) *implementing*. Pada aspek analisis terbagi atas: (a) *difrentiating*, (b) *organizing*, dan (c) *attributing*. Pada aspek evaluasi terbagi atas: (a) *checking*, dan (b) *critiquing*. Pada aspek create terbagi atas : (a) *generating*, (b) *planning*, dan (c) *producing*.

5. Penentuan Perilaku yang Akan Diukur

Perilaku yang akan diukur tergantung pada tuntutan kompetensi, baik standar kompetensi maupun kompetensi dasarnya. Setiap kompetensi di dalam kurikulum memiliki tingkat keluasan dan kedalaman kemampuan yang berbeda. Semakin tinggi kemampuan/perilaku yang diukur sesuai dengan target kompetensi, maka semakin sulit soal dan semakin sulit pula menyusunnya.

D. Penyusunan Soal dan Kisi-Kisi Tes

1. Penentuan dan Penyebaran Soal

Sebelum menyusun kisi-kisi dan butir soal perlu ditentukan jumlah soal setiap kompetensi dasar dan penyebaran soalnya. Untuk lebih jelasnya, perhatikan contoh penilaian akhir semester berikut ini.

Tabel 2.1 Contoh Penyebaran Butir Soal

No	Kompetensi Dasar	Materi	Jumlah soal tes tulis		Jumlah soal Praktik
			PG	Uraian	
1	1.1	6	--	--
2	1.2	3	1	--
3	1.3	4	--	1
4	2.1	5	1	--
5	2.2	8	1	--
6	3.1	6	--	1
7	3.2	--	2	--
8	3.3	8	--	--
Jumlah soal			40	5	2

Soal yang bermutu adalah bahwa soal harus sah (valid), dan handal. Linn dan Gronlund (1995: 47) menyatakan bahwa tes yang baik harus memenuhi tiga karakteristik, yaitu: validitas, reliabilitas, dan usabilitas. Validitas artinya ketepatan interpretasi hasil prosedur pengukuran, reliabilitas artinya konsistensi hasil pengukuran, dan usabilitas artinya praktis prosedurnya.

Untuk dapat menghasilkan soal yang sah dan handal, penulis soal harus

merumuskan kisi-kisi dan menulis soal berdasarkan kaidah penulisan soal yang baik (kaidah penulisan soal bentuk objektif/pilihan ganda, uraian, atau praktik).

2. Penyusunan Kisi-kisi

Kisi-kisi (*test blue-print* atau *table of specification*) merupakan deskripsi kompetensi dan materi yang akan diujikan. Tujuan penyusunan kisi-kisi adalah untuk menentukan ruang lingkup dan sebagai petunjuk dalam menulis soal. Kisi-kisi dapat berbentuk format atau matriks seperti contoh berikut ini.

FORMAT KISI-KISI PENULISAN SOAL

Jenis sekolah :

Jumlah soal :

Mata pelajaran :

Bentuk soal/tes :

Kurikulum :

Penyusun : 1.
2.

Alokasi waktu :

Standar Kompetensi :

No.	Kompetensi Dasar	Kls/smt	Materi Pokok	Indikator Soal	Aspek yang Diuji						Jumlah Soal
					C 1	C 2	C 3	C 4	C 5	C 6	

Keterangan:

Isi pada kolom 2, 3, 4, dan 5 adalah harus sesuai dengan pernyataan yang ada di dalam silabus/kurikulum. Untuk Aspek kognitif yang diuji C1 (pengetahuan / *knowledge*), C2 (pemahaman / *comprehension*), C3 (aplikasi / *application*), C 4 (analisis / *analysis*), C5 (sintesis / *synthesis*), dan C6 (evaluasi / *evaluation*).

Kisi-kisi yang baik harus memenuhi persyaratan berikut ini:

1. Kisi-kisi harus dapat mewakili isi silabus/kurikulum atau materi yang telah diajarkan secara tepat dan proporsional.
2. Komponen-komponennya diuraikan secara jelas dan mudah dipahami.
3. Materi yang hendak ditanyakan dapat dibuatkan soalnya.

3. Perumusan Indikator Soal

Indikator dalam kisi-kisi merupakan pedoman dalam merumuskan soal yang dikehendaki. Kegiatan perumusan indikator soal merupakan bagian dari kegiatan penyusunan kisi-kisi. Untuk merumuskan indikator dengan tepat, dosen harus memperhatikan materi yang akan diujikan, indikator pembelajaran, kompetensi dasar, dan standar kompetensi. Indikator yang baik dirumuskan secara singkat dan jelas. Syarat indikator yang baik:

1. menggunakan kata kerja operasional (perilaku khusus) yang tepat,
2. menggunakan satu kata kerja operasional untuk soal objektif, dan satu atau lebih kata kerja operasional untuk soal uraian/tes perbuatan,
3. dapat dibuatkan soal atau pengecohnya (untuk soal pilihan ganda).

Penulisan indikator yang lengkap mencakup A = *audience* (peserta didik), B = *behaviour* (perilaku yang harus ditampilkan), C = *condition* (kondisi yang diberikan), dan D = *degree* (tingkatan yang diharapkan). Ada dua model penulisan indikator. Model pertama adalah menempatkan kondisinya di awal kalimat. Model pertama ini digunakan untuk soal yang disertai dengan dasar pernyataan (stimulus), misalnya berupa sebuah kalimat, paragraf, gambar, denah, grafik, kasus, atau lainnya, sedangkan model yang kedua adalah menempatkan peserta didik dan perilaku yang harus ditampilkan di awal kalimat.

Contoh: soal menyimak pada mata pelajaran Bahasa Indonesia.

Indikator: Diperdengarkan sebuah pernyataan pendek dengan topik "belajar mandiri", peserta didik dapat menentukan dengan tepat pernyataan yang sama artinya.

Model yang kedua ini digunakan untuk soal yang tidak disertai dengan dasar pertanyaan (stimulus). Contoh: soal penulisan nilai uang pada mata pelajaran Ekonomi Indikator: Peserta didik dapat menentukan dengan tepat penulisan tanda baca pada nilai uang.

4. Langkah-langkah Penyusunan Butir Soal

Agar soal yang disiapkan oleh setiap guru menghasilkan bahan ulangan/ujian yang sah dan handal, maka harus dilakukan langkah-langkah berikut, yaitu: (1) menentukan tujuan tes, (2) menentukan kompetensi yang akan diujikan, (3) menentukan materi yang diujikan, (4) menetapkan penyebaran butir soal berdasarkan kompetensi, materi, dan bentuk penilaiannya (tes tertulis: bentuk pilihan ganda, uraian; dan tes praktik), (5) menyusun kisi-kisinya, (6) menulis butir soal, (7) memvalidasi butir soal atau menelaah secara kualitatif, (8) merakit soal menjadi perangkat tes, (9) menyusun pedoman penskorannya (10) uji coba butir soal, (11) analisis butir soal secara kuantitatif dari data empirik hasil uji coba, dan (12) perbaiki soal berdasarkan hasil analisis.

5. Penyusunan Butir Soal Tes Tertulis

Penulisan butir soal tes tertulis merupakan suatu kegiatan yang sangat penting dalam penyiapan bahan ulangan/ujian. Setiap butir soal yang ditulis harus berdasarkan rumusan indikator soal yang sudah disusun dalam kisi-kisi dan berdasarkan kaidah penulisan soal bentuk obyektif dan kaidah penulisan soal uraian.

Penggunaan bentuk soal yang tepat dalam tes tertulis, sangat tergantung pada perilaku/kompetensi yang akan diukur. Ada kompetensi yang lebih tepat diukur/ditanyakan dengan menggunakan tes tertulis dengan bentuk soal uraian, ada pula kompetensi yang lebih tepat diukur dengan menggunakan tes tertulis dengan bentuk soal objektif. Bentuk tes tertulis pilihan ganda maupun uraian memiliki kelebihan dan kelemahan satu sama lain.

6. Penulisan Soal Bentuk Uraian

Menulis soal bentuk uraian diperlukan ketepatan dan kelengkapan dalam merumuskannya. Ketepatan yang dimaksud adalah bahwa materi yang ditanyakan tepat diujikan dengan bentuk uraian, yaitu menuntut peserta didik untuk mengorganisasikan gagasan dengan cara mengemukakan atau mengekspresikan gagasan secara tertulis dengan menggunakan kata-katanya sendiri. Adapun kelengkapan yang dimaksud adalah kelengkapan perilaku yang diukur yang digunakan untuk menetapkan aspek yang dinilai dalam pedoman penskorannya.

Pembahasan tes esai akan diuraikan lebih lanjut dalam bab berikutnya.

7. Penulisan Soal Bentuk Pilihan Ganda

Menulis soal bentuk pilihan ganda sangat diperlukan keterampilan dan ketelitian. Hal yang paling sulit dilakukan dalam menulis soal bentuk pilihan ganda adalah menuliskan pengecohnya. Soalnya mencakup: (1) dasar pertanyaan/stimulus (bila ada), (2) pokok soal (*stem*), (3) pilihan jawaban yang terdiri atas: kunci jawaban dan pengecoh.

Pembahasan tes obyektif, termasuk tes pilihan ganda (*multiple choice*) akan diuraikan lebih lanjut dalam bab berikutnya.

E. Perakitan Butir Soal

1. Pengertian

Merakit soal adalah menyusun soal yang siap pakai menjadi satu perangkat/paket tes atau beberapa paket tes paralel. Dasar acuan dalam merakit soal adalah tujuan tes dan kisi-kisinya. Untuk memudahkan pelaksanaannya, guru harus memperhatikan langkah-langkah perakitan soal.

2. Langkah-langkah Perakitan Soal

Para pendidik dapat merakit soal menjadi suatu paket tes yang tepat, apabila para pendidik memperhatikan langkah-langkah perakitan soal. Berikut langkah-langkah perakitan soal.

- a. Mengelompokkan soal-soal yang mengukur kompetensi dan materi yang sama, kemudian soal-soal itu ditempatkan dalam urutan yang sama.
- b. Memberi nomor urut soal didasarkan nomor urut soal dalam kisi-kisi.
- c. Mengecek setiap soal dalam satu paket tes apakah soal-soalnya sudah bebas dari kaidah "Setiap soal tidak boleh memberi petunjuk jawaban terhadap soal yang lain".
- d. Membuat petunjuk umum dan khusus untuk mengerjakan soal.
- e. Membuat format lembar jawaban.
- f. Membuat lembar kunci jawaban dan petunjuk penilaiannya.
- g. Menentukan penyebaran kunci jawaban (untuk bentuk pilihan ganda).
- h. Menentukan soal inti sebanyak 10% dari jumlah soal dalam satu paket. Soal inti ini diperlukan apabila soal yang dirakit terdiri dari beberapa tes paralel. Tujuannya adalah agar antar tes memiliki keterkaitan yang sama. Penempatan soal inti dalam paket tes diletakkan secara acak.
- i. Menentukan besarnya bobot setiap soal (untuk soal bentuk uraian)

- j. Bobot soal adalah besarnya angka yang ditetapkan untuk suatu butir soal dalam perbandingan (*ratio*) dengan butir soal lainnya dalam satu perangkat tes. Penentuan besar kecilnya bobot soal didasarkan atas tingkat kedalaman dan keluasan materi yang ditanyakan atau kompleksitas jawaban yang dituntut oleh suatu soal.
- k. Menyusun tabel konversi skor

Tabel konversi sangat membantu para pendidik pada saat menilai lembar jawaban peserta didik. Terutama bila dalam satu tes terdiri dari dua bentuk soal, misal bentuk pilihan ganda dan uraian atau tes tertulis dan tes praktik. Skor dari soal bentuk pilihan ganda tidak dapat langsung digabung dengan skor uraian. Hal ini karena tingkat keluasan dan kedalaman materi yang ditanyakan atau penekannya dalam kedua bentuk itu tidak sama. Nilai keduanya dapat digabung setelah keduanya ditentukan bobotnya. Misalnya, untuk soal bentuk pilihan ganda (45 soal dengan skor maksimum 45) bobotnya 60% dan bentuk uraian (5 soal dengan skor maksimum 20) bobotnya 40%. Untuk menentukan skor jadinya adalah skor perolehan peserta didik yang bersangkutan dibagi skor maksimum kali bobot.

Konversi biasa (model pengukuran secara klasik), yaitu untuk memperoleh nilai murni peserta didik. Bila menghendaki skor maksimum 10 digunakan rumus (skor perolehan: skor maksimum) x 10 dan bila menggunakan skor maksimum 100 digunakan nilai konversi dengan rumus (skor perolehan: skor maksimum) x 100 atau bila menggunakan skor maksimum 4 digunakan nilai konversi dengan rumus (skor perolehan : skor maksimum) x 4.

