

## **Langkah-langkah dalam penyusunan tes**

Urutan langkah yang dilakukan adalah:

- a. Menentukan tujuan mengadakan tes
- b. Mengadakan pembatasan terhadap bahan yang akan diteskan
- c. Merumuskan tujuan instruksional khusus dari tiap bahan
- d. Menderetkan semua tujuan tersebut
- e. Menuliskan butir soal berdasarkan TIK yang sudah disusun

## **Mengkonstruksi item tes pilihan ganda**

Untuk membuat tes pilihan ganda, aturan penyusunannya adalah sebagai berikut:

1. Pembuatan kalimat (inti kalimat) hendaknya mengandung permasalahan atau bisa dalam bentuk pernyataan
2. Sebaiknya menggunakan kalimat-kalimat positif. Kalimat negative digunakan jika memang terpaksa menggunakan kalimat tersebut
3. Penempatan pilihan yang benar hendaknya di susun secara terpecah
4. Pilihan jawaban dapat dibuat antara 3-5 pilihan
5. Data-data yang tidak relevan hendaknya dihilangkan
6. Hindari kata-kata pada item yang mengandung petunjuk
7. Pemetaan jawaban hendaknya di atas dengan posisi dalam bentuk kolom
8. Semua pilihan jawaban sebaiknya memiliki jumlah kata yang sama
9. Jangan menggunakan tes pilihan ganda jika ada jenis tes lain yang lebih tepat

## **menilai tes yang dibuat sendiri**

ada 4 macam cara untuk menilai tes:

1. Meneliti secara jujur soal-soal yang sudah disusun dengan cara mengajukan pertanyaan
  - a. Apakah banyaknya soal untuk tiap topic sudah seimbang
  - b. Apakah semua soal menanyakan bahan yang telah diajarkan
  - c. Apakah soal-soal itu sudah dimengerti
  - d. Apakah soal itu dapat dikerjakan oleh semua siswa
2. Dengan mengadakan analisi soal. kegunaannya
  - a. Membantu kita dalam mengidentifikasi butir-butir soal yang jelek
  - b. Memperoleh informasi yang akan dapat digunakan untuk menyempurnakan soal-soal untuk kepentingan lebih lanjut
  - c. Memperoleh gambaran secara selintas tentang keadaan yang kita susun
3. Mengadakan checking validitas
4. Mengadakan checking reliabilita

## Penilaian Melalui Tes

### Pengertian

Test is any series of question or exercise or others means of measuring the skill, knowledge, intelligence, capacities of attitudes or an individual or group.

### Bentuk Tes

Cara pengukuran hasil belajar melalui tes dapat dilakukan dengan berbagai bentuk yaitu : dengan cara tes tertulis, tes lisan, dan tes praktek. Tes tertulis dapat digunakan untuk mengukur pengetahuan tentang fakta, pengertian, keterampilan menerapkan prinsip-prinsip dasar untuk memecahkan masalah-masalah nyata dan keterampilan menerapkan ide-ide ke dalam suatu uraian bebas. Tes tertulis ini terdiri ; tes uraian, tes obyektif dan karya tulis ( makalah ). Tes lisan tepat digunakan untuk mengukur kesiapan siswa dalam mengemukakan kembali pengetahuan yang dipelajarinya. Agar tes ini memenuhi fungsinya secara maksimal, maka tes harus dilaksanakan dalam bentuk dialog. Adapun tes praktek adalah tes yang tepat untuk mengukur keterampilan yang bersifat motoris. Materi evaluasinya terdiri atas praktek melaksanakan suatu tugas. Ke tiga bentuk tes tersebut adalah tes yang umum dipakai oleh para guru, namun yang paling dominan penggunaannya adalah tes tertulis.

