

# **EVALUASI PENDIDIKAN**

**Oleh**

**Drs. Andi Suntoda S., M.Pd.**

**Fakultas Pendidikan Olahraga dan Kesehatan  
Universitas Pendidikan Indonesia  
Bandung, 2010**

# Pengantar

Tes, Pengukuran, dan Evaluasi merupakan bagian yang tak terpisahkan dalam berbagai kegiatan manusia, demikian pula halnya dalam kegiatan pengajaran.

Kegiatan pembelajaran merupakan sebuah proses yang dinamis, seorang pengajar/guru dan pembina menghadapi berbagai permasalahan yang membutuhkan pemecahan. Semakin teliti informasi yang diperoleh (melalui tes dan pengukuran) akan semakin baik keputusan yang diambil. Dengan melaksanakan ketiga hal tersebut (tes, pengukuran, dan evaluasi) kita dapat mengetahui perkembangan dan kekurangan, sehingga akhirnya dapat membuat suatu keputusan yang tepat.

## Pengertian Tes, Pengukuran, dan Evaluasi

- TES : Sebuah instrumen yang dipakai untuk memperoleh informasi tentang seseorang atau objek (Observasi, wawancara, angket, tes skill, atau bentuk lain yang sesuai).
- PENGUKURAN : Proses pengumpulan data / informasi dari suatu obyek tertentu. ( skor, frekuensi, waktu, jarak) tinggi badan : 179 cm, 165 cm dsb..  
Data bersifat kuantitatif,
- EVALUASI :
  - ✓ Proses penentuan nilai atau kelayakan data yang terhimpun.
  - ✓ Proses penilaian secara kualitatif data yang telah diperoleh melalui pengukuran
  - ✓ Hasilnya bersifat kualitatif

# FUNGSI TES DAN PENGUKURAN

1. Mengadakan Klasifikasi Siswa . . . Kedudukan siswa dalam kelompoknya.
2. Menentukan Status Siswa . . . Mengetahui perkembangan anak didik.
3. Mengadakan diagnosa dan bimbingan . . . Melihat kelemahan dan kekurangan siswa
4. Pemberian motivasi . . . Merangsang mengikuti kegiatan yang diprogramkan.
5. Perbaikan mengajar . . . Koreksi dalam PBM.
6. Menilai guru dan bahan . . . Melihat efektivitas dan efisiensi proses pendidikan.
7. Alat pembantu dalam survey dan penelitian . . . Pengumpulan data yang obyektif.

## Prinsip-Prinsip Evaluasi

- A. Prinsip Pelaksanaan : yaitu prinsip tentang bagaimana evaluasi dilakukan.
  - 1. Evaluasi harus dilakukan secara obyektif
  - 2. Evaluasi dilakukan secara kontinue
  - 3. Evaluasi dilakukan secara komprehensif (integralitas).
  
- B. Prinsip Dasar : yaitu sebagai pedoman kerja dalam melakukan evaluasi
  - 1. Evaluasi adalah alat komunikasi
  - 2. Membantu anak didik untuk mencapai perkembangan potensinya semaksimal mungkin
  - 3. Jangan hanya membandingkannya dengan orang lain saja
  - 4. Memanfaatkan berbagai macam alat / teknik evaluasi
  - 5. Menyarankan langkah-langkah / tindak lanjut yang perlu diambil

## **Kemungkinan-Kemungkinan Kesalahan dalam Evaluasi :**

1. Kesalahan dalam pengamatan atau observasi
2. Kesalahan pada alat pengukur
3. Kesalahan dalam proses analisis data
4. Pengaruh pekerjaan-pekerjaan yang mendahului
5. Kecenderungan untuk menilai lebih tinggi atau lebih rendah
6. Pengaruh kesan-kesan luar

## **Langkah-Langkah dalam Melakukan Evaluasi :**

1. Perencanaan :
  - 1.1. *Kriteria yang akan digunakan*
  - 1.2. *Bentuk tes / alat ukur yang akan digunakan*
  - 1.3. *Menentukan frekuensi evaluasi*
  - 1.4. *Fasilitas dan Perlengkapan*
  - 1.5. *Waktu pelaksanaan pengambilan data*
  - 1.6. *Para pembantu dalam pelaksanaan pengambilan data.*
2. Pengumpulan Data
3. Penelitian data
4. Pengolahan data
5. Penafsiran data.