Konversi seperti ini memiliki dua kelemahan, pertama adalah bahwa setiap butir soal dihitung memiliki tingkat kesukaran yang sama. Artinya peserta didik manapun yang menjawab benar 40 dari 50 butir soal dalam satu tes (terserah nomor butir soal berapa yang benar, apakah nomor 1 benar, nomor 2 salah, nomor 3 benar atau sebaliknya dan seterusnya, yang penting benar 40 soal) peserta didik yang bersangkutan akan memperoleh nilai 8 (untuk konversi skor maksimum 10), 80 (untuk konversi skor maksimum 100) 0,2 (untuk konversi skor maksimum 4). Kelemahan kedua adalah bahwa tingkat kesukaran butir soal tidak ditempatkan/dikalibrasi pada skala yang sama. Artinya bahwa butir-butir soal tidak disusun berdasarkan tingkat kesukarannya dan kemampuan peserta didik sehingga model konversi ini

belum bisa menentukan nilai murni peserta didik yang sebenarnya. Seharusnya hanya peserta didik yang memiliki kemampuan tinggi (misal pada skala kemampuan 1, kemampuan 2, kemampuan 3) yang dapat menjawab benar semua soal dalam tes pada skala yang bersangkutan atau tingkat kesukaran butir (mudah, sedang, sukar) sesuai dengan kemampuan peserta didik yang bersangkutan. Apabila sekolah mempergunakan konversi biasa seperti ini justru akan merugikan peserta didik yang memiliki kemampuan lebih tinggi.

F. Rangkuman

Untuk menentukan teknik dan alat penilaian, perlu ditetapkan terlebih dahulu tujuan penilaian dan kompetensi dasar yang hendak diukur. Setelah menentukan kompetensi yang akan diukur, maka langkah berikutnya adalah menentukan materi yang akan diujikan. Penentuan materi yang akan diujikan sangat penting karena di dalam satu tes tidak mungkin semua materi yang telah diajarkan dapat diujikan dalam waktu yang terbatas. Penentuan materi penting dilakukan dengan memperhatikan kriteria: (1) Urgensi, (2) Kontinuitas, dan (3) Relevansi.

Dalam menentukan perilaku yang akan diukur, ada tiga ranah yang dikembangkan oleh Bloom (1956) dan menjadi bahan pertimbangan, yaitu ranah kognitif, afektif, dan psikomotor. Ranah kognitif mencakup pengetahuan, pemahaman, aplikasi, analisis, sintesis, dan evaluasi. Ranah afektif mencakup menerima (*receiving*), menanggapi (*responding*), menilai (*evaluating*), mengorganisasi (*organization*), dan membentuk watak (*characteristization*). Ranah psikomotor meliputi: meniru (*perception*), menyusun (*manipulating*), melakukan dengan prosedur (*precision*), melakukan dengan baik dan tepat (*articulation*), dan melakukan tindakan secara alami (*naturalization*). Taksonomi Bloom ada yang telah direvisi (2001) pada ranah kognitif. Hasil revisi ranah kognitif oleh Anderson *et.al* meliputi: (1) *remember* (mengingat), (2) *understanding* (pemahaman), (3) *apply* (aplikasi), (4) *analyze* (analisa), (5) *evaluate* (evaluasi), dan (6) *create* (kreasi).

Butir soal perlu ditentukan jumlah soal setiap kompetensi dasar dan penyebaran soalnya. Soal yang baik adalah soal yang sah (*valid*), dan handal. Untuk dapat menghasilkan soal yang sah dan handal, penulis soal harus merumuskan kisi-kisi dan menulis soal berdasarkan kaidah penulisan soal yang baik. Kisi-kisi dapat berbentuk format atau matriks. Indikator dalam kisi-kisi

merupakan pedoman dalam merumuskan soal yang dikehendaki. Penulisan indikator yang lengkap mencakup A = *audience* (peserta didik) , B = *behaviour* (perilaku yang harus ditampilkan), C = *condition* (kondisi yang diberikan), dan D = *degree* (tingkatan yang diharapkan). Setelah butir soal disusun pada kisi-kisi dilanjutkan dengan menulis butir soal, memvalidasi butir soal atau menelaah secara kualitatif, merakit soal menjadi perangkat tes, menyusun pedoman penskorannya uji coba butir soal, analisis butir soal secara kuantitatif dari data empirik hasil uji coba, dan perbaikan soal berdasarkan hasil analisis.

G. Soal

Kerjakan atau diskusikan soal di bawah ini sebagai latihan!

1. Taksonomi Bloom untuk tujuan pembelajaran dalam ranah kognitif secara hirarki didasarkan pada:
 - A. keluasan daya ingat yang dibutuhkan untuk mencapai suatu tujuan
 - B. level pemahaman yang dibutuhkan untuk mencapai suatu tujuan
 - C. level pembacaan (*reading level*) yang dibutuhkan untuk mencapai suatu tujuan
 - D. sikap yang dibutuhkan untuk mencapai suatu tujuan
2. Tes dibuat untuk berbagai tujuan. Manakah berikut ini yang bukan tujuan umum dari tes
 - A. afirmasi
 - B. prediksi
 - C. evaluasi
 - D. deskripsi
3. Butir tes harus dievaluasi dalam arti seberapa baik mereka memenuhi:
 - A. tujuan.
 - B. level pembacaan
 - C. keandalan
 - D. penilaian
4. Menjelaskan dalam satu kata-kata sendiri yang suatu kompromi memerlukan suatu jalur tentang apa level Taksonomi Bloom's?
 - A. aplikasi
 - B. pengetahuan
 - C. pemahaman
 - D. analisis

5. Suatu tugas mendisain suatu solusi yang bisa diterima untuk pembuangan limbah nuklir termasuk level taksonomi:
 - A. sintesis
 - B. pengetahuan
 - C. pemahaman
 - D. analisis

6. Menjodohkan nama-nama pengarang dan judul dari buku adalah suatu tugas termasuk level taksonomi:
 - A. aplikasi
 - B. pengetahuan
 - C. pemahaman
 - D. analisis

7. Sebuah tabel spesifikasi butir tes merupakan cerminan dari:
 - A. isi dan level pembacaan
 - B. proses teori dan isi
 - C. proses kognitif dan level pembacaan
 - D. jenis butir dan proses kognitif

8. Sebuah tabel spesifikasi butir tes merupakan cerminan dari:
 - A. standard untuk mengevaluasi suatu tes
 - B. mengubah skor tes untuk evaluasi
 - C. mendaftar tujuan pembelajaran
 - D. merencanakan suatu tes

9. Mana berikut ini yang bukan bagian dari suatu tujuan pembelajaran yang ditulis baik menurut Mager?
 - A. suatu deskripsi dari pelajar
 - B. perilaku yang akan diamati
 - C. kondisi-kondisi di mana perilaku akan terjadi
 - D. ukuran kinerja yang dapat diterima

10. Mana berikut ini tujuan pembelajaran yang bukan dinyatakan dalam istilah perilaku?
 - A. membaca

- B. memahami
- C. mendaftar
- D. menghitung

BAB III TES OBYEKTIF

Tes obyektif yang dibahas di bawah ini meliputi tes benar-salah (*true-false test*), tes menjodohkan (*matching test*), dan tes pilihan ganda (*multiple choice test*).

A. Tes Benar – Salah

1. Pengertian

Tes benar-salah adalah tes berbentuk kalimat berita atau pertanyaan yang mengandung dua kemungkinan yaitu benar atau salah. Siswa diminta untuk mengemukakan pendapatnya mengenai pertanyaan-pertanyaan yang menjadi isi dari setiap soal.

Macam-macam bentuk pertanyaan benar salah:

- a. Tes benar-salah bentuk pertanyaan
Dalam bentuk ini soal terdiri dari pertanyaan-pertanyaan dan siswa diminta memilih kemungkinan betul atau salah saja.
- b. Tes benar- salah yang menuntut alasan
Dalam bentuk ini selain seperti bentuk pertanyaan saja juga menuntut siswa memberikan alasan apabila ia memilih kemungkinan salah (menyalahkan pertanyaan soal).
- c. Tes benar - salah dengan membetulkan
Dalam bentuk ini selain seperti bentuk pertanyaan juga menuntut supaya siswa membetulkan pertanyaan soal yang disalahkan, jika siswa memilih kemungkinan terhadap pertanyaan/soal yang bersangkutan).
- d. Tes benar-salah berganda
Dalam bentuk ini satu induk persoalan menghasilkan beberapa anak persoalan yang dirumuskan dalam pertanyaan/soal yang mempunyai kemungkinan benar atau salah.

2. Kelebihan dan Kelemahan Tes Benar-Salah

Kelebihan dari tes benar-salah adalah:

- a. Mudah dan cepat dalam menilai
- b. Waktu pengerjaannya cepat

- c. Penilaiannya objektif
- d. Menyusun soalnya lebih mudah daripada tes pilihan berganda
- e. Mencakup bahan yang luas dan tidak memakan tempat karena biasanya pertanyaan-pertanyaan singkat saja
- f. Dapat dipakai berulang-ulang.

Kelemahan dari tes benar-salah

- a. Lama penyusunan soalnya dibanding dengan tes esai
- b. Kemungkinan mengira-ngira jawaban besar
- c. Menyusun pertanyaan (soal) supaya soal benar atau salah saja adalah sulit
- d. Kurang dapat membedakan siswa pandai dengan siswa yang tidak pandai
- e. Reabilitasnya rendah
- f. Dapat membingungkan siswa
- g. Hanya dapat mengungkapkan ingatan dan pengenalan kembali
- h. Banyak hal persoalan yang tidak dapat dinyatakan hanya dua kemungkinan (benar atau salah) saja.

3. Cara Menyusun Soal

Syarat-syarat menyusun soal:

- a. Soal harus singkat, jelas bukan kalimat majemuk
- b. Jumlah soal harus banyak dan disusun berdasarkan tabel spesifikasi
- c. Satu soal harus berisi satu persoalan
- d. Tidak memakai kata seperti, selalu, seringkali, pada umumnya, biasanya, karena kata-kata seperti itu memudahkan siswa menerka jawabannya
- e. Setiap pertanyaan/ soal harus pasti salah atau betul (tidak mendua arti)
- f. Jumlah soal yang betul dan salah harus seimbang
- g. Urutan soal yang betul dan salah harusnya tidak mengikuti pola yang teratur
- h. Sebaiknya pertanyaan tidak diambil langsung dari buku
- i. Tulisan huruf B - S pada permulaan nomor pada masing- masing item dengan maksud mempermudah pengerjaannya dan menilai (*scoring*).

4. Contoh Soal

- a. Soal Benar - salah bentuk pertanyaan

Perintah: lingkarihuruf B jika pertanyaan di bawah ini benar dan huruf S jika pertanyaan itu salah!

1. B–S Tanah air kita terletak di khatulistiwa
2. B–S Menurut penelitian terakhir jumlah pulau-pulau di Indonesia ada 13,677 buah.

- b. Soal benar - salah menuntut alasan

Lingkarihuruf B jika pernyataan di bawah ini betul, dan lingkarihuruf S jika salah dan beri alasan mengapa anda anggap salah. Tulis alasan anda di tempat yang telah disediakan.

- 1) B–S Udara di kota kena pencemaran
- 2) B–S Pengamalan tata perekonomian Pancasila bertujuan untuk kebebasan berusaha
- 3) B–S P4 merupakan kepribadian / jiwa bangsa Indonesia

- c. Tes benar - salah dengan membetulkan

Lingkarihuruf B jika pernyataan di bawah ini betul, dan lingkarihuruf S dan tunjukkan bagian mana yang salah dengan cara memberi garis di bawahnya serta kemudian tulislah pembedulannya pada tempat yang telah disediakan.

- 1) B–S Tes esai dapat dipakai berulang kali
- 2) B–S Prinsip diskriminasi dalam evaluasi sesuai pendekatan dalam sistem pembelajaran

- d. Tes benar - salah berganda

Lingkarihuruf B jika pernyataan di bawah ini betul, dan lingkarihuruf S jika salah.

Pertanyaan:

1. Volume suatu gas
 - a. Bertambah besar jika temperatur dinaikkan B – S
 - b. Bertambah besar bila tekanan diperbesar B – S
 - c. Jika diperkecil menjadi nol dengan mempertinggi tekanan dan mempertinggi temperatur. B – S

B. Tes Menjodohkan

1. Pengertian

Tes bentuk ini disediakan dua kelompok bahan dan siswa harus mencari pasangan/jodoh-jodoh yang sesuai antara bahan yang ada dalam kelompok pertama dan pada kelompok kedua. Dengan demikian, tes menjodohkan terdiri atas satu seri pertanyaan dan satu seri jawaban. Masing-masing pertanyaan mempunyai jawaban yang tercantum dalam seri jawaban. Dalam tes ini siswa diminta mencari dan menempatkan setiap jawaban untuk setiap pertanyaan.