### Soal Tes Tertulis

Ada 2 model tes tertulis yaitu : tes uraian ( essay type test ) dan tes obyektif ( objective type test ). Tes uraian lebih tepat digunakan untuk mengukur kemampuan murid dalam masalah-masalah yang menuntut kemampuan berpikir tinggi, seperti kecakapan : (1) memecahkan soal, (2) menganalisis, (3) memeperbandingkan, (4) mengevaluasi, (5) menyatakan sebab akibat, (6) menarik kesimpulan dan seterusnya. Sedangkan tes obyektif lebih tepat digunakan untuk mengukur kecakapan murid dalam masalah-masalah yang menuntut kemampuan berpikir tidak terlalu tinggi seperti kecakapan- kecakapan : (a) mengingat kembali fakta, (b) mengenal kembali fakta, (c) memahami prinsip-prinsip, (d) menggunakan prinsip-prinsip, (e) mengasosiasikan antara 2 hal dan seterusnya.

Kedua model tes diatas memiliki kelebihan dan kekurangannya masing-masing. **Kelebihan** tes **uraian** diantaranya adalah ; tidak mudah ditebak, sulit untuk saling mencontoh, dan menuntut siswa untuk menyusun buah pikiran secara rinci. **Kelemahannya** ; sangat subyektive, memeriksa memerlukan waktu lama dan tidak bisa diwakilkan, tidak mencakup bahan yang luas, jawaban bisa terlalu panjang, dan kemungkinan tulisan sukar dibaca yang akan mempengaruhi terhadap pemberian angka. Dalam **tes obyektif** **kelebihannya** yaitu : pemeriksaan dapat lebih cepat, penilaian lebih obyektif, dapat dijawab secara cepat dan mencakup bahasan yang luas. Adapun **kelemahannya** adalah ; penyusunan soal memerlukan waktu yang lama dan sukar digunakan untuk mengukur kecakapan murid berpikir tinggi.

Ada beberapa macam soal bentuk obyektif tes dan masing-masing memiliki ketepatan tersendiri dalam penggunaannya. Dalam **soal benar-salah**, lebih tepat digunakan untuk mengukur

kecakapan mengingat kembali fakta. Pada **soal pilihan ganda** (jamak) , tepat digunakan untuk mengukur kecakapan dalam memahami dan mempergunakan prinsip-prinsip. **Soal menjodohkan** , tepat digunakan untuk mengukur kecakapan dalam mengasosiasikan dua hal. Bentuk **soal isian**, tepat digunakan untuk mengukur kecakapan mengingat kembali fakta. Bentuk **soal jawaban singkat**, tepat digunakan untuk mengukur kecakapan mengingat kembali fakta.

### **Rambu Rambu Pembuatan Soal**

Dalam membuat soal tes uraian maupun tes obyektif ada rambu-rambu yang harus diperhatikan oleh pembuat soal, seperti dalam membuat **soal tes uraian** rambu-rambunya adalah : (1) soal hendaknya dirumuskan secara jelas dan tegas batas-batasnya, (2) dalam setiap soal hendaknya sudah terkandung maksud tentang jawaban yang dikehendaki oleh penyusun soal, (3) kunci jawaban hendaknya dibuat serempak bersamaan dengan penyusunan soal, (4) seluruh bahan ujian hendaknya diolah dan dipadukan. Sedangkan rambu-rambu **soal tes obyektif** masing-masing adalah sebagai berikut : Untuk **Soal benar-salah**, (a) petunjuk pengerjaan dijelaskan terlebih dahulu, (b) setiap soal tidak boleh mengandung kata-kata yang meragukan, (c) hindari kalimat menyingkal karena akan membingungkan siswa, (d) pergunakan kalimat tunggal yang pendek saja dan (e) jangan membuat soal yang masih dipersoalkan. Untuk **Soal pilihan ganda/jamak**, (1) pernyataan masalah harus jelas-jelas mempersoalkan suatu masalah, (2) pada setiap soal hanya mempunyai satu jawaban yang betul, (3) perumusan masalah hendaknya merupakan kalimat tidak lengkap yang dapat dilengkapi oleh pilihan jawaban, (4) pilihan jawaban hendaknya meninggi. Pada **soal menjodohkan** rambu-rambu yang perlu diperhatikan adalah; jumlah jodoh/pasangan hendaknya lebih besar dari pada terjodoh, dan jumlah soal tidak terlalu banyak. Dalam **soal isian**, (1) Bagian kalimat yang diganti dengan titik-titik hendaknya merupakan bagian yang terpenting, (2) pergunakan kalimat yang mudah dipahami, (3) panjang titik-titik hendaknya sebanding dengan panjang isian, (4) bagian kalimat yang diganti dengan titik-titik hendaknya ditempatkan dibagian tengah atau akhir kalimat. Sedangkan dalam pembuatan **soal jawaban singkat hendaknya**; (1) jawaban tidak memerlukan uraian panjang-lebar, (2) setiap soal hendaknya di jawab secara mutlak, dan (3) susunan kalimat hendaknya sederhana dan jelas. Itulah beberapa hal yang perlu diperhatikan ketika akan membuat soal obyektif tes.