# Pemilihan Instrumen Yang Layak Pakai

## 1. Validitas

Sebuah instrumen berupa tes dikatakan valid apabila tes tersebut mampu mengukur secara tepat terhadap apa yang semestinya diukur. Dengan kata lain, validitas berkaitan dengan ketepatan instrumen tersebut terhadap konsep, obyek, atau variabel yang hendak diukur sehingga mengukur atau mengevaluasi apa yang semestinya dievaluasi. Persamaan istilah lain yang digunakan untuk kata valid adalah sahih/sangkil.

Validitas suatu tes terdiri dari empat jenis, yaitu: validitas isi, validitas bangun pengertian, validitas ramalan dan validitas kesamaan.

a. **Validitas isi** (*content validity*)

Validitas isi berkenaan dengan kesanggupan tes dalam mengukur isi (content) yang semestinya diukur. Menyusun butir-butir tes berdasarkan seluruh materi yang telah diberikan.

b. **Validitas bangun pengertian** (*construct validity*)

Validitas bangun atau validitas bangun pengertian (*construct validity*) berkenaan dengan kesanggupan suatu tes untuk mengukur konsep pengertian-pengertian yang terkandung dalam materi yang hendak diukurnya.

c. **Validitas prediksi** (*predictive validity*)

Memprediksi artinya meramal, dan meramal selalu berhubungan dengan suatu yang akan terjadi pada masa mendatang. Misalnya butir-butir tes SNMPTN.

d. **Validitas kesamaan** (*concurrent validity*)

Validitas kesamaan suatu tes artinya membuat tes yang memiliki persamaan dengan tes sejenis yang telah ada atau tes yang telah dibakukan (*standardized*).

Catatan : Tinggi rendahnya derajat validitas suatu tes adalah  $r = \pm 1$

Untuk mencari derajat/tingkat validitas suatu tes menggunakan pendekatan statistika, dengan rumus statistika sebagai berikut:

a. Korelasi Product Moment dengan simpangan :

$$r_{xy} = \frac{\sum X_1 Y_1}{\sqrt{(\sum X_1^2) (\sum Y_1^2)}}$$

Arti unsur-unsur tersebut adalah:

$r_{xy}$  = Koefisien korelasi antara variabel X dan variabel Y

$X_1$  = Perbedaan skor variabel X dengan nilai rata-rata dari variabel X

$Y_1$  = Perbedaan skor variabel Y dengan nilai rata-rata dari variabel Y

$\sum X_1 Y_1$  = Jumlah dari hasil perkalian antara  $X_1$  dengan  $Y_1$

$X_1^2$  = Nilai  $X_1$  yang dikuadratkan

$Y_1^2$  = Nilai  $Y_1$  yang dikuadratkan

Contoh penghitungan: Kita akan menghitung validitas tes keterampilan bermain bola basket yang dinotasikan sebagai variabel X. Sebagai kriteria diambil tes standar dari Johnson Basketball Test, yang dinotasikan sebagai variabel Y. Selanjutnya kita buat tabel persiapan sebagai berikut:

Subyek	Hasil Tes X	Hasil Tes Y	X <sub>1</sub>	Y <sub>1</sub>	X <sub>1</sub> <sup>2</sup>	Y <sub>1</sub> <sup>2</sup>	X <sub>1</sub> Y <sub>1</sub>
A	56	34	5	6	25	36	30
B	43	25	-8	-3	64	9	24
C	50	25	-1	-3	1	9	3
D	45	20	-6	-8	36	64	48
E	36	25	-15	-3	225	9	45
F	58	30	7	2	49	4	14
G	55	31	4	3	16	9	12
H	61	30	10	2	100	4	20
I	46	28	-5	0	25	0	0
J	60	32	9	4	81	16	36
Σ	510	280			622	160	232
	51	28					

$$r_{xy} = \frac{\sum X_1 Y_1}{\sqrt{(\sum X_1^2) (\sum Y_1^2)}}$$

$$r_{xy} = \frac{232}{\sqrt{622 \times 160}} = \frac{232}{\sqrt{99520}} = \frac{232}{315.467} = 0.7354 = 0.74$$

b. Korelasi Prudct Moment dengan angka kasar :

$$r_{xy} = \frac{N \cdot \sum XY - (\sum X) (\sum Y)}{\sqrt{\{N \cdot \sum X^2 - (\sum X)^2\} \{N \cdot \sum Y^2 - (\sum Y)^2\}}}$$

$r_{xy}$  = Korelasi antara variabel X dan Y (kriteria)