Terdapat dua bentuk tes penjodohan yaitu : (1) penjodohan sempurna, dan (2) penjodohan tidak sempurna. Dalam penjodohan sempurna setiap satu butir dari soal memiliki satu jawaban sebagai jodohnya, sedangkan dalam sistem penjodohan tak sempurna terdapat dua atau lebih butir jawaban yang sama mempunyai pasangan/jodoh yang sama.

2. Kelebihan dan Kelemahan Tes Menjodohkan

Kelebihan dari tes menjodohkan

- a. Baik untuk mengukur proses mental yang rendah
- b. Kemungkinan untuk mengukur proses mental yang tinggi tetap ada tapi sulit sekali
- c. Obyektif
- d. Mudah disusun
- e. Cocok untuk mengukur informasi-informasi yang berbentuk fakta dari sebuah pengertian hubungan pengertian atau konsep-konsep

Kelemahan tes menjodohkan

Kelemahan dari bentuk soal tes ini adalah sukar untuk mengukur proses mental yang tinggi dan siswa cenderung untuk membuat tafsiran-tafsiran.

3. Struktur dan Pedoman Menyusun Soal

Soal terdiri dari:

- a. Pengantar yaitu: penjelasan mengenai bagaimana cara mengerjakan soal. Pengantar berfungsi membentuk kerangka berfikir secara umum dan harus disusun secara singkat dan jelas.

- b. Persoalan yang berfungsi untuk membentuk kerangka berfikir lebih khusus, sempit, dan terbatas.
- c. Jawaban yang dipilih sesuai dengan persoalan atau menjadi jodoh /pasangan yang sesuai.

4. Petunjuk penyusunan adalah :

- a. Buatlah pengantar se jelas mungkin
- b. Bentuklah tes terdiri dari sederetan persoalan dan sederetan jawaban.
- c. Hal yang disusun baik dalam pertanyaan maupun jawaban hendaknya homogen
- d. Jumlah jawaban buatlah lebih banyak dari jumlah pertanyaan
- e. Batasi setiap kelompok jangan lebih dari 10 pertanyaan jika ingin banyak buatlah beberapa kelompok.
- f. Semua pertanyaan dan jawaban hendaknya dibuat dalam satu halaman saja
- g. Setiap satu pertanyaan hendaknya hanya ada satu jawaban yang benar.
- h. Buatlah kunci jawaban dan pedoman penilaian
- i. Sesudah dikerjakan nilailah dengan rumus.

Contoh soal:

Cocokkanlah simbol-simbol pada kolom II dengan elemen kimia pada kolom I, dengan cara menempatkan huruf pada ruang kosong di bagian depan setiap nomor pada kolom I.

Kolom I	Kolom II
..... 1. Oksigen	A. Na
..... 2. Besi	B. H
..... 3. Sodium	C. Au
..... 4. Karbon	D. I
..... 5. Khlorin	E. C
..... 6. Potasium	F. O
..... 7. Hidrogen	G. Fe
..... 8. Nitrogen	H. Ni
..... 9 Emas	I. K

..... 10. Nikel	J. P
..... 11. Iodine	K. N
	L. Cl

C. Tes Pilihan Ganda

Tes pilihan ganda digunakan secara luas dan termasuk butir tes tipe seleksi. Tes pilihan ganda dapat didesain untuk mengukur berbagai hasil belajar, mulai dari yang sederhana sampai yang kompleks. dan dapat menyediakan butir yang bermutu tinggi. Oleh karena itu, tes pilihan ganda memiliki peran yang penting dalam tes hasil belajar.

1. Sifat Butir Pilihan Ganda

Butir tes pilihan ganda terdiri dari sebuah pernyataan (*stem*), yang menampilkan suatu masalah (*problem*), dan beberapa alternatif jawaban (*options or choices*), yang menyediakan beberapa jawaban terhadap masalah. Pernyataan mungkin sebuah pertanyaan atau pernyataan yang tidak lengkap. Alternatif jawaban meliputi jawaban benar dan beberapa jawaban salah yang dinamakan distraktor (*distractors*). Fungsi dari distraktor untuk mengecoh siswa yang tidak pasti (*uncertain*) dengan jawaban yang benar.

Tipe lain dari butir pilihan ganda dalam bentuk semua jawaban benar (*the best answer*), yang mana alternatif semua benar tetapi hanya satu jawaban lebih benar dari yang lain. Tipe ini digunakan untuk hasil belajar yang lebih kompleks, sehingga siswa harus memilih alasan paling baik, metode terbaik untuk melakukan sesuatu atau aplikasi terbaik dari sebuah prinsip. Dengan demikian jawaban benar (*correct answer*) atau *the best answer* digunakan bergantung pada hasil belajar yang diukur.

Tipe butir pilihan ganda meliputi: tiga, empat, dan lima pilihan. Tentu pilihan lebih banyak akan mengurangi peluang siswa untuk memperoleh jawaban yang benar dengan terkaan. Secara teoritis, dengan lima alternatif pilihan hanya ada satu peluang dalam lima terkaan jawaban, sedangkan dengan empat alternatif pilihan hanya ada satu peluang dalam 4 terkaan jawaban. Bagaimanapun lebih sulit membuat lima pilihan yang masuk akal (*plausible*), dan butir-butir tidaklah bertambah baik (*improved*) dengan menambahkan jawaban yang salah untuk mempunyai lima

alternatif. Beberapa mungkin berisi tiga, empat, dan lima alternatif, tergantung pada ketersediaan distraktor yang masuk akal.

2. Penggunaan Butir Pilihan Ganda

Butir pilihan ganda dapat digunakan untuk mengukur pengetahuan (*knowledge*) dan berbagai tipe dari hasil belajar yang kompleks. Format butir tunggal mungkin paling banyak digunakan untuk mengukur pengetahuan, pemahaman, dan aplikasi.

a. Butir Pengetahuan

Tipe butir pengetahuan mengukur derajat yang berkaitan dengan materi belajar sebelumnya untuk diingat kembali. Fokus butir didasarkan pada mengingat kembali informasi sederhana dan terkait dengan pengukuran istilah, fakta, atau aspek pengetahuan spesifik lainnya.

Contoh: 1

Kebanyakan air tanah yang ada dalam tanah dan batuan berpori berasal dari:

- A. sungai
- B. gletser yang mencair
- C. sumber mata air
- D. salju
- E. hujan

Contoh: 2

Pengangkutan material yang sudah lapuk dilakukan oleh air yang mengalir akan menghasilkan batuan sedimen:

- A. Aeolik
- B. Glasial
- C. Klatik
- D. Akuatik
- E. Marin

Variasi dari pengetahuan yang dapat diukur dengan butir pilihan ganda terbaik ditunjukkan dengan mengilustrasikan beberapa tipe pertanyaan yang dapat ditanyai dalam berbagai kategori. Pertanyaan item 1 berhubungan dengan

pengetahuan fakta. Pertanyaan item 2 berhubungan dengan pengetahuan kecenderungan dan urutan.

Pertanyaan untuk butir pilihan ganda berhubungan erat dengan hasil belajar spesifik yang diukur, meliputi pengetahuan terminologi, pengetahuan fakta spesifik, pengetahuan konvensi, pengetahuan kecenderungan atau urutan, pengetahuan klasifikasi dan kategori, pengetahuan kriteria, dan pengetahuan metodologi. Grounlund (1993:43) memberikan beberapa sampel pertanyaan pilihan ganda tidak lengkap yang disajikan dalam sebuah kotak (*box*).

Pertanyaan yang ditunjukkan dalam kotak, hanya menyediakan sebuah sampel dari banyak pertanyaan yang dapat ditanyai. Juga pertanyaan dinyatakan dalam terminologi yang umum.

Tabel 3.1 Ilustrasi Pertanyaan Pengetahuan

1. Pengetahuan terminologi
Apa makna kata yang sama dengan _____?
Mana pernyataan terbaik mendefinisikan terminologi _____?
Dalam kalimat ini, apa makna dari kata _____?
2. Pengetahuan fakta spesifik
Dimana anda temukan _____?
Siapa yang menemukan pertama kali _____?
Apa nama dari _____?
3. Pengetahuan konvensi
Apa bentuk yang benar untuk _____?
Apa pernyataan yang menunjukkan pemakaian yang benar dari _____?
Mana aturan berikut yang berlaku bagi _____?
4. Pengetahuan kecenderungan dan urutan
Mana pernyataan berikut yang terbaik menguraikan kecenderungan dari _____?
Apa penyebab paling penting dari _____?
Mana indikasi berikut urutan yang sesuai _____?
5. Pengetahuan klasifikasi dan kategori
Apa tipe utama dari _____?
Apa klasifikasi utama dari _____?
Apa karakteristik dari _____?
6. Pengetahuan kriteria
Mana kriteria berikut ini untuk mempetimbangkan _____?
Mana kriteria paling penting untuk memilih _____?
Apa kriteria yang digunakan untuk mengklasifikan _____?
7. Pengetahuan metodologi
Apa metode yang digunakan untuk _____?
Apa cara terbaik untuk _____?
Apa langkah pertama dalam membuat _____?
8. Pengetahuan prinsip atau generalisasi
Mana pernyataan yang terbaik menyatakan prinsip _____?
Mana pernyataan yang terbaik menyimpulkan kepercayaan dari _____?

_____?

Mana prinsip berikut yang terbaik menjelaskan _____?

9. Pengetahuan teori dan struktur

Mana pernyataan berikut paling konsisten dengan teori dari _____?

Mana pernyataan berikut terbaik menjelaskan struktur dari _____?

Apa bukti terbaik mendukung teori dari _____?

Sumber ; Grounlund (1993: hal. 43).

b. Butir Pemahaman

Tipe butir pemahaman mengukur pada level lebih tinggi dari pengetahuan. Butir tes ini menentukan siswa-siswa telah menyerap arti dari materi tanpa membutuhkan mereka untuk mengaplikasikannya. Pemahaman diukur dengan membutuhkan siswa-siswa untuk menjawab dalam berbagai cara, tetapi yang penting butir-butir ada terbaru (*novely*). Butir tes berikut ini mengilustrasikan tipe-tipe umum dari hasil belajar pada level pemahaman.

Contoh: 3

Bila di suatu tempat ditemukan adanya sumber-sumber gas, sumber-sumber air panas, mata air mineral, maka dapat dipastikan bahwa daerah tersebut merupakan daerah:

- A. vulkanis*
- B. tanah pasir
- C. tanah padas
- D. tektonis

Contoh: 4.

Faktor-faktor yang mempengaruhi cuaca adalah:

- A. iklim, tekanan, dan suhu
- B. suhu, iklim, dan banyaknya penguapan
- C. suhu, ketinggian daerah, dan curah hujan
- D. suhu, tekanan, dan kelembaban*

Contoh: 5

Ada beberapa pernyataan tentang angin: (1) terjadi pada siang hari, (2) suhu daratan lebih tinggi dari suhu lautan, (3) mengalir dari laut ke darat, dan (4) terjadi pada malam hari. Pernyataan yang ada kaitannya dengan angin laut adalah:

- A. (4), (3), dan (2)
- B. (4), (3), dan (1)
- C. (1), (2), dan (3)*
- D. (1), (2), dan (4)

Variasi dari pemahaman yang dapat diukur dengan butir pilihan ganda ditunjukkan dengan mengilustrasikan beberapa tipe pertanyaan yang dapat ditanyai dalam berbagai kategori. Contoh 3 mengilustrasikan pertanyaan yang berhubungan dengan pemahaman tentang eksplorasi (*extrapolation*). Contoh 4 mengilustrasikan pertanyaan yang berhubungan dengan pemahaman tentang interpretasi (*interpretation*). Contoh 5 mengilustrasikan pertanyaan yang berhubungan dengan pemahaman tentang translasi (*translation*).

Beberapa hasil belajar pada level pemahaman dapat diukur dengan butir pilihan ganda yang diilustrasikan dengan pertanyaan tidak lengkap dalam kotak berikut.