### **Ciri-Ciri Tes Yang Baik**

Soal yang akan ditekankan kepada peserta didik dapat dikatakan baik apabila memiliki persyaratan : validitas, reliabilitas, obyektivitas, praktibilitas dan ekonomis.

**Validitas**, validitas dalam bahasa kita dapat diartikan kesahihan. Sebuah tes dapat dikatakan tepat apabila memiliki tingkat validitas tinggi, atau dapat tepat mengukur apa yang diukur.

**Reliabilitas**, realibilitas diartikan dapat dipercaya. Suharsimi mengatakan suatu tes dapat dipercaya jika memberikan hasil yang tetap apabila ditekankan berkali kali. Sebuah tes dikatakan reliable apabila hasil tes tersebut menunjukkan ketepatan. Jika kepada siswa diberikan tes yang sama pada waktu yang berlainan, maka setiap siswa akan berada dalam urutan yang sama dalam kelompoknya.

**Obyektivitas**, sebuah tes dikatakan obyektif apabila dalam melaksanakan tes tidak ada factor subyektif, yang memengaruhi terutama pada system sekoringnya.

**Praktibilitas**, tes yang bersifat praktis adalah tes yang mudah pengadministrasiannya. Menurut Suharsimi tes yang praktis adalah; (1) mudah dilaksanakan (tidak menuntut peralatan yang banyak, (2) mudah pemeriksaanya (tes dilengkapi dengan pedoman skoringnya) dan (3) dilengkapi petunjuk yang jelas.

**Ekonomis**, suatu tes dikatakan ekonomis apabila tes dilaksanakan tidak memerlukan biaya yang mahal, tenaga yang banyak, dan waktu yang lama.

### **Analisis Butir Soal**

Semua guru dalam melaksanakan tugasnya akan senantiasa membuat soal /tes untuk menguji kompetensi peserta didiknya. Dari keseluruhan soal yang dibuatnya mungkin saja terdapat soal yang terlalu mudah, soal yang sedang, bahkan mungkin soal soalnya terlalu sulit untuk dikerjakan oleh siswa. Menurut Suharsimi, soal yang baik adalah soal yang tidak terlalu mudah dan juga tidak terlalu sukar.

Kapan sebuah soal dikatakan baik ?, untuk memberikan jawaban terhadap pertanyaan ini, Suharsimi mengatakan “ perlu menerangkan 3 masalah yang berhubungan dengan analisis soal yaitu : taraf kesukaran, daya pembeda dan pola jawaban soal”.

Taraf kesukaran, Bilangan yang menunjukkan sukar dan mudahnya sesuatu soal disebut indek kesukaran. Indek kesukaran soal diberi simbol P. Besarnya indeks kesukaran antar 0.00 – 1.0 . Soal dengan indeks kesukaran 0.00 , menunjukkan bahwa soal itu terlalu sukar, sebaliknya apabila indeks kesukarannya 1.00 menunjukkan bahwa soal itu terlalu mudah. Untuk menghitung besaran indeks kesukaran soal dapat digunakan rumus sebagai berikut :

Rumus

$$P = \frac{B}{JS}$$

Keterangan :

