PENGEMBANGAN DAN PENYUSUNAN TES OBYEKTIF

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

A. Tes Benar – Salah

1. Pengertian

Tes benar-salah adalah 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 pertanytaan-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 soaalnya lebih mudah daripada tes pilihan berganda
e. Mencakup bahan yang luas dan tidak memakan tempat karena biasanya pertanyaan-perytanyaan singkat saja
f. Dapat dipakai berulang-ulang.

Kelemahan dari tes benar-salah

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

Perintah: lingkarilah huruf B jika pertanyaan di bawah ini benar danhuruf S jika pertanyaan itu salah!

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

b. Soal benar - salah menuntut alasan

Lingkarilah huruf B jika peranyataan di bawah ini betul, dan lingkarilah huruf 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 bertujuanuntuk kebebasan berusaha
3) B–S P4 merupakan kepribadian / jiwa bangsa Indonesia

c. Tes benar - salah dengan membetulkan

Lingkarilah huruf B jika pernyataan di bawah ini betul, dan lingkarilah huruf S dan tunjukkan bagian mana yang salah dengan cara memberi garis di bawahnya serta kemudiaan tulislah pembetulannya 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

Lingkarilah huruf B jika peranyataan di bawah ini betul, dan lingkarilahhuruf 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 danmempertinggi 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/jodohyang 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 adatapi sulit sekali
c. Obyektif
d. Mudah disusun
e. Cocok untuk mengukur informasi-informasi yang berbentuk fakta darisebuah pengertian hubungan penegrtian atau konsep-konsep

Kelemahan tes menjodohkan

Kelemahan dari bentuk soal tes ini adalah sukar untuk mengukur prosesmental 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 lebihkhusus, sempit, dan terbatas.
c. Jawaban yang dipilih sesuai dengan persoalan atau menjadi jodoh /pasangan yang sesuai.

4. Petunjuk penyusunan adalah :

a. Buatlah pengantar sejelas mungkin
b. Bentuklah tes terdiri dari sederetan persoalan dan sederetan jawaban.
c. Hal yang disusun baik dalam pertanyaan maupun jawaban hendaknyahomogen
d. Jumlah jawaban buatlah lebih banyak dari jumlah pertanyaan
e. Batasi setiap kelompok jangan lebih dari 10 pertanyaan jika inginbanyak buatlah beberapa kelompok.
f. Semua pertanyaan dan jawaban hendaknya dibuat dalam satuhalaman saja
g. Setiap satu pertanyaan hendaknya hanya ada satu jawaban yangbenar.
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 komplek. 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 komplek, 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 yangdiukur.

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 padaketersediaan 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 komplek. 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 mengalirakan 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 kategoriApa 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 daratanlebih 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 kotakberikut.

Tabel 3.2 Ilustrasi Pertanyaan Pemahaman dan Aplikasi

Pertanyaan Pemahaman
Mana dari berikut ini sebuah contoh dari_____?
Apa yang merupakan pikiran utama yang diungkap oleh_____?
Apa perbedaan utama antara_____?
Apa karakteristik umum dari_____?
Mana dari berikut ini bentuk lain dari_____?
Mana dari berikut ini penjelasan terbaik_____?
Mana dari berikut ini ringkasan terbaik_____?
Mana dari berikut ini ilustrasi terbaik_____?
Apa yang anda prediksi akan terjadi jika_____?
Apa kecenderungan anda memprediksi pada_____?

Pertanyaan Aplikasi
Mana dari berikut ini metode terbaik untuk ?
Apa langkah-langkah yang harus diikuti dalam aplikasi ?
Mana situasi yang membutuhkan penggunaan dari_____?
Mana prinsip yang terbaik untuk menyelesaikan_____?
Apa prosedur terbaik untuk memperbaiki_____?
Apa prosedur terbaik untuk mengkonstruksi_____?
Apa prosedur terbaik untuk mengoreksi
Mana dari berikut ini perencanaan terbaik untuk ?
Mana dari berikut ini yang menyediakan urutan sesuai untuk ?
Apa yang merupakan efek yang paling mungkin_____?

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 menggunakanya 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 x 10^{-19} C dan bermassa 6,4 x 10^{-27} kg ditempatkan dari keadaan diam dalam suatu medan listrik homogen 4,0 x 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 x 10^{10} m/detik
B. 2,0 x 10^{10} m/detik
C. 3,0 x 10^{10} m/detik
D. 4,0 x 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 aplikasidari 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 menjadi125 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) gayagravitasi bulan dengan matahari mempengaruhi pasang surut air laut, (2) pasang naik tertinggi terjadi pada saat bulan purnama, (3) pasang surut terendah terjadipada saat kuartil awal atau akhir, (4) matahari lebih mudah menarik air laut dari padabulan. 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 selembar 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 komplek 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 dankemampuan 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 yangberfungsi 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 butirsoal (*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 yangstabil.
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 nonlogam?