Tabel 3.2 Ilustrasi Pertanyaan Pemahaman dan Aplikasi

<p>Pertanyaan Pemahaman</p> <p>Mana dari berikut ini sebuah contoh dari _____?</p> <p>Apa yang merupakan pikiran utama yang diungkap oleh _____?</p> <p>Apa perbedaan utama antara _____?</p> <p>Apa karakteristik umum dari _____?</p> <p>Mana dari berikut ini bentuk lain dari _____?</p> <p>Mana dari berikut ini penjelasan terbaik _____?</p> <p>Mana dari berikut ini ringkasan terbaik _____?</p> <p>Mana dari berikut ini ilustrasi terbaik _____?</p> <p>Apa yang anda prediksi akan terjadi jika _____?</p> <p>Apa kecenderungan anda memprediksi pada _____?</p> <p>Pertanyaan Aplikasi</p> <p>Mana dari berikut ini metode terbaik untuk _____?</p> <p>Apa langkah-langkah yang harus diikuti dalam aplikasi _____?</p> <p>Mana situasi yang membutuhkan penggunaan dari _____?</p> <p>Mana prinsip yang terbaik untuk menyelesaikan _____?</p> <p>Apa prosedur terbaik untuk memperbaiki _____?</p> <p>Apa prosedur terbaik untuk mengkonstruksi _____?</p> <p>Apa prosedur terbaik untuk mengoreksi _____?</p> <p>Mana dari berikut ini perencanaan terbaik untuk _____?</p> <p>Mana dari berikut ini yang menyediakan urutan sesuai untuk _____?</p> <p>Apa yang merupakan efek yang paling mungkin _____?</p>
--

Sumber ; Grounlund (1993: hal. 45).

c. Butir Aplikasi

Pada butir aplikasi, siswa harus menunjukkan bahwa mereka tidak hanya menyerap makna dari informasi tetapi juga menerapkannya pada situasi konkrit yang baru bagi mereka. Oleh karena itu, butir aplikasi menentukan landasan dimana siswa dapat memindahkan pembelajarannya dan menggunakannya secara efektif dalam memecahkan permasalahan yang baru. Jadi butir ini dapat pula digunakan untuk aplikasi berbagai aspek dari pengetahuan, seperti fakta, konsep, prinsip, aturan, metode, dan teori. Butir pemahaman dan aplikasi mampu beradaptasi untuk semua area dari mata pelajaran (*subject matter*).

Contoh: 6

Alat pemanas listrik memakai 5 A, apabila dihubungkan dengan sumber 110 V, maka hambatannya adalah:

- A. 0,5 ohm
- B. 5 ohm
- C. 22 ohm *
- D. 110 ohm

Contoh: 7

Sebuah partikel bermuatan $3,2 \times 10^{-19}$ C dan bermassa $6,4 \times 10^{-27}$ kg ditempatkan dari keadaan diam dalam suatu medan listrik homogen $4,0 \times 10^4$ N/C. Jika medan listrik yang ditimbulkan oleh muatan partikel itu sendiri diabaikan, maka kelajuan partikel setelah 0,02 detik adalah:

- A. $1,0 \times 10^{10}$ m/detik
- B. $2,0 \times 10^{10}$ m/detik
- C. $3,0 \times 10^{10}$ m/detik
- D. $4,0 \times 10^{10}$ m/detik *

Contoh: 8

Mana salah satu tipe hasil belajar berikut ini paling sulit untuk dievaluasi secara efektif.

- A. Aplikasi
- B. Apresiasi
- C. Fakta

D. Konsep

Ketika menulis butir aplikasi, hati-hati sewaktu memilih masalah bagi siswa yang tidak mampu menemui jawabannya dan karenanya tidak dapat diselesaikan berdasarkan pengetahuannya. Setiap butir harus didesain dengan baik untuk aplikasi dari fakta, konsep, prinsip, atau prosedur yang ditunjukkan dalam hasil belajar yang diharapkan. Lihat kotak di atas untuk beberapa pertanyaan yang mungkin ditanyai pada level aplikasi.

d. Butir Analisis

Pada butir analisis, siswa harus menunjukkan bahwa mereka tidak hanya menyerap makna dari informasi dan menerapkannya pada situasi konkrit yang baru, tetapi juga mampu menganalisis suatu informasi yang luas menjadi bagian-bagian kecil sehingga informasi tersebut menjadi lebih jelas. Oleh karena siswa dapat menggunakan suatu informasi atau pengetahuan yang diperolehnya untuk memecahkan masalah.

Contoh:

Kuat medan listrik di suatu titik sejauh r dari muatan titik q akan diperbesar menjadi 125 kali semula. Ini dapat dilakukan dengan cara memperbesar

- A. jarak menjadi $1/125$ kali dan muatan 5 kali
- B. jarak menjadi $1/5$ kali dan muatan 25 kali
- C. muatan menjadi $1/125$ kali dan jarak 5 kali
- D. muatan menjadi 5 kali dan jarak 25 kali.*

e. Butir Sintesis

Pada butir ini, siswa harus menunjukkan bahwa mereka tidak hanya menyerap makna dari informasi, tetapi juga mampu menggabungkan beberapa informasi menjadi suatu kesimpulan.

Contoh:

Pernyataan berikut berkaitan dengan pasang naik dan pasang surut air laut: (1) gaya gravitasi bulan dengan matahari mempengaruhi pasang surut air laut, (2) pasang naik tertinggi terjadi pada saat bulan purnama, (3) pasang surut terendah terjadi pada saat kuartil awal atau akhir, (4) matahari lebih mudah menarik air laut dari pada bulan. Pernyataan yang benar adalah:

- A. (1), (2), dan (3)*

- B. (1), (3), dan (4)
- C. (2), (3), dan (4)
- D. (2) dan (4)

f. Butir Evaluasi

Pada butir evaluasi, siswa harus menunjukkan bahwa mereka mampu mempertimbangkan mana yang baik dan mana yang buruk dan memutuskan untuk mengambil tindakan tertentu.

Contoh:

Sebuah kapasitor keping sejajar dibuat dengan menyisipkan selembat bahan penyekat diantara kedua kepingnya. Bahan yang disisipkan seperti dicantumkan di bawah ini

Bahan Penyekat	Permitivitas Relatif	Tebal (mm)
Teflon	2	0,4
Kuarsa	3	0,8
Gelas	4	1,0
Mika	5	1,2
Porselen	6	1,3

Kapasitas maksimum akan diperoleh dengan menggunakan lembaran:

- A. Gelas
- B. Kuarsa
- C. Porselen
- D. Teflon *

3. Kelebihan dan Kelemahan Butir Pilihan Ganda

Kelebihan dari tes pilihan ganda

1. Hasil belajar dari yang sederhana sampai yang kompleks dapat diukur.
2. Sangat terstruktur dan tugas jelas disajikan.
3. Sampel dari hasil belajar lebih luas dapat diukur.
4. Alternatif yang tidak benar memberikan informasi diagnostik
5. Skor sedikit dipengaruhi terkaan dari butir benar-salah
6. Menskor lebih mudah. obyektif, dan reliabel.

Kelemahan dari tes pilihan ganda

1. Mengkonstruksi butir yang bagus membutuhkan waktu yang lama

2. Sulit menentukan distraktor yang masuk akal
3. Butir ini tidak efektif untuk mengukur beberapa tipe pemecahan masalah dan kemampuan untuk mengorganisir dan mengungkap ide-ide
4. Skor dapat dipengaruhi oleh kemampuan membaca.

4. Aturan Menulis Butir Tes Pilihan Ganda

Aturan sebagai petunjuk untuk penulisan dari butir pilihan ganda yang berfungsi sesuai yang diharapkan seperti di bawah ini.

a. **Stem harus bermakna dan menyajikan masalah yang pasti.**

Sebuah kesalahan umum dalam penulisan butir pilihan ganda adalah memiliki *stem* yang singkat. Dalam kasus ini, bisa sulit untuk melihat maksud dari butir soal (*item*) setelah membaca stem. Untuk menulis sebuah item yang fokus, harus mencakup ide sentral di *stem* bukan dipilihan jawaban. Dalam butir soal 1, *stem* tidak ada masalah yang pasti.

Contoh yang kurang baik:

Butir soal 1

Non-logam:

- A. tidak dapat eksis sebagai padatan pada suhu kamar.
- B. hanya dapat digabungkan dengan logam untuk membentuk senyawa yang stabil.
- C. biasanya memiliki lebih dari tiga elektron di kulit terluar atom.*
- D. biasanya ditemukan di sisi kiri dari Tabel Periodik.

Jawaban yang benar adalah ditandai dengan tanda bintang (*). Siswa dihadapkan dengan empat pilihan benar-salah; masing-masing tentang non-logam, tetapi C merupakan satunya pilihan yang benar. Selanjutnya, empat pilihan mencakup suatu set ide kimia yang berbeda secara luas, sehingga evaluasi dengan perbandingan tidak mungkin. *Stem* dapat dinilai secara jelas menyajikan masalah, jika pilihan jawaban mengandung jenis konten yang paralel. Butir soal 2 menunjukkan salah satu cara untuk membuat *stem* menjadi masalah yang pasti. Siswa dapat berpikir tentang jawaban yang benar ketimbang mencari tahu apa masalahnya. Juga secara jelas masalah dinyatakan pada stem sehingga empat pilihan jawaban telah mengandung konten yang paralel.

Contoh yang lebih baik:

Butir soal 2

Seberapa banyak elektron dapat ditemukan di kulit terluar dari atom non logam?

- A. 1
- B. 2
- C. 3
- D. 4 *

Demikian pula, butir soal 3 adalah butir soal pilihan ganda yang kurang baik ditulis. Perhatikanlah contoh berikut ini!

Contoh kurang baik:

Butir soal 3

Manakah dari pernyataan berikut tentang sel elektrokimia yang benar?

- A. Ada reaksi kimia spontan dalam setiap sel elektrokimia.*
- B. Ggl dari sel elektrokimia diukur dalam joule.
- C. Anoda diberi label (+) sedangkan katoda diberi label (-).
- D. Jembatan garam menyediakan elektron untuk melengkapi rangkaian.

Contoh yang lebih baik:

Butir soal 4

Apa fungsi utama dari jembatan garam (*salt bridge*) dalam sel elektrokimia?

- A. ion pasokan pindah ke dua setengah-sel*
- B. menarik elektron dari satu setengah-sel dari setengah-sel yang lain
- C. menjaga tingkat solusi yang sama dalam dua setengah-sel
- D. pasokan elektron untuk melengkapi sirkuit

Stem pada butir soal 3 gagal menyajikan suatu masalah tertentu dan empat pilihan jawaban muncul menjadi gado-gado ide kimia. Butir soal 4 lebih fokus dari pada butir soal 3. *Stem* dari butir soal 4 menimbulkan masalah yang jelas dan menilai tujuan pembelajaran tunggal.

b. Nyatakan stem dengan bahasa sederhana dan jelas.

Masalah dalam stem dari sebuah butir pilihan ganda harus dinyatakan dengan tepat dan harus bebas dari kata kompleks yang tidak perlu. Stem butir

yang dinyatakan jelek sering ambigu (bermakna ganda) dan dapat mencegah siswa untuk memberikan jawaban secara benar. Juga, struktur kalimat yang kompleks membuat butir lebih mengukur pemahaman bacaan dari hasil belajar yang diharapkan. Perhatikan dua contoh berikut ini!

Contoh

Butir soal yang kurang baik

Kurt Vonnegut, Jr., seorang penulis dan kritik sosial Amerika yang kontroversial, mana buku yang ditulisnya berikut ini?

- A. Dynasty
- B. Player Piano*
- C. Shogun
- D. The Women's Room

Butir soal yang lebih baik:

Kurt Vonnegut, Jr., manakah buku yang ditulisnya berikut ini?

- E. Dynasty
- F. Player Piano*
- G. Shogun
- H. The Women's Room

Kesalahan umum yang lain dalam menyatakan butir pilihan ganda akan membebani *stem* dengan materi yang tidak relevan dan tidak fungsional. Ini mungkin disebabkan oleh keinginan guru untuk mentes siswa.

c. Menempatkan sebanyak mungkin kata di dalam stem

Hindari mengulangi materi yang sama dalam setiap alternatif jawaban. Dengan memindahkan semua konten umum kepada *stem*, biasanya memungkinkan untuk menjelaskan masalah lebih lanjut dan mengurangi waktu siswa untuk membaca alternatif jawaban. Perhatikanlah contoh berikut ini!

Contoh

Butir soal yang kurang baik:

John F. Kennedy terbunuh:

- A. pada tahun 1961
- B. pada tahun 1963
- C. pada tahun 1965
- D. pada tahun 1968

Butir soal yang lebih baik:

John F. Kennedy terbunuh pada tahun:

- E. 1961
- F. 1963
- G. 1965
- H. 1968

Dalam banyak masalah tidak mudah untuk memindahkan kata-kata umum pada *stem*. Perhatikan contoh berikut ini!

Contoh

Butir soal yang kurang baik:

Tujuan pembelajaran bermanfaat untuk mengkonstruksi tes ketika tujuan dinyatakan dalam suatu cara yang menunjukkan:

- A. konten mencakup selama periode pembelajaran
- B. Jenis kinerja siswa harus menampilkan sasaran belajar (*goal*)
- C. Berbagai hal guru lakukan untuk memperoleh hasil belajar siswa maksimum.
- D. Jenis aktivitas belajar selama belajar.

Butir soal yang lebih baik:

Tujuan pembelajaran paling bermanfaat untuk tujuan mengkonstruksi tes ketika dinyatakan dalam terminologi:

- A. konten mata pelajaran
- B. kinerja siswa.
- C. perilaku guru
- D. aktivitas belajar.