P = Indeks kesukaran

B = Banyaknya siswa yang menjawab soal dengan betul

Js = Jumlah seluruh siswa peserta tes

Di dalam pelaksanaan pengerjaan analisis butir soal, jawaban benar diberi nilai “ 1 “, dan untuk jawaban salah diberi nilai “ 0 “. Sedangkan kriteria untuk mengklasifikasikan indeks kesukarannya adalah sebagai berikut : soal dengan nilai P = 0.00 – 0.30 adalah soal sukar, P = 0.30 – 0.70 adalah soal sedang dan soal dengan nilai P = 0.70 – 1;00 adalah soal mudah. Contoh pengerjaanya dapat dilihat pada table dihalaman berikut.

Halaman 7

Jika table tersebut adalah perhitungan/skorings hasil ulangan mata pelajaran geografi di salah satu S.M.A. , maka masing-masing nilai P nya adalah sebagai berikut :

Soal no.1. B = 10, S = 10 , Nilai P = 0.50

2. B = 14, S = 6 , Nilai P = 0.70

3. B = 4 , S = 16, Nilai P = 0.20
4. B = 9 , S = 9 , Nilai P = 0.45
5. B = 15, S = 5 , Nilai P = 0.75
6. B = 6 , S = 6 , Nilai P = 0.30
7. B = 18, S = 2 , Nilai P = 0.90
8. B = 17, S = 3 , Nilai P = 0.85
9. B = 3 , S = 17, Nilai P = 0.15
10. B = 11 , S = 9 , Nilai P = 0.55
11. B = 10, S = 10, Nilai P = 0.50
12. B = 18, S = 2 , Nilai P = 0.90
13. B = 20, S = 0 , Nilai P = 1.00
14. B = 10, S = 10, Nilai P = 0.50
15. B = 9 , S = 11, Nilai P = 0.45
16. B = 7 , S = 13, Nilai P = 0.35
17. B = 10, S = 10, Nilai P = 0.50
18. B = 14, S = 6 , Nilai P = 0.70
19. B = 13, S = 7 , Nilai P = 0.65
20. B = 13, S = 7 , Nilai P = 0.65

#### Daya Pembeda

Daya pembeda soal, adalah kemampuan ai (berkemampuan tinggi) dengan siswa yang bodoh (berkemampuan rendah).

Angka yang menunjukkan besarnya daya pembeda disebut indeks diskriminasi, disingkat D (d besar). Seperti halnya indeksnya kesukaran, indeks diskriminasi (daya pembeda) ini berkisar antara 0, 00 sampai 1, 00. Hanya bedanya, indeks kesukaran tidak mengenal tanda negative (-), tetapi pada indeks diskriminasi ada tanda negatif. Tanda negatif pada indeks diskriminasi digunakan jika sesuatu soal "Terbalik" menunjukkan kualitas testee. Yaitu anak pandai disebut bodoh dan anak bodoh disebut pandai.

Dengan demikian ada tiga titik pada daya pembeda, yaitu :

- 1,00 ←————→ 0,00 ←————→ 1,00

Daya Pembeda negatif

daya pembeda rendah

Daya pembeda tinggi (positif)

Bagi suatu soal yang dapat dijawab benar oleh siswa pandai maupun siswa bodoh, maka soal itu tidak baik karena tidak mempunyai daya pembeda. Demikian pula jika semua siswa baik pandai maupun bodoh tidak dapat menjawab dengan benar. Soal tersebut tidak baik juga karena tidak mempunyai daya pembeda. Soal yang baik adalah soal yang dapat dijawab benar oleh siswa-siswa yang pandai saja.

Seluruh pengikut tes dikelompokkan menjadi 2 kelompok, yaitu kelompok pandai atau kelompok atas (*upper group*) dan kelompok bodoh atau kelompok bawah (*lower group*).