A. 1

B. 2

C. 3

D. 4 *

Demikian pula, butir soal 3 adalah butir soal pilihan ganda yang kurang baikditulis. 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 daripada 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 umumpada *stem*. Perhatikan contoh berikut ini!

Contoh

Butir soal yang kurang baik:

Tujuan pembelajaran bermanfaat untuk mengkonstruksi tes ketika tujuandinyatakan 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 siswamaksimum.
D. Jenis aktivitas belajar selama belajar.

Butir soal yang lebih baik:

Tujuan pembelajaran paling bermanfaat untuk tujuan mengkonstruksi tesketika 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 harusterjadi 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 menyenang tidak pasti tanpa:

A. Temperatur diatur pada 18 ^{0}C, kelembaban relatif 50 %.

B. Temperatur diatur pada 23 ^{0}C, kelembaban relatif 72 %

C. Temperatur diatur pada 20 ^{0}C, kelembaban relatif 25 %

D. Temperatur diatur pada 20 ^{0}C, kelembaban relatif 50 % *

Butir soal yang lebih baik:

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

A. Temperatur diatur pada 18 ^{0}C, kelembaban relatif 50 %.
B. Temperatur diatur pada 23 ^{0}C, kelembaban relatif 72 %
C. Temperatur diatur pada 20 ^{0}C, kelembaban relatif 25 %
D. Temperatur diatur pada 20 ^{0}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 bawahi (*underscore*), atau ditulis tebal (*bold*). Mengganti istilah negatif dengan kata *kecuali*, kadang-kadang dapat meningkatkan kejelasan, sebagai diilustrasikan dalam angka 10. Beberapa siswaakan 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 olehpembakaran...

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

Butir soal yang lebih baik:

Air-jenis pemadam yang cocok untuk memadamkan api yang disebabkanoleh 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 contohberikut!

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 tatabahasa 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 penjodohan
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 essai.
D. boleh mengakibatkan tes kurang relevan dengan tujuan pembelajaran.* Istilah ”boleh” jelas nyata di dalam contoh, tetapi kesalahan jenis ini adalahumum 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 essai.
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 denganjawaban 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 memilikimakna 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* 10dan *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 chronical 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.

l. 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 siswadalam mengerjakan tes. Panjang jawaban benar dapat diatur dengan memvariasikanya. Jawaban benar kadang-kadang lebih panjang, kadang-kadanglebih 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 essaiadalah:

A. untuk pengukuran hasil belajar yang lebih komplek
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* (ukurannilai 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 semuadi 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 bawahkonstitusi pada tahun:

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

Butir soal yang lebih baik:

Pemerintahan Amerika Serikat diproklamirkan pada hakekatnya di bawahkonstitusi 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 lebihhomogen. 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 bai

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. Keenamtipe 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 perserta 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 menunjukkansebab 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 merupakanpenataan dari ketatanegaraan Indonesia.

SEBAB

Masyarakat Indonesia berada dalam masa transisi, karena itu mengalami banyakperubah-an 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. Strukturpertanyaan 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 sesuaipetunjuk 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 akanmemiliki 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 hukumdan 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 Kepemerintahan 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?

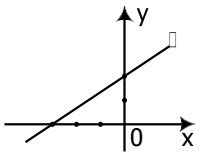
ALTERNATIF JAWABAN		NILAI
A	Bagaimanapun pendapat orang tersebut perlu dipertimbangkan	5
B	Kenyatan tersebut adalah wajar, karena semua peserta rapat mempunyai hak untuk mengeluarkan pendapat/ide	3
C	Orang tersebut akan saya nasehati, agar tidak mengganggu kelancaran rapat	2
D	Pimpinan perlu arif dalam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

Contoh

Perhatikan gambar di bawah!



Persamaan garis g yang melalui O(0, 0) dan tegak lurus pada 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 memilihjawaban 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) jumlahsoal 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 betuldan salah harusnya tidak mengikuti pola yang teratur, (8) sebaiknya pertanyaan tidakdiambil 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 benarlebih 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.

DAFTAR PUSTAKA

Aiken, Lewis R. 1994. *Psychological Testing and Assessment*, (Eight Edition), Boston: Allyn and Bacon.

Anastasi. Anne and Urbina, Susana. 1997. *Psycohololological 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.* NewYork: 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*.
Masschusetts: 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 Yersey: Englewood Cliff.

Nitko, Anthony J. 1996. *Educational Assessment of Students*. New Yersey: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 andTest*. Boston: Allyn and Bacon.