Susunan kata dan kejelasan mengungkapkan sasaran (*goal*) penting untuk memperbaiki konstruksi tes. *Item* berfungsi lebih baik ketika ramping (*slim*) dan seimbang (*trim*).

d. Hindari bagian yang kosong pada stem

Stem yang ditulis sebagai pernyataan tidak lengkap akan membutuhkan penyisipan dari pilihan jawaban yang benar. Ahli pengukuran menyarankan tidak menggunakan format yang tidak lengkap, karena mahasiswa harus

mempertahankan *stem* dalam memori jangka pendek sambil melengkapkan *stem* dengan masing-masing pilihan jawaban.

Contoh:

Butir soal yang kurang baik:

_____ memiliki rumus molekul $C_n H_{2n}$.

- A. Alkana
- B. Alkena *
- C. Alkanols
- D. Alkanoic asam

Butir soal yang lebih baik:

Jenis zat organik memiliki rumus molekul $C_n H_{2n}$?

- A. alkana
- B. alkena *
- C. alkanols
- D. alkanoic asam

Jika format lengkap tidak dapat dihindari, maka bagian yang kosong harus terjadi menjelang akhir *stem* bukan di tengah atau di awal *stem*.

e. **Nyatakan stem dari butir dalam bentuk positif**

Stem sedapat mungkin dinyatakan dalam pernyataan atau pertanyaan positif.

Perhatikan contoh berikut ini!

Contoh.

Butir soal yang kurang baik:

Di dalam mempertimbangkan kondisi lingkungan ideal dalam suatu kelas, fisik menyenangkan tidak pasti tanpa:

- A. Temperatur diatur pada $18^{\circ}C$, kelembaban relatif 50 %.
- B. Temperatur diatur pada $23^{\circ}C$, kelembaban relatif 72 %
- C. Temperatur diatur pada $20^{\circ}C$, kelembaban relatif 25 %
- D. Temperatur diatur pada $20^{\circ}C$, kelembaban relatif 50 % *

Butir soal yang lebih baik:

Kondisi lingkungan yang paling *ideal* untuk memastikan fisik senang dalam kelas adalah:

- A. Temperatur diatur pada 18 °C, kelembaban relatif 50 %.
- B. Temperatur diatur pada 23 °C, kelembaban relatif 72 %
- C. Temperatur diatur pada 20 °C, kelembaban relatif 25 %
- D. Temperatur diatur pada 20 °C, kelembaban relatif 50 % *

Pada contoh butir soal yang kurang baik, pernyataan mengandung dua kata negatif yang membingungkan yakni kata tidak dan tanpa. Dengan konversi sederhana dapat diubah menjadi pertanyaan secara langsung seperti pada contoh soal yang lebih baik.

Jika pernyataan negatif harus digunakan, maka kata negatif dibuat dengan huruf besar (*capitalize*), digaris bawah (*underscore*), atau ditulis tebal (*bold*). Mengganti istilah negatif dengan kata *kecuali*, kadang-kadang dapat meningkatkan kejelasan, sebagai diilustrasikan dalam angka 10. Beberapa siswa akan mengabaikan unsur negatif dalam *stem* karena kata *kecuali* sengaja ditempatkan di ujung *stem* dan dikapitalisasi. Perhatikan contoh berikut ini!

Contoh:

Butir soal yang kurang baik:

Air-jenis pemadam ini tidak cocok untuk memadamkan api disebabkan oleh pembakaran...

- A. alkohol.*
- B. kapas.
- C. kertas.
- D. kayu.

Butir soal yang lebih baik:

Air-jenis pemadam yang cocok untuk memadamkan api yang disebabkan oleh pembakaran, **KECUALI**:

- A. alkohol.*
- B. kapas.
- C. kertas.
- D. kayu.

f. Pastikan jawaban diharapkan benar atau terbaik

Ketika bentuk jawaban benar dari butir pilihan ganda digunakan, hanya ada satu jawaban yang benar. Dalam bentuk jawaban semua benar (*the best answer*)

jawaban yang diharapkan hanya satu yang paling benar. Perhatikan contoh berikut!

Contoh

Butir soal yang kurang baik:

Menghisap ibu jari menghasilkan trauma yang psikologis terbesar

- A. selama masa kanak-kanak
- B. pada periode sebelum masuk sekolah
- C. sebelum masa remaja *
- D. selama masa remaja

Butir soal yang lebih baik

Menghisap ibu jari menghasilkan trauma yang psikologis terbesar

- A. lahir sampai 2 tahun
- B. 3 – 5 tahun
- C. 6 – 12 tahun*
- D. 13 – 20 tahun

Pada butir kurang baik bisa membingungkan antara respon c, a, atau b, karena istilah sebelum remaja, selama masa kanak-kanak, dan sebelum masuk sekolah. Periode umur spesifik, seperti ditunjukkan pada butir lebih baik, sehingga akan mengurangi makna ganda (ambigu).

g. Pastikan semua alternatif jawaban konsisten dengan stem dan paralel

Jawaban benar biasanya secara hati diutarakan, sehingga konsisten secara tata bahasa dengan stem. Pembuat tes cenderung untuk tergelincir di dalam menyatakan distraktor, sehingga distraktor berfungsi tidak efektif dan ini akan memberikan suatu petunjuk kepada jawaban yang benar.

Langkah umum yang dapat diambil untuk mencegah tidak konsisten secara tata bahasa dengan menghindari memakai kata sebuah atau suatu pada ujung *stem* dari butir. Perhatikan contoh berikut!

Contoh

Butir soal yang kurang baik:

Daya ingat dari informasi berdasarkan fakta dapat diukur paling baik dengan sebuah:

- A. butir penjodohan

- B. butir pilihan ganda
- C. butir jawaban singkat
- D. pertanyaan esai.

Butir soal yang lebih baik

Daya ingat dari informasi berdasarkan fakta dapat diukur paling baik dengan :

- A. butir-butir penjumlahan
- B. butir-butir pilihan ganda
- C. butir-butir jawaban singkat
- D. pertanyaan-pertanyaan esai.

Penempatan kata sebuah pada butir kurang baik membuat distraktor terakhir salah. Dengan mengubah alternatif dari tunggal ke jamak, memungkinkan menghilangkan kata tersebut. Dalam kasus lain, mungkin perlu memakai kata suatu atau sebuah pada setiap alternatif jawaban.

h. Hindari jawaban benar lebih rinci

Jawaban benar lebih rinci dapat memberikan suatu petunjuk. Jawaban yang secara khas dihubungkan dengan pernyataan yang benar, sebagai contoh kadang-kadang (*sometime*), boleh (*may*), pada umumnya (*usually*), itu lebih disukai untuk dipilih.

Contoh

Butir soal yang kurang baik:

Kurangnya perhatian pada hasil belajar selama persiapan tes:

- A. akan menurunkan kualitas teknis butir
- B. akan membuat konstruksi butir tes lebih sulit.
- C. akan mengakibatkan semakin besar penggunaan pertanyaan esai.
- D. boleh mengakibatkan tes kurang relevan dengan tujuan pembelajaran.*

Istilah "boleh" jelas nyata di dalam contoh, tetapi kesalahan jenis ini adalah umum dan sering terlihat di dalam sebuah bentuk yang hampir tidak kentara (*subtler*).

Butir soal yang lebih baik:

Kurangnya perhatian pada hasil belajar selama persiapan tes akan...

- A. menurunkan kualitas teknis butir tes.

- B. membuat konstruksi butir tes lebih sulit.
- C. mengakibatkan semakin besar penggunaan pertanyaan esai.
- D. mengakibatkan tes kurang relevan dengan tujuan pembelajaran. *

Pada contoh butir soal yang lebih baik, kata boleh pada pilihan jawaban d dihilangkan dan kata akan pada pilihan jawaban a, b, dan c dipindahkan ke stem.

i. **Hindari menggunakan dua respon yang inklusif**

Penggunaan dua respon yang inklusif memungkinkan untuk mengeliminasi alternatif lain, karena salah satu dari dua respon harus sungguh-sungguh jawaban yang benar.

Contoh

Butir soal yang kurang baik:

Mana salah satu tipe tes berikut ini yang mengukur hasil belajar pada level analisis?

- A. Butir tipe memberikan jawaban
- B. Butir tipe memilih jawaban
- C. Butir menjodohkan
- D. Butir pilihan ganda

Oleh karena dua alternatif sudah mencakup dua tipe utama butir tes, bahkan disiapkan dengan kurang baik sehingga siswa membatasi pilihannya pada dua ini. Tentu saja ini, memberi peluang kepada siswa sama besar (*fifty-fifty*) dalam menebak jawaban yang benar itu. Oleh karena itu kedua alternatif diganti dengan jawaban yang lain. Perhatikan contoh butir soal lebih baik

Butir soal yang lebih baik:

Mana salah satu tipe tes berikut ini yang mengukur hasil belajar pada level analisis?

- A. Butir tipe benar-salah
- B. Butir tipe melengkapi jawaban
- C. Butir menjodohkan
- D. Butir pilihan ganda

j. **Hindari menggunakan dua respon yang memiliki makna sama**

Penggunaan dua respon yang memiliki makna sama memungkinkan mereka untuk mengeliminasi sebagai jawaban yang potensial. Jika dua alternatif memiliki makna yang sama dan hanya ada satu jawaban untuk dipilih, maka itu secara jelas yang kedua alternatif harus salah.

Contoh

Butir soal yang kurang baik:

Mana salah satu karakteristik penting berikut ini dari hasil tes belajar?

- A. Konsistensi
- B. Reliabilitas
- C. Relevansi
- D. Obyektif

Pada butir ini konsistensi dan reliabilitas dapat dieliminasi, sebab mereka memiliki makna secara esensial sama, konsistensi diganti dengan valid.

Butir soal yang lebih baik:

Mana salah satu karakteristik penting berikut ini dari hasil tes belajar?

- A. Valid
- B. Reliabilitas
- C. Relevansi
- D. Obyektif

k. **Buatlah distraktor yang masuk akal dan menarik**

Alternatif jawaban yang masuk akal dan menarik dapat dengan mudah dihasilkan dengan mengatur sebuah stem sebagai suatu pertanyaan yang terbuka. Respon mereka dapat memberikan alternatif jawaban yang masuk akal. Perhatikan contoh berikut ini

Contoh

Butir soal yang kurang baik:

IQ (*Intelligence Quotient*) dari seorang siswa yang memiliki *mental age* 10 dan *chronical age* 8 adalah:

- A. 75
- B. 100
- C. 125 *
- D. 150

Butir soal yang lebih baik:

IQ (*intelligence quotient*) dari seorang siswa yang memiliki mental age 10 dan chronological age 8 adalah:

- A. 80
- B. 102
- C. 118
- D. 125 *

Setiap pilihan di atas merupakan jawaban yang dapat menghasilkan kesalahan formula umum dibuat dalam menghitung IQ. Dalam respek ini, alternatif yang tidak benar adalah masuk akal. Dengan menganalisis sifat-sifat dari respon yang tidak benar, siswa dan guru dapat menunjukkan dengan tepat sumber kesalahan dan memperoleh keuntungan pengertian yang mendalam tentang kesulitan belajar.

I. Hindari jawaban benar ditulis lebih panjang dari jawaban salah

Ada kecenderungan jawaban yang benar ditulis lebih panjang dari jawaban yang salah, karena memenuhi persyaratan pernyataan untuk membuatnya dengan tegas benar. Tentu saja ini, memberikan suatu petunjuk bagi siswa dalam mengerjakan tes. Panjang jawaban benar dapat diatur dengan memvariasikannya. Jawaban benar kadang-kadang lebih panjang, kadang-kadang lebih pendek, dan sekali waktu panjangnya sama. Dalam beberapa kasus lebih diinginkan untuk membuat alternatif yang kira-kira panjangnya sama dengan mengatur distraktor dari pada jawaban yang benar.

Contoh

Butir soal yang kurang baik

Salah satu keuntungan dari butir pilihan ganda dibandingkan dengan essai adalah:

- A. mengukur hasil yang komplek
- B. lebih tergantung pada ingatan
- C. membutuhkan sedikit waktu untuk menyekor
- D. menyediakan suatu sampel yang luas dari konten materi ajar *

Butir lebih baik

Salah satu keuntungan dari butir pilihan ganda dibandingkan dengan esai adalah:

- A. untuk pengukuran hasil belajar yang lebih kompleks
- B. lebih menekankan pada ingatan dari informasi fakta
- C. membutuhkan sedikit waktu untuk menyiapkan dan menyekor
- D. menyediakan suatu sampel yang luas dari konten materi ajar*

Panjang distraktor seperti yang dilakukan dalam butir soal yang lebih baik, direvisi jawaban sehingga memiliki panjang yang sama dengan jawaban yang benar.

m. Hindari pilihan semua di atas dan tidak satupun di atas benar

Pemakaian ungkapan semua di atas melemahkan daya sebuah butir. Siswa hanya membandingkan dua pilihan; jika kedua-duanya bisa diterima, semua di atas benar adalah jawaban yang logis tersebut. Sebaliknya, jika salah satu dari dua pilihan tidak dapat diterima (*unacceptable*), tidak hanya satu pilihan dieliminasi, tetapi juga pilihan semua di atas dieliminasi. Di dalam kasus manapun, siswa dapat mengira jawaban benar tanpa benar-benar mengetahui informasi itu. Perhatikanlah contoh berikut!