Jika seluruh kelompok atas dapat menjawab soal tersebut dengan benar, sedang seluruh kelompok bawah menjawab salah, maka soal tersebut mempunyai D paling besar, yaitu 1,00. Sebaliknya jika semua kelompok atas menjawab salah, tetapi semua kelompok bawah menjawab betul, maka nilai D-nya -1,00. Tetapi jika siswa kelompok atas dan siswa kelompok bawah sama-sama menjawab benar atau Sama-sama menjawab salah, maka soal tersebut mempunyai nilai D 0,00. karena tidak mempunyai daya pembeda sama sekali.

### Cara menentukan daya pembeda (nilai D)

Untuk ini perlu dibedakan antara kelompok kecil (kurang dari 100 orang) dan kelompok besar (100 orang ke atas).

a. Untuk kelompok kecil

Seluruh kelompok testee dibagi dua sama besar, 50% kelompok atas dan 50% kelompok bawah.

**Contoh:**

Siswa	Skor	
A	9	Kelompok Atas ( $J_A$ )
B	8	
C	7	
D	7	
E	6	
F	5	Kelompok Bawah ( $J_B$ )
G	5	
H	4	
I	4	
J	3	

Seluruh pengikut tes. Dideretkan mulai dari skor teratas sampai terbawah, lalu dibagi 2.

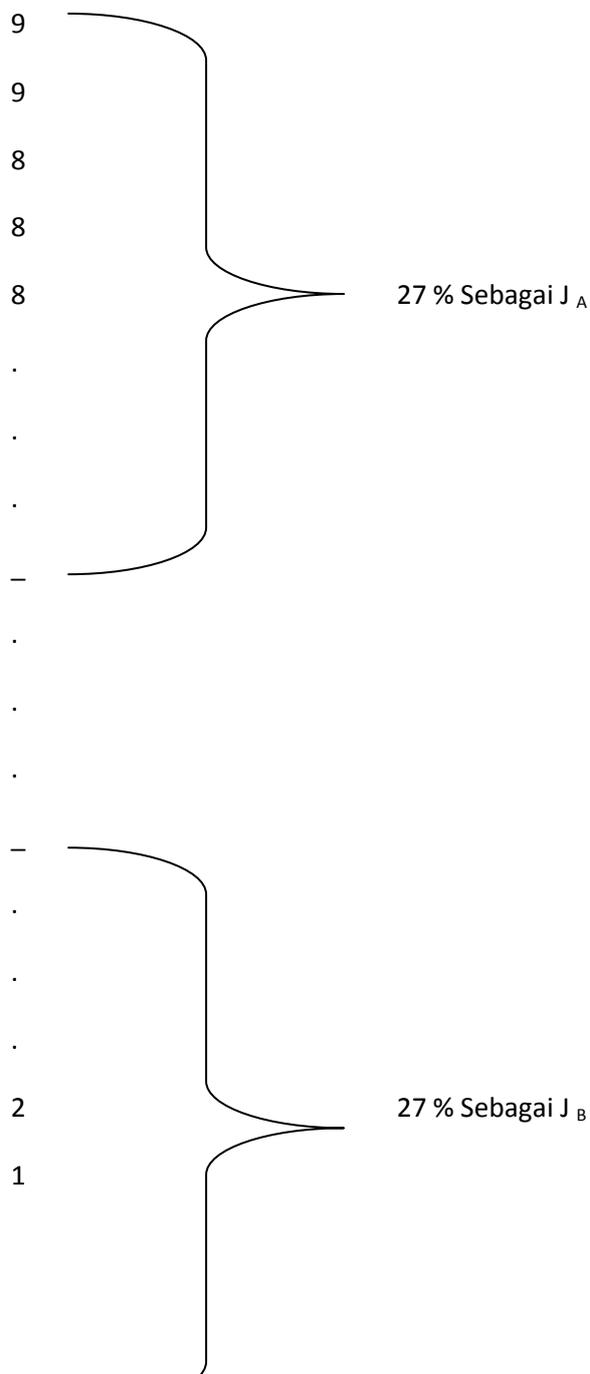
b. Untuk kelompok besar

Mengingat biaya dan waktu analisis, maka untuk kelompok besar biasanya hanya diambil kedua kutubnya saja, yaitu 27% skor teratas sebagai kelompok atas ( $J_A$ ) dan 27% skor terbawah sebagai kelompok bawah ( $J_B$ )