N = Jumlah subyek

$\sum X$  = Jumlah skor variabel X

$\sum Y$  = Jumlah skor variabel Y

$\sum X^2$  = Jumlah dari kuadrat skor-skor X

$\sum Y^2$  = Jumlah dari kuadrat skor-skor Y

$\sum XY$  = Jumlah dari perkalian skor X dengan Y

$\sum (X)^2$  = Jumlah skor X dikuadratkan

$\sum (Y)^2$  = Jumlah skor Y dikuadratkan

Kemudian kita buat tabel persiapan penghitungan sebagai berikut:

Subyek	Hasil Tes X	Hasil Tes Y	X <sup>2</sup>	Y <sup>2</sup>	XY
A	56	34	3136	1156	1904
B	43	25	1849	625	1075
C	50	25	2500	625	1250
D	45	20	2025	400	900
E	36	25	1296	625	900
F	58	30	3364	900	1740
G	55	31	3025	961	1705
H	61	30	3721	900	1830
I	46	28	2116	784	1288
J	60	32	3600	1024	1920
Σ	510 260100 (X) <sup>2</sup>	280 78400 (Y) <sup>2</sup>	26632	8000	14512

$$r_{xy} = \frac{N \cdot \sum XY - (\sum X) (\sum Y)}{\sqrt{\{N \cdot \sum X^2 - (\sum X)^2\} \{N \cdot \sum Y^2 - (\sum Y)^2\}}}$$

$$r_{xy} = \frac{(10 \times 14512) - (510 \times 280)}{\sqrt{\{(10 \times 26632) - 260100\}\{(10 \times 8000) - 78400\}}}$$

$$\frac{145120 - 142800}{\sqrt{(266320 - 260100)(80000 - 78400)}} = \frac{2320}{\sqrt{6220 \times 1600}}$$

$$\frac{2320}{\sqrt{9952000}} = \frac{2320}{3154.68} = 0.7354 = 0.74$$

Untuk mencari validitas butir tes, selain menggunakan rumus korelasi juga dapat menggunakan pendekatan signifikansi daya pembeda (discriminating power). Langkah-langkahnya adalah sebagai berikut:

### **VALIDITAS BUTIR TES (Teknik Daya Pembeda)**

1. Menyusun Rank Hasil Tes (Bola Volley : Passing, Smash, Service)
2. Menentukan kelompok Atas dan Bawah
  - a. Sampel Besar : 27 % Kelompok Atas dan 27 % Bawah
  - b. Sampel Kecil : 50 % Kelompok Atas dan 50 % Bawah

NO	NAMA	SKOR
1	A	80
2	B	75
3	C	74
4	D	74
5	E	73
6	F	71
7	G	71
8	H	70
9	I	69
10	J	66
11	K	62
12	L	61
13	M	60
14	N	59
15	O	58
16	P	58
17	Q	56
18	R	54
19	S	51
20	T	50

➤ Misalnya : Ingin mencari Validitas Butir Tes Passing

3. Mencari nilai rata-rata ( $\bar{X}$ ) kelompok atas dan kelompok atas dari data butir tes dengan rumus sbb :

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

KELOMPOK ATAS :

1. A = 26

2. B = 24

3. C = 25             $\sum X = 126$

4. D = 28             $\bar{X} = 25.2$

5. E = 23

KELOMPOK BAWAH :

1. P = 25

2. Q = 22

3. R = 19             $\sum X = 110$

4. S = 18             $\bar{X} = 22$

5. T = 26

4. Mencari Variansi ( $S^2$ ) Kelompok atas dan kelompok bawah dari data butir tes dengan rumus sbb :

$$S^2 = \frac{N \cdot \sum X^2 - (\sum X)^2}{N(N-1)}$$

$$\text{KELOMPOK ATAS : } S^2 = \frac{5 \times 3190 - 15876}{5(5-1)} = 3.7$$

$$\text{KELOMPOK BAWAH : } S^2 = \frac{5 \times 2470 - 12100}{5(5-1)} = 12.5$$

5. Masukkan nilai rata-rata dan varians dari masing-masing kelompok ke dalam rumus :

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{N} + \frac{S_2^2}{N}}} = \frac{25.2 - 22}{\sqrt{\frac{3.7}{5} + \frac{12.5}{5}}} = 1.78$$

6. Mencari batas kritis nilai t-tabel pada  $t = 1 - \alpha$  dengan  $dk = n_1 + n_2 - 2$
7. Membandingkan hasil t-hitung dengan nilai t-tabel, Jika :
- Nilai t-hitung  $\geq$  nilai t-tabel, maka perbedaan tersebut signifikan, artinya butir tes tersebut "valid"
  - Nilai t-hitung  $<$  dari nilai t-tabel, maka perbedaan tersebut tidak signifikan, artinya butir tes tersebut "tidak valid"

## 2. Reliabilitas

Kata reliabilitas berasal dari kata *relia-bility* (bahasa Inggris, berasal dari kata dasar *reliable*) yang berarti dapat dipercaya.