Contoh

Butir soal yang kurang baik:

Apa istilah statistika untuk rata-rata aritmatik

- A. Mean *
- B. Median
- C. Mode
- D. Semua di atas benar

Butir soal yang lebih baik:

Apa istilah statistika untuk rata-rata aritmatik

- A. Mean *
- B. Median
- C. Mode
- D. Ukuran nilai pusat

Pilihan d tidak menambah apapun kepada butir tes. Barangkali itu digunakan karena penulis tes tidak bisa berpikir tentang suatu pilihan yang keempat. Suatu alternatif masuk akal untuk semua di atas diganti dengan *central tendency* (ukuran nilai pusat), suatu pilihan lebih baik untuk jawaban yang tidak benar.

Pemakaian *none of above* (tidak satupun dari di atas) menjadi lebih baik dibandingkan dengan semua di atas benar. Tidak sama dengan ungkapan semua di atas (*all of the above*), tidak satupun dari di atas mencegah seorang siswa dalam menebak jawaban yang benar. Tidak satupun dari di atas dengan aman dipilih setelah setiap butir lain dibaca dan dieliminasi sebagai yang tidak dapat diterima. Tetapi, ada dua kritik utama dari respon ini. Pertama, memilih tidak satupun di atas tidak menyediakan apapun informasi diagnostik untuk membantu guru atau siswa menunjukkan dengan tepat kesalahan konsep (*misconception*) siswa. Ke dua, menggunakan tidak satupun di atas benar dapat mengurangi daya beda dan reliabilitas tes. Perhatikan contoh berikut ini!

Contoh:

Butir soal yang kurang baik

Tujuan hak/kebenaran Bill (*Bill of rights*) untuk:

- A. Mebebaskan perbudakan.
- B. memberikan kepada setiap orang hak untuk memilih.
- C. menjamin kebebasan individu*
- D. tidak satupun dari di atas

Butir soal yang lebih baik

Tujuan hak/kebenaran Bill (*Bill of rights*) untuk:

- A. Mebebaskan perbudakan.
- B. memberikan kepada setiap orang hak untuk memilih.
- C. menjamin kebebasan individu*
- D. menetapkan hak untuk membebaskan pendidikan bagi publik

Kelemahan pertama jelas sederhana dan kedua telah ditunjukkan pada butir soal yang kurang baik. Seorang siswa yang kompeten membaca butir ini bisa memberi alasan bahwa tujuan *Bill Rights* benar-benar untuk membatasi kekuasaan pemerintahan pusat. Sebagai konsekuensi, seorang siswa memilih respon d, tidak satupun dari di atas.

n. Atur pilihan jawaban pada suatu urutan yang logis

Untuk efisiensi dan memudahkan membaca, disarankan pilihan jawaban itu disusun menurut abjad di dalam kasus terminologi atau nama, sesuai nomor kasus, dan secara kronologis pada kasus tersebut. Siswa dapat mengikuti urutan dengan mudah dan berkonsentrasi pada tugas kognitif dalam mengidentifikasi respon yang benar tanpa keharusan untuk mengatur kembali informasi itu.

Contoh

Butir soal yang kurang baik:

Pemerintahan Amerika Serikat diproklamirkan pada hakekatnya di bawah konstitusi pada tahun:

- A. 1792
- B. 1776
- C. 1785
- D. 1789 *

Butir soal yang lebih baik:

Pemerintahan Amerika Serikat diproklamirkan pada hakekatnya di bawah konstitusi pada tahun:

- A. 1776
- B. 1785
- C. 1789 *
- D. 1792

Pada contoh ini pilihan jawaban ditunjukkan dalam urutan yang logis. Urutan yang logis lebih disukai karena memungkinkan siswa untuk membaca dengan cepat untuk jawaban yang benar itu.

o. Tulislah pilihan jawaban lebih homogen

Untuk meningkatkan kesukaran dengan membuat alternatif yang lebih homogen. Perhatikan contoh berikut!

Contoh:

Butir soal yang kurang baik:

Siapakah di antara nama-nama di bawah ini yang menemukan telepon?

- A. Bell

- B. Marconi
- C. Morse
- D. Pasteur

Butir soal yang lebih baik:

Siapakah di antara nama-nama di bawah ini yang menemukan telepon?

- A. Bell
- B. Marconi
- C. Morse
- D. Edison

Pada contoh butir soal yang kurang baik alternatif jawaban belum homogen, karena alternatif jawaban Pasteur berkaitan unsur religi, dan bukan berkaitan dengan unsur teknologi. Dengan demikian jawaban ini direvisi supaya homogen, seperti ditunjukkan pada contoh butir soal yang lebih baik.

p. Posisi jawaban yang benar di acak

Banyak penulis pertanyaan pilihan ganda cenderung untuk menempatkan jawaban yang benar pada posisi tengah, pilihan b atau c, sedangkan a dan d dalam format empat pilihan jarang digunakan. Suatu pola respon, jika terdeteksi oleh siswa, dapat mengurangi validitas tes dengan menetapkan suatu penyimpangan (bias) respon. Untuk memilih posisi respon yang pantas, urutannya di acak. Perhatikan suatu pola pengacakan posisi jawaban di bawah ini!

Suatu pola posisi respon kurang baik

Pertanyaan	Jawaban benar
1.	B
2.	C
3.	C
4.	B
5.	B
6.	C
7.	B
8.	D

Suatu pola posisi respon lebih baik

Pertanyaan	Jawaban benar
1.	B
2.	D
3.	A
4.	C
5.	D
6.	A
7.	C
8.	B

Ada beberapa teknik yang dapat dikembangkan dalam pengacakan posisi jawaban antara lain: *the roll of the die* (gulungan mati); undian suatu kartu; atau suatu tabel dari nomor-nomor yang diacak. Prosedur sederhana lain yang dapat digunakan dengan membuka suatu buku pada suatu penempatan acak dan penggunaan digit terakhir menyangkut nomor untuk menghasilkan posisi butir test.

- o. Buat setiap butir independen dengan butir lainnya.**
- p. Gunakan sebuah format butir dengan efisien.**
- q. Ikuti aturan tata bahasa.**
- r. Abaikan suatu aturan jika ingin memperbaiki efektifitas butir.**

5. Jenis Butir Tes Pilihan Ganda

Ada beberapa tipe dari butir tes pilihan ganda yaitu : (1) pilihan ganda biasa, (2) pilihan ganda analisis hubungan antar hal, (3) pilihan ganda analisis kasus, (4) pilihan ganda kompleks, (5) pilihan ganda semua jawaban benar (*the best answer*), dan (6) pilihan ganda menggunakan gambar, diagram, grafik, atau tabel. Keenam tipe pilihan ganda dibahas pada bagian berikut ini.

a. Pilihan ganda biasa

Contoh-contoh yang ditampilkan pada aturan menulis butir pilihan ganda di atas semuanya mengacu pada variasi pilihan ganda biasa.

b. Pilihan ganda analisis hubungan antar hal

Butir pilihan ganda analisis hubungan antar hal terdiri dari dua pernyataan. Kedua pernyataan ini dihubungkan oleh kata "SEBAB". Jadi, peserta tes atau siswa harus menganalisis pernyataan pertama dan kedua. Apakah pernyataan pertama benar atau salah? Apakah pernyataan kedua benar atau salah? Apakah pernyataan pertama ada hubungan sebab akibat atau tidak ada hubungan sebab akibat dengan pernyataan kedua. Adanya berbagai hal yang harus dianalisis dan dinilai oleh peserta tes atau siswa dari kedua pernyataan tersebut, maka dikembangkan tes bentuk hubungan antar hal yang memiliki beberapa alternatif pilihan jawaban sebagai berikut:

- A. Jika pernyataan pertama benar dan pernyataan kedua benar dan keduanya menunjukkan sebab akibat.

- B. Jika pernyataan pertama dan kedua benar tetapi keduanya tidak menunjukkan sebab akibat.
- C. Jika pernyataan pertama benar dan pernyataan kedua salah.
- D. Jika pernyataan pertama salah dan pernyataan kedua benar.
- E. Jika pernyataan pertama dan kedua salah.

Contoh

Amandemen terhadap Undang-Undang Dasar 1945 pada dasarnya merupakan penataan dari ketatanegaraan Indonesia.

SEBAB

Masyarakat Indonesia berada dalam masa transisi, karena itu mengalami banyak perubahan dalam perkembangan.

Pada contoh ini, pernyataan pertama benar dan pernyataan kedua benar. Kedua pernyataan memiliki hubungan sebab akibat. Dengan demikian pilihan jawaban A atau kunci jawaban A

c. Pilihan ganda kompleks

Butir pilihan ganda kompleks ini biasa disebut asosiasi pilihan ganda. Struktur pertanyaan dilengkapi dengan jawaban yang benar. Pada pilihan ganda biasa ada satu jawaban yang benar, tetapi pada pilihan ganda kompleks ini jawaban yang benar lebih dari satu, mungkin 2, 3, atau 4. Peserta tes atau siswa memilih jawaban sesuai dengan petunjuk yang diberikan. Adapun petunjuk pilihan ganda dengan jawaban yang benar lebih dari satu, yaitu setiap soal terdapat 4 alternatif jawaban, tetapi yang benar bisa satu atau lebih dari satu.

- .Jawablah A apabila yang benar alternatif jawaban no. 1, 2, dan 3
- .Jawablah B apabila yang benar alternatif jawaban no. 1, dan 3
- .Jawablah C apabila yang benar alternatif jawaban no. 2, dan 4
- .Jawablah D apabila yang benar alternatif jawaban no. 4
- .Jawablah E apabila semua alternatif jawaban benar

Contoh

Berdasarkan Undang-Undang No. 32 tahun 2004 tentang otonomi daerah, maka kewenangan pemerintah pusat yang tidak diberikan ke daerah meliputi:

- (1). Politik Luar negeri

- (2). Agama
- (3). Peradilan
- (4). Moneter dan fiskal

Pada contoh ini, alternatif pilihan (1), (2), (3), dan (4). Dengan demikian sesuai petunjuk dalam mengerjakan tes, maka pilihan jawaban E atau kunci jawaban E.

d. Pilihan ganda analisis kasus

Pada butir pilihan ganda analisis kasus, peserta tes atau siswa dihadapkan pada suatu kasus yang disajikan dalam bentuk cerita, peristiwa, dan sejenisnya. Dari kasus yang disajikan tersebut diajukan beberapa pertanyaan.

Bentuk tes pilihan ganda analisis kasus ini menggunakan petunjuk berikut: Bacalah wacana di bawah ini dengan teliti, kemudian jawablah pertanyaan-pertanyaan berikutnya dengan teliti!

Contoh

Akuntabilitas publik adalah suatu ukuran atau standar yang menunjukkan seberapa besar tingkat kesesuaian penyelenggaraan penyusunan kebijakan publik dengan peraturan hukum dan perundang-undangan yang berlaku untuk organisasi publik yang bersangkutan. Pada dasarnya, setiap pengambilan kebijakan publik akan memiliki dampak tertentu pada sekelompok orang atau seluruh masyarakat, baik dampak yang menguntungkan atau merugikan, maupun langsung atau tidak langsung. Oleh karena itu, penyusun kebijakan publik harus dapat mempertanggungjawabkan setiap kebijakan yang diambilnya kepada publik. Akuntabilitas merupakan pandangan yang didasarkan pada *fairness* semua komponen birokrasi yang seharusnya berpartisipasi dalam proses pengambilan keputusan, mendapat kesempatan dalam proses pengambilan keputusan (dituangkan dalam Undang-Undang) lebih aktif, lebih tersistem. Semua kebijakan siap dilaksanakan secara tersistem.

Akuntabilitas sektor publik terkait erat dengan kinerja sektor publik dengan fokus tidak hanya pada kepatuhan terhadap peraturan perundang-undangan tetapi juga lebih pada bagaimana mencapai *outcomes* dengan efisien dan efektif. Oleh karena itu, sering dikatakan bahwa *'Accountability is not just about the right thing but doing it well'*.