$J_A$  = jumlah kelompok atas

$J_B$  = jumlah kelompok bawah

Contoh:



1

1

0

### Rumus mencari D

Rumus untuk menentukan indeks diskriminasi adalah:

$$D = B_A/J_A - B_B/J_B = P_A - P_B$$

Dimana:

J = jumlah peserta tes

$J_A$  = banyaknya peserta kelompok atas

$J_B$  = banyaknya peserta kelompok bawah

$B_A$  = banyaknya peserta kelompok atas yang menjawab soal itu dengan benar

$B_B$  = banyaknya peserta kelompok bawah yang menjawab soal itu dengan benar

$P_A = B_A/J_A$  = proporsi peserta kelompok atas yang menjawab benar (ingat,  $P$  sebagai indeks kesukaran)

$P_B = B_B/J_B$  = proporsi peserta kelompok bawah yang menjawab benar

### ***Contoh perhitungan***

Dari hasil analisis tes yang terdiri dari 10 butir soal yang dikerjakan oleh 20 orang siswa, terdapat dalam table sebagai berikut:

siswa	kelompok	nilai soal										skor siswa
		1	2	3	4	5	6	7	8	9	10	
A	B	1	0	1	0	0	0	1	1	1	0	5
B	A	0	1	1	1	1	1	0	0	1	1	7
C	A	1	0	1	0	1	1	1	1	1	1	8
D	B	0	0	1	0	0	1	1	1	1	0	5
E	A	1	1	1	1	1	1	1	1	1	1	10
F	B	1	1	0	0	0	1	1	1	1	0	6
G	B	0	1	0	0	0	1	1	1	1	1	6
H	B	0	1	1	0	0	1	0	1	1	1	6
I	A	1	1	1	0	0	1	1	1	1	1	8
J	A	1	1	1	1	0	0	1	0	1	1	7
K	A	1	1	1	0	0	1	1	1	1	0	7
L	B	0	1	0	1	1	0	0	1	1	0	5
M	B	0	1	0	0	0	0	0	1	1	0	3
N	A	0	0	1	0	1	1	1	1	1	1	7
O	A	1	1	0	1	1	1	1	1	1	1	9
P	B	0	1	0	0	0	1	0	0	1	0	3
Q	A	1	1	0	1	0	1	1	1	1	1	8
R	A	1	1	1	1	0	1	1	1	1	0	8
S	B	1	0	1	0	0	1	1	1	1	0	6
T	B	0	1	0	1	0	1	1	1	1	0	6
jumlah		11	15	12	8	6	16	15	17	20	10	

Berdasarkan nama-nama siswa dapat kita peroleh skor-skor sebagai berikut:

A = 5

F = 6

K = 7

P = 3

B = 7	G = 6	L = 5	Q = 8
C = 8	H = 6	M = 3	R = 8
D = 5	I = 8	N = 7	S = 6
E = 10	J = 7	O = 9	T = 6

Dari angka-angka yang belum teratur kemudian dibuat array (urutan penyebaran), dari skor yang paling tinggi ke skor yang paling rendah.

Kelompok atas	kelompok bawah
10	6
9	6
8	6
8	6
8	6
8	5
7	5
7	5
7	3
7	3
_____	_____
10 orang	10 orang

Array ini sekaligus menunjukkan adanya kelompok atas ( $J_A$ ) dan kelompok bawah ( $J_B$ ) dengan pemilikinya sebagai berikut:

Kelompok atas	kelompok bawah
B = 7	A = 5
C = 8	D = 5
E = 10	F = 6
I = 8	G = 6
J = 7	H = 6
K = 7	L = 5
N = 7	M = 3
O = 9	P = 3

$$Q = 8$$

$$S = 6$$

$$R = 8$$

$$T = 6$$

—————  
10 orang

—————  
10 orang

Perhatikan pada table analisis 10 butir 20 siswa.