Sebuah tes dikatakan reliabel (memiliki reliabilitas) apabila hasil-hasil penggunaan tes tersebut menunjukkan *ketetapan*. Dengan kata lain, apabila kepada para siswa diberikan tes yang sama pada waktu yang berbeda-beda, maka setiap siswa akan tetap berada dalam peringkat (*rangking*) yang sama dalam kelompoknya.

Untuk menentukan tinggi rendahnya reliabilitas dapat menggunakan rumus korelasi product moment sbb:

a. Korelasi Pruduct Moment dengan angka kasar :

$$r = \frac{N \cdot \sum XY - (\sum X) (\sum Y)}{\sqrt{\{N \cdot \sum x^2 - (\sum X)^2\} \{N \cdot \sum Y^2 - (\sum Y)^2\}}}$$

b. Korelasi Product Moment dengan simpangan :

$$r = \frac{\sum X_1 Y_1}{\sqrt{(\sum X_1^2) (\sum Y_1^2)}}$$

Adapun metode yang dapat digunakan untuk menentukan reliabilitas suatu tes antara lain adalah : metode tes ulang, metode paralel, metode belah dua dan metode kesamaan rasional.

➤ **Metode tes ulang** (*test and re-test method*)

Metode tes ulang adalah penggunaan tes terhadap subyek yang sama, dilakukan dalam waktu yang berlainan. Misalnya tes push ups yang dilakukan 2 kali kesempatan, Hasilnya dikorelasikan antara tes I dengan pengulangannya (ke II).

➤ **Metode paralel** (*equivalent method*)

Mengukur reliabilitas dengan metode paralel tidak dilakukan dengan cara pengulangan kepada subyek yang sama, tetapi menggunakan hasil dari bentuk tes yang setara (paralel). Ada dua bentuk tes (misalnya A dan B) yang dianggap sama bobotnya. Hasil tes bentuk A dikorelasikan dengan hasil tes bentuk B.

➤ **Metode belah dua** (*split half method*).

Kelemahan penerapan metode tes ulang dan metode paralel, dapat diatasi dengan menggunakan metode yang ketiga yaitu metode belah dua (*split half method*). Setelah dilakukan pengetesan, hasilnya dipisahkan antara perolehan skor dari butir-butir soal yang bernomor gasal dengan butir-butir soal yang bernomor genap, kemudian dikorelasikan dengan menggunakan **rumus korelasi product moment**.

Hasil penghitungan korelasinya baru koefisien reliabilitas separuh tes. Oleh karena itu, untuk mengetahui koefisien reliabilitas tes secara keseluruhan maka koefisien reliabilitas belah dua perlu dikonversikan ke dalam rumus **Spearman Brown** sebagai berikut :

$$r_{11} = \frac{2 \cdot r_{1/2}}{(1+r_{1/2})}$$

$r_{11}$  = Koefisien korelasi yang dicari

$r_{1/2}$  = Hasil korelasi antara belahan genap dengan yang gasal.

Contoh Soal : Dari hasil tes pelajaran IPS sebanyak 20 butir soal terhadap 10 orang siswa adalah sebagai berikut:

1    2    3    4    5    6    7    8    9    10  
 11 12 13 14 15 16 17 18 19 20

NAMA	NO. GASAL	NO. GENAP
A	8	7
B	7	5
C	6	8
D	5	6
E	4	7
F	8	9
G	3	4
H	4	3
I	9	8
J	6	7

Subyek	Hasil Tes No. Gasal X	Hasil Tes No. Genap Y	X <sup>2</sup>	Y <sup>2</sup>	XY
A	8	7	64	49	56
B	7	5	49	25	35
C	6	8	36	64	48
D	5	6	25	36	30
E	4	7	16	49	28
F	8	9	64	81	72
G	3	4	9	16	12
H	4	3	16	9	12
I	9	8	81	64	72
J	6	7	36	49	42
Σ	60 3600 (X) <sup>2</sup>	64 4096 (Y) <sup>2</sup>	396	442	407

$$r = \frac{N \cdot \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \cdot \sum x^2 - (\sum X)^2\} \{N \cdot \sum Y^2 - (\sum Y)^2\}}}$$

$$\frac{(10 \times 407) - (60)(64)}{\sqrt{\{(10 \times 396) - 3600\} \{(10 \times 442) - 4096\}}} = \frac{4070 - 3840}{\sqrt{360 \times 324}}$$

$$= \frac{230}{\sqrt{116640}} = \frac{230}{341.5259} = 0.6734 = 0.67 \text{ (korelasi parohan tes)}$$