Penerapan *good governance* harus berlandaskan pada prinsip negara hukum dan demokrasi, serta penyelenggaraan pelayanan kepada masyarakat merupakan

hal yang tidak dapat dipisahkan dari akuntabilitas publik. Penyelenggaraan negara secara individual bertanggungjawab kepada kode etik dan perilaku yang mengikat secara profesional. Apabila semua yang tersebut di atas dapat terpenuhi, kepercayaan rakyat kepada aparat dan keandalan lembaga pemerintahan yang ada akan tumbuh. Penyelenggaraan pemerintahan yang tidak menerapkan akuntabilitas akan menimbulkan penyalahgunaan wewenang. Dengan penerapan prinsip akuntabilitas tersebut, diharapkan pertanggungjawaban penyelenggaraan pemerintah/institusi/unit kerja tidak lagi sekedar laporan kesan-kesan dan pesan-pesan, tetapi menjadi laporan pertanggungjawaban kinerja selama yang bersangkutan menjabat. Hal ini sejalan dengan kebijakan Anggaran Berbasis Kinerja.

Sumber: **Sekretariat Tim Pengembangan Kebijakan Nasional Tata Pemerintahan yang Baik BAPPE**

- 1). Pikiran utama wacana di atas adalah
 - A. Akuntabilitas publik
 - B. Anggaran berbasis kinerja
 - C. Penerapan *good governance* pada sektor publik
 - D. Tangung jawab pada publik
 - E. Pertanggungjawaban

- 2). Pikiran utama wacana di atas dikembangkan penulis dengan cara ...
 - A. Argumentasi
 - B. Defnisi
 - C. Deskripsi
 - D. Eksposisi
 - E. Persuasi

- 3). Padanan kata yang tepat untuk kata *fairness* pada wacana di atas...
 - A. Keterbukaan
 - B. Kesepakatan
 - C. Kesepahaman
 - D. Kemauan
 - E. Kejujuran

Kunci jawaban untuk soal 1) adalah E, soal 2) adalah B, dan soal 3) adalah E. Dari contoh ini dapat dikatakan bahwa dari satu kasus dapat dibuat beberapa pertanyaan dan aspek dan proses berpikir yang lebih tinggi dari pemahaman.

e. Pilihan ganda semua jawaban benar

Pada butir pilihan ganda semua jawaban benar, peserta tes atau siswa dihadapkan pada suatu pertanyaan atau pernyataan yang semua alternatif pilihan jawaban benar dan tidak ada jawaban yang salah. Dari bentuk tes ini siswa harus memilih alternatif pilihan yang paling benar atau paling sesuai dengan dirinya. Kemudian semua jawaban tersebut akan diterjemahkan menjadi nilai skala 1 sampai 4 atau 5.

Bentuk tes pilihan ganda semua jawaban benar menggunakan petunjuk berikut: Kemukakanlah jawaban Saudara terhadap pernyataan-pernyataan berikut dengan memilih salah satu alternatif jawaban yang telah disediakan. Saudara dipersilahkan memberi tanda silang (X) salah satu huruf (A, B, C, D, atau E) pada lembaran jawaban sesuai dengan alternatif jawaban yang Saudara pilih.

Contoh

Dalam beberapa kali rapat, salah seorang peserta membantah pendapat/ide yang dikemukakan peserta rapat lain dan memaksakan idenya untuk diterima. Bagaimana tanggapan Saudara?

ALTERNATIF JAWABAN		NILAI
A	Bagaimanapun pendapat orang tersebut perlu dipertimbangkan	5
B	Kenyataan tersebut adalah wajar, karena semua peserta rapat mempunyai hak untuk mengeluarkan pendapat/ide	3
C	Orang tersebut akan saya nasehati, agar tidak mengganggu kelancaran rapat	2
D	Pimpinan perlu arif dalam memimpin rapat	4
E	Untuk rapat berikutnya sebaiknya orang tersebut tidak perlu diikuti	1

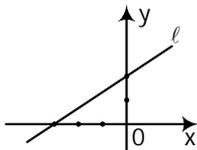
f. Pilihan ganda menggunakan gambar, diagram, grafik, atau tabel

Butir pilihan ganda ini mirip dengan analisis kasus, baik struktur maupun pola pertanyaannya. Bedanya dalam tes bentuk ini tidak disajikan kasus dalam bentuk

cerita atau peristiwa, tetapi kasus tersebut berupa diagram, gambar, grafik, atau tabel.

Contoh

Perhatikan gambar di bawah!



Persamaan garis g yang melalui $O(0, 0)$ dan tegak lurus pada l adalah... .

- A. $y = -x$
- B. $y = -2x$
- C. $y = -3x$
- D. $y = -\frac{2}{3}x$
- E. $y = -\frac{3}{2}x$ *

Butir pilihan ganda yang menggunakan diagram, gambar, grafik, dan tabel dapat mengukur aspek dan proses berpikir yang lebih tinggi dari hanya aspek ingatan.

Rangkuman

Tes obyektif adalah tes yang menghendaki peserta tes (testee) untuk memilih jawaban yang benar dari beberapa alternatif jawaban yang tersedia. Tes obyektif ini ada beberapa jenis, yaitu: tes benar-salah, tes menjodohkan, dan tes pilihan ganda. Untuk tes pilihan ganda dapat diklasifikasikan menjadi: (1) pilihan ganda biasa, (2) pilihan ganda analisis hubungan antar hal, (3) pilihan ganda analisis kasus, (4) pilihan ganda kompleks, (5) pilihan ganda semua jawaban benar (*the best answer*), dan (6) pilihan ganda menggunakan gambar, diagram, grafik, atau tabel.

Ketiga jenis tes obyektif ini memiliki keunggulan dan kelemahannya serta ada pula aturan (petunjuk) dalam penyusunannya. Aturan dalam penyusunan tes benar salah antara lain: (1) soal harus singkat, jelas bukan kalimat majemuk, (2) jumlah soal harus banyak dan disusun berdasarkan tabel spesifikasi, (3) satu soal harus berisi satu persoalan, (4) tidak memakai kata seperti, selalu, seringkali, pada umumnya, biasanya, karena kata-kata seperti itu memudahkan siswa menerka jawabannya, (5) setiap pertanyaan/ soal harus pasti salah atau betul (tidak mendua

arti), (6) jumlah soal yang betul dan salah harus seimbang, (7) urutan soal yang betul dan salah harusnya tidak mengikuti pola yang teratur, (8) sebaiknya pertanyaan tidak diambil langsung dari buku, dan (9) tulisan huruf B - S pada permulaan nomor pada masing- masing item dengan maksud mempermudah pengerjaannya dan menilai (skoring).

Aturan dalam penyusunan tes menjodohkan antara lain: (1) buatlah pengantar sejelas mungkin, (2) bentuklah tes terdiri dari sederetan persoalan dan sederetan jawaban, (3) hal yang disusun baik dalam pertanyaan maupun jawaban hendaknya homogen, (4) jumlah jawaban buatlah lebih banyak dari jumlah pertanyaan, (5) batasi setiap kelompok jangan lebih dari 10 pertanyaan jika ingin banyak buatlah beberapa kelompok, (6) semua pertanyaan dan jawaban hendaknya dibuat dalam satu halaman saja, (7) setiap satu pertanyaan hendaknya hanya ada satu jawaban yang benar, dan (8) buatlah kunci jawaban dan pedoman penilaian.

Aturan dalam penyusunan tes pilihan ganda antara lain: (1) stem harus bermakna dan menyajikan masalah yang pasti, (2) nyatakan stem dengan bahasa sederhana dan jelas, (3) menempatkan sebanyak mungkin kata di dalam stem, (4) hindari bagian yang kosong pada stem, (5) nyatakan stem dari butir dalam bentuk positif, (6) pastikan jawaban diharapkan benar atau terbaik, (7) pastikan semua alternatif jawaban konsisten dengan stem dan paralel, (8) hindari jawaban benar lebih rinci, (9) hindari menggunakan dua respon yang inklusif, (10) hindari menggunakan dua respon yang memiliki makna sama, (11) buatlah distraktor yang masuk akal dan menarik, (12) hindari jawaban benar ditulis lebih panjang dari jawaban salah, (13) hindari oilihan semua di atas dan tidak satupun di atas benar, (14) atur pilihan jawaban pada suatu urutan yang logis, (15) tulislah pilihan jawaban lebih homogen, (16) posisi jawaban yang benar di acak, (17) buat setiap butir independen dengan butir lainnya, (18) Gunakan sebuah format butir dengan efisien, (19) Ikuti aturan tata bahasa, dan (20) abaikan suatu aturan jika ingin memperbaiki efektifitas butir.

Untuk saat ini dari ketiga jenis tes obyektif di atas, namun yang banyak diterapkan di sekolah adalah tes pilihan ganda. Tentunya ada beberapa alasan atau pertimbangan, mengapa tes pilihan ganda lebih banyak digunakan di sekolah untuk menguji kemampuan dan keterampilan dari peserta didik.

Soal-Soal Latihan

Kerjakan soal di bawah ini.

1. Tulislah lima pernyataan (*statement*) faktual yang diambil dari salah satu mata pelajaran saudara:

Pernyataan 1 : _____

Pernyataan 2 : _____

Pernyataan 3 : _____

Pernyataan 4 : _____

Pernyataan 5 : _____

2. Tulislah lima level berbeda dari butir tes pilihan ganda yang didasarkan pada pernyataan di atas.

Butir 1 : Level

Butir 2 : Level

Butir 3 : Level

Butir 4 : Level

Butir 5 : Level

Kelompok A: Tes Benar – Salah

Petunjuk : Bacalah setiap butir tes benar–salah berikut ini. Jika butir melanggar satu atau lebih dari aturan/petunjuk penulisan butir tes benar-salah, beri tanda silang pada huruf M dan tulislah aturan spesifik mana yang dilanggar. Jika butir tidak melanggar aturan, beri tanda silang pada huruf OK dan tulis aturan mana yang tidak dilanggar.

3. Kecepatan suara berbanding lurus dengan kenyaringan M OK
 Alasannya :
4. Semua tumbuh-tumbuhan photosintesis M OK
 Alasannya :
5. Epistimologi merupakan studi tentang sifat pengetahuan M OK
 Alasannya :
6. Presiden Indonesia dipilih melalui pemilu M OK
 Alasannya :

Kelompok B: Menjodohkan (Matching)

Bacalah setiap butir tes menjodohkan berikut ini. Jika butir melanggar satu atau lebih dari aturan/petunjuk penulisan butir tes menjodohkan, beri tanda silang pada huruf M dan tulislah aturan spesifik mana yang dilanggar. Jika butir tidak melanggar aturan, beri tanda silang pada huruf OK dan tulis aturan mana yang tidak dilanggar.

7. Pada garis sebelah kiri kolom 1, tulislah huruf dari data pada kolom II yang tepat. Setiap data pada kolom II dapat digunakan lebih dari satu

Kolom I	Kolom II
..... 1. Kenderaan ruang angkasa pertama mencapai Venus	A. 1969
..... 2. Satelit komunikasi pertama yang ditempatkan di orbit	B. 1962
..... 3. Manusia pertama mendarat di bulan	C. 1961
..... 4. Orang Amerika pertama di ruang angkasa	D. 1960
	E. 1965

Melanggar aturan :

Alasannya :

.....

OK :

Alasannya :

.....

8. Daftar pada kolom II merupakan bagian-bagian struktur mata. Cocokkan bagian mata ini dengan deskripsi yang ada pada kolom I. Bagian-bagian yang dapat digunakan hanya satu.

Kolom I	Kolom II
..... 1. Area yang berisi kerucut (cone)	A. Forvea
..... 2. Daerah pusat mata	B. Iris
..... 3. Jaringan menghubungkan elastis	C. Retina
..... 4. Jaringan sel yang peka rangsangan	D. Sclera
	E. Vitreous humor

Melanggar aturan :

Alasannya :

.....

OK :

Alasannya :

.....

9. Pada Kolom 1 merupakan beberapa ketentuan dasar (undang-undang) yang ditetapkan selama Periode Kolonial (1763 – 1775) di Amerika Serikat , Cocokkan ketentuan ini dengan nama-nama di kolom II. Ingat bahwa nama yang digunakan lebih dari satu.

Kolom I	Kolom I
1. Pajak berita, kertas, kalender	A. U U ketidaktoleransi
2. limpore yang baik	B. U U Navigasi
3. Regulasi perdagangan kolonial	C. U U Perangko
4. Penyelenggaraan pelayaran	D. U U Townshend
5. Hak pemerintahan dari Massachusetts	

Melanggar aturan :

Alasannya :

.....

OK :

Alasannya :

.....