Di belakang nama siswa dituliskan huruf A atau B sebagai tanda kelompok. Hal ini untuk mempermudah menentukan  $B_A$  dan  $B_B$ .

$B_A$  = banyaknya siswa yang menjawab benar pada kelompok atas (A)

$B_B$  = banyaknya siswa yang menjawab benar kepada kelompok bawah (B)

Sudah disebutkan diatas bahwa soal yang baik adalah soal yang dapat membedakan antara anak pandai dengan anak bodoh, dilihat dari dapat dan tidaknya mengerjakan soal itu.

Marilah kita perhatikan tabelk analisis lagi, khusus butir soal nomor 1.

- Dari kelompok atas yang menjawab betul 8 orang
- Dari kelompok bawah yang menjawab betul 3 orang

kita terapkan dalam rumus indeks doskriminasi:

$$\left. \begin{array}{l} J_A = 10 \\ \\ B_A = 8 \end{array} \right\} P_A = 0,8$$

$$\left. \begin{array}{l} J_B = 10 \\ \\ B_B = 3 \end{array} \right\} P_B = 0,3$$

$$\text{Maka } D = P_A - P_B$$

$$= 0,8 - 0,3$$

$$= 0,5$$

Dengan demikian maka indeks diskriminasi untuk soal nomor 1 adalah 0,5.

Sekarang kita perhatikan butir soal nomor 8:

$$\begin{array}{l}
 J_A = 10 \\
 \left. \begin{array}{l} \\ \\ \\ \end{array} \right\} P_A = 0,8 \\
 B_A = 8 \\
 \\
 J_B = 00 \\
 \\
 P_B = 0,9 \\
 \\
 B_B = 9
 \end{array}
 \qquad
 \begin{array}{l}
 \text{maka } D = P_A - P_B \\
 \\
 = 0,8 - 0,9 \\
 \\
 = -0,1
 \end{array}$$

Butir soal ini jelek karena lebih banyak dijawab benar oleh kelompok bawah dibandingkan dengan jawaban benar dari kelompok atas. Ini berarti bahwa untuk menjawab soal dengan benar, dapat dilakukan dengan menebak.

*Butir-butir soal yang baik adalah butir-butir*

*soal yang mempunyai indeks diskriminasi 0,4 sampai 0,7*

klasifikasi daya pembeda:

D : 0,00 – 0,20 : jelek (poor)

D : 0,20 – 0,40 : cukup (satisfactory)

D : 0,40 – 0,70 : baik (good)

D : 0,70 – 1,00 : baik sekali (excellent)

D : negative, semuanya tidak baik, jadi semua butir soal yang mempunyai nilai D negatif sebaiknya dibuang saja.

Menskor dan menilai

Menskor = memberi angka.

Alat bantuannya:

1. Kunci jawaban

2. Kunci skoring
3. Pedoman penilaian

Dalam menentukan angka (skor) pada soal pilihan ganda, benar atau salah dapat dilakukan dengan dua cara:

1. Dengan hukuman/denda (karena diragukan adanya tebakan)
2. Tanpa hukuman/denda

Dengan hukuman, rumusnya:

$$S = R - w/(0-1)$$

Keterangan:

S = skor; R = jawaban betul; w = wrong/salah; 0 = banyaknya pilihan jawaban; 1 = bilangan tetap

Tanpa hukuman rumusnya:

$$S = R$$

Skor = jawaban yang benar

Skor = adalah hasil pekerjaan menskor yang diperoleh dengan menjumlahkan angka bagi setiap soal tes yang dijawab betul.

Nilai = angka ubahan dari skor dengan acuan tertentu, acuan norma atau standar

Contoh: ada 20 soal tes plihan ganda dijawab oleh si A = 16 soal dijawab benar, maka skornya 16. Berapa nilainya?

$$16/20 \times 10 = 8$$

Soal essay, ada 3 pertanyaan, setiap jawaban diberi:

1. Bobot = 15 → A = 10                      \*skor = 30
2. Bobot = 20 → A = 15                      nilai = 30/40 x 10 = 7,5
3. Bobot = 5 → A = 5