Kelompok C : Tes Pilihan Ganda

Petunjuk : Bacalah setiap butir tes pilihan ganda berikut ini. Jika butir melanggar satu atau lebih dari aturan/petunjuk penulisan butir tes pilihan ganda, beri tanda silang pada huruf M dan tulislah aturan spesifik mana yang dilanggar. Jika butir tidak melanggar aturan, beri tanda silang pada huruf OK dan tulis aturan mana yang tidak dilanggar.

10. Jon F. Kennedy terbunuh M OK
- A. pada tahun 1961
- B. pada tahun 1968
- C. pada tahun 1963 *
- D. pada tahun 1965
- OK, alasannya:
-
- Melanggar, alasannya :
-
11. Menghisap ibu jari mungkin menghasilkan M OK
- trauma yang psikologis terbesar dari
- A. 0 – 2 tahun
- B. 3 – 5 tahun
- C. 6 – 12 tahun *
- D. 13 – 20 tahun
- OK, alasannya:
-
- Melanggar, alasannya :
-
12. Suatu contoh dari binatang yang punah adalah: M OK
- A. sloth
- B. eskimo curlew *
- C. pockect gropher
- D. wombat
- OK, alasannya:

.....
Melanggar, alasannya :
.....

13. Afrika Selatan unggul di dunia dalam pertambangan dari M OK
- A. bauksit
 - B. intan *
 - C. biji besi
 - D. timah

OK, alasannya:
.....
Melanggar, alasannya :
.....

Tanda * merupakan kunci jawaban untuk tes pilihan ganda.

DAFTAR PUSTAKA

- Aiken, Lewis R. 1994. *Psychological Testing and Assessment*, (Eight Edition), Boston: Allyn and Bacon.
- Anastasi, Anne and Urbina, Susana. 1997. *Psychological Testing*. (Seventh Edition). New Jersey: Prentice-Hall, Inc.
- Anderson, Lorin. W. 2003. *Classroom Assessment*. London: Lawrence Erlbaum Associates, Inc.
- Benson, J 1987. *Detecting Item Bias in Affective Scales, Educational and Psychological Measurement*, 47, 55 – 67.
- Cohen, Ronald Jay; Swerdlik, Mark E. and Smith, Douglas K. 1992. *Psychological Testing and Assessment: An Introduction to Test and Measurement*. California: Mayfield Publishing Company.
- Crocker, L. & Algina, J. 1986. *Introduction to Classical and Modern Test Theory*. New York: Holt, Rinehart and Winston, Inc.
- Cunningham, George K. 1998. *Assessment in the Classroom, Constructing and Interpreting Test*. London: The Falmer Press.
- Ebel, Robert L. and Frisbie, David A. 1991. *Essentials of Education Measurement*. New Jersey: Prentice Hall.
- Gregory, Robert J. 2000. *Psychological Testing*. Boston: Allyn and Bacon.
- Gronlund, Norman E. 1993. *How to Make Achievement Test and Measurement*. Massachusetts: Allyn and Bacon.
- Hieronymous, A.M, Lindquist, E.F, dan Hoover, H.D. 1982. *Iowa Test of Basic Skills: Manual for School Administrators*. Chicago: Riverside Publishing

- Kobiszyn, Tom and Gary Borich. 2003. *Educational Testing and Measurement*. United States of America: John Wiley & Sons, Inc.
- McBeath, Ron J. 1992. *Instructing and Evaluating in Higher Education*. New Jersey: Englewood Cliff.
- Nitko, Anthony J. 1996. *Educational Assessment of Students*. New Jersey: Englewood Cliff.
- Oriondo, Leonara Loyola. 1989. *Evaluating Educational Outcomes*. Manila: Rex Printing Company, Inc.
- Osterlind, S.J. 1983. *Test Item Bias*. Beverly Hills, CA: Sage Publication.
- Popham, W. James. 1995. *Classroom Assessment, What Teacher Need to Know*. Boston: Allyn & Bacon.
- Wiersma, William and Stephen G. Jurs. 1990. *Educational Measurement and Test*. Boston: Allyn and Bacon.

Lampiran 1.

Kode Praktik Tes dalam Pendidikan

Kode Praktik yang Adil dalam Pendidikan telah dikembangkan oleh *Joint Committee on Testing Practices*. Komite ini inisiatif *American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education*.

Kode untuk pengembang dan pemakai tes pendidikan yang standar mencakup empat bidang yaitu: pengembang/seleksi tes yang sesuai, interpretasi skor, berusaha untuk adil (*striving for fairness*), informasi pengambil tes (Nitko, 1996 : hal 432-4340).

A. Pengembang/Pemilihan Tes

<p>Pengembang tes harus memiliki informasi tentang pemakai tes membutuhkan tes yang sesuai</p>	<p>Pemakai tes harus menseleksi tes yang memenuhi tujuan mereka dan sesuai dengan populasi pengambil tes.</p>
<p>Pengembang tes harus:</p> <ol style="list-style-type: none"> 1. Mendefinisikan setiap tes dan penggunaannya. Menguraikan populasi untuk tes yang sesuai dipakai. 2. Secara akurat menghadirkan karakteristik, kegunaan, dan pembatasan tes untuk tujuan yang diharapkan pemakai 3. Menjelaskan konsep pengukuran dan secara rinci (detail) yang sesuai dengan para pemakai harapan. 4. Uraikan proses pengembangan tes. Jelaskan konten dan keterampilan yang dites. 5. Memiliki bukti tes memenuhi tujuan yang diharapkan. 	<p>Pemakai tes harus:</p> <ol style="list-style-type: none"> 1. Mendefinisikan tujuan tes dan populasi tes. Memilih tes yang sesuai dengan tujuan dan populasi didasarkan pada informasi yang tersedia. 2. Menyelidiki sumber informasi berpotensi bermanfaat, sebagai tambahan terhadap skor tes, dan informasi yang disajikan oleh tes. 3. Baca materi yang dimiliki oleh pengembang tes dan hindari pemakaian tes yang tidak jelas dan informasi yang tidak lengkap. 4. Mengenal bagaimana dan kapan tes dikembangkan dan uji cobakan. 5. Baca evaluasi dari tes dan alternatif yang mungkin dalam

<p>6. Menyediakan contoh yang mewakili atau melengkapi salinan (<i>copies</i>) tes, pertanyaan, petunjuk, lembar jawaban, manual, dan laporan skor untuk para pemakai.</p> <p>7. Menandai bukti yang diperoleh mengenai kepatutan dari setiap test untuk kelompok dari ras yang berbeda, etnis, atau latar belakang ilmu bahasa (linguistik) yang dites.</p> <p>8. Mengidentifikasi dan mempublikasikan suatu ketrampilan khusus yang diperlukan untuk mengurus setiap tes dan untuk menginterpretasikan skor dengan tepat.</p>	<p>mengukur. Lihat bukti untuk mendukung klaim pengembang tes.</p> <p>6. Meneliti set tes, contoh pertanyaan, petunjuk, lembar jawaban, manual, laporan skor sebelum memilih suatu perjanjian</p> <p>7. Memastikan konten tes dan norma-norma kelompok atau perbandingan kelompok sesuai dengan pengambil tes yang diharapkan.</p> <p>8. Memilih dan menggunakan tes untuk keterampilan yang dibutuhkan dan menginterpretasikan skor secara benar tersedia.</p>
---	---

B. Interpretasi Skor

<p>Pengembang tes harus membantu menginterpretasikan skor secara benar.</p>	<p>Pemakai tes harus menginterpretasikan skor secara benar.</p>
<p>Pengembang tes harus:</p> <p>9. Menyediakan tepat waktu dan dengan mudah dipahami laporan skor yang menguraikan kinerja tes secara jelas akurat. Juga menjelaskan arti dan keterbatasan skor yang dilaporkan.</p> <p>10. Menjelaskan populasi yang representatif dengan norma atau perbandingan grup, tanggal data dikumpulkan, dan proses seleksi sampel dari pengambil tes.</p> <p>11. Mengingatkan para pemakai untuk menghindari antisipasi</p>	<p>Pemakai tes harus:</p> <p>9. Mengandung informasi tentang skala yang digunakan untuk menyekor, karakteristik norma atau perbandingan kelompok, dan keterbatasan skor.</p> <p>10. Menginterpretasikan skor sesuai dengan norma atau perbandingan kelompok dan hasil yang diperoleh pengambil tes. Juga menghitung perbedaan pada praktik administrasi tes atau keakraban dengan pertanyaan spesifik dalam tes.</p> <p>11. Menghindari pemakaian tes untuk tujuan yang tidak direkomendasi-</p>

penyalahgunaan skor tes.

12. Memiliki informasi yang membantu pemakai mengikuti prosedur untuk menset skor yang lulus.
13. Memiliki informasi yang membantu pemakai memperoleh bukti yang menunjukkan tes memenuhi tujuan yang diharapkan.

kan oleh pengembang tes karena sedikitnya bukti yang ada untuk mendukung pemakaian yang diharapkan.

12. Menjelaskan skor yang lulus dan mengumpulkan bukti untuk mendukung kepastian skor tersebut.
13. Mengandung bukti yang menunjukkan bahwa tes memenuhi tujuan yang diharapkan.

C. Berusaha untuk Adil

Pengembang tes harus berusaha keras untuk membuat seadil mungkin untuk pengambil tes yang berbeda ras, jenis kelamin, latar belakang etnis, atau kondisi yang menghalangi

Pemakai tes harus memilih tes yang telah dikembangkan dikembangkan dengan cara-cara dan berusaha seadil mungkin untuk pengambil test yang berbeda ras, jenis kelamin, latar belakang etnis, atau kondisi yang menghalangi.

Pengembang tes harus:

14. Revisi dan meninjau ulang pertanyaan tes dan materi yang terkait untuk menghindari konten dan bahasa yang tidak sensitif.
15. Menyelidiki kinerja pengambil tes dari ras yang berbeda, jenis kelamin, dan latar belakang etnis apabila ukuran sampel cukup tersedia. Menetapkan prosedur yang membantu untuk menjamin perbedaan kinerja terutama oleh ketrampilan bukannya dari faktor yang tidak relevan.
16. Bila mungkin memodifikasi bentuk tes atau memeriksa prosedur administrasi yang tersedia untuk pengambil tes oleh kondisi yang menghalangi. Mengingat pemakai tes dari masalah potensial di dalam penggunaan norma baku dengan modifikasi tes atau memeriksa prosedur

Pemakai tes harus:

14. Mengevaluasi prosedur yang telah digunakan pengembang tes untuk menghindari konten dan bahasa yang tidak sensitif.
15. Meninjau ulang kinerja pengambil tes dari yang ras yang berbeda, jenis kelamin, latar belakang etnis apabila ukuran sampel cukup tersedia. Evaluasi perbedaan kinerja yang disebabkan oleh karakteristik tes.
16. Bila mungkin dan perlu, memodifikasi bentuk tes yang sesuai atau memeriksa prosedur untuk pengambil tes dengan kondisi yang menghalangi. Menginterpretasikan norma-norma baku dengan penuh perhatian dari modifikasi yang telah dibuat.

yang mengakibatkan skor menjadi tidak dapat diperbandingkan.	
--	--

D. Informasi Pengambil Tes

Dalam beberapa keadaan, pengembang tes mempunyai komunikasi langsung dengan pengambil tes. Dalam keadaan lain, pemakai tes berkomunikasi secara langsung dengan pengambil tes. Kelompok mana saja berkomunikasi secara langsung dengan pengambil tes perlu menyediakan informasi yang diuraikan di bawah.

17. Kapan suatu tes optimal, menyediakan pengambil tes atau orang tua/wali mereka sejumlah informasi untuk membantu mereka menilai tes harus diambil, atau jika tersedia suatu alternatif pada tes yang seharusnya digunakan.

18. Menyediakan informasi bagi pengambil tes yang mereka butuhkan menyangkut tes, jenis-jenis dari format pertanyaan, petunjuk, strategi melaksanakan tes yang sesuai. Berusaha keras agar informasi yang sama tersedia untuk semua pengambil tes.

Dalam beberapa keadaan, pengembang tes mengontrol langsung tes dan skor tes. Dalam keadaan lain, pemakai tes juga ikut mengontrol. Kelompok mana saja dapat mengontrol langsung tes dan skor tes harus mengambil langkah-langkah yang diuraikan di bawah ini.

19. Menyediakan pengambil tes atau orang tua/wali mereka informasi tentang tentang pengambil test benar yang ingin memperoleh salinan tes dan lembar jawaban yang lengkap, melakukan tes ulang, memiliki kembali skor tes (*test rescored*), atau skor yang ditunda.

20. Menceritakan pengambil tes atau orang tua/wali mereka berapa lama skor akan disimpan pada file dan menunjukkan kepada siapa dan di dalam keadaan apa skor tes akan dilepaskan atau tidak akan dilepaskan.

21. Menguraikan prosedur bagi pengambil tes atau orang tua/wali mereka untuk mendaftarkan keluhan dan masalah yang ingin diselesaikan