Hasil penghitungan korelasi antara belahan yang gasal dan genap adalah sebesar :  
 $r = 0.67$  maka korelasi seluruh tesnya adalah sebagai berikut;

$$r_{11} = \frac{2 \cdot r_{1/2}}{(1+r_{1/2})} = \frac{2 \times 0.67}{(1+0.67)} = 0.80$$

a. Metode kesamaan rasional (*rational same method*)

Selain ketiga metode yang diuraikan di depan, maka terdapat metode menghitung reliabilitas tanpa harus melakukan tes ulang, tes setara (paralel) maupun belah dua. Metode tersebut adalah metode kesamaan rasional. Metode ini dilakukan dengan cara menghubungkan setiap item atau butir dalam suatu tes dengan butir-butir lainnya dalam tes itu sendiri secara keseluruhan.

Adapun teknik analisis yang sering digunakan untuk mencari besarnya reliabilitas seluruh tes, adalah menggunakan rumus-rumus sbb:

$$1. \text{ Kuder Richardson (KR-20) } r_{11} = \left( \frac{n}{n-1} \right) \left( \frac{s^2 - \sum pq}{s^2} \right)$$

$r_{11}$  = Reliabilitas Tes

$p$  = Proporsi Subyek yang menjawab Benar

$q$  = Proporsi subyek yang menjawab Salah (= 1 -  $p$ )

$n$  = Banyaknya Butir Tes

$\sum pq$  = Jumlah dari  $p \times q$

$$S^2 = \text{Varians} = \frac{N \cdot \sum X^2 - (\sum X)^2}{N(N-1)}$$

Contoh mencari varians ( $S^2$ ) :

$N$  = Jumlah Peserta Tes

$\sum X^2$  = Jumlah kuadrat dari skor-skor

$(\sum x)^2$  = Jumlah Skor yang dikuadratkan

Subyek	Skor (X)	$X^2$
A	11	121
B	10	100
C	8	64
D	7	49
E	4	16
Jumlah ( $\sum$ )	40	350

$$S^2 = \frac{5 \times 350 - (40)^2}{5(5 - 1)} = \frac{1750 - 1600}{20} = \frac{150}{20} = 7.5$$

MENCARI RELIABILITAS TES DENGAN RUMUS KR-20

NO	NAMA	NOMOR ITEM / BUTIR TES							SKOR TOTAL
		1	2	3	4	5	6	7	
1	A	1	-	1	1	1	1	-	5
2	B	-	1	1	-	1	1	1	5
3	C	-	-	-	-	1	-	1	2
4	D	-	1	1	1	1	1	1	6
5	E	1	-	-	-	1	-	-	2
6	F	-	1	1	1	1	-	-	4
7	G	-	-	-	1	1	1	-	3
8	H	-	1	-	1	1	-	-	3
9	I	-	1	-	1	1	-	-	3
10	J	-	-	-	1	1	-	-	2
Σ skor		2	5	4	7	10	4	3	35
p		0.2	0.5	0.4	0.7	1	0.4	0.3	2.06 (s <sup>2</sup> )
q		0.8	0.5	0.6	0.3	-	0.6	0.7	
pq		0.16	0.25	0.24	0.21	-	0.24	0.21	1.31

$$r_{11} = \left( \frac{n}{n-1} \right) \left( \frac{s^2 - \sum pq}{s^2} \right)$$

$$r_{11} = \left( \frac{7}{6} \right) \left( \frac{2.06 - 1.31}{2.06} \right) = 1.17 \times 0.36 = 0.42$$

2. Rumus Kuder Richardson (KR-21) :  $r_{11} = \left( \frac{n}{n-1} \right) \left( \frac{s^2_t - \sum s^2_i}{s^2_t} \right)$

$r_{11}$  = Reliabilitas Tes

$n$  = Banyak Butir Tes

$s^2_t$  = Varians Skor Total

$\sum s^2_i$  = Jumlah Varians Butir Tes

3. Rumus Alpha :  $r_{11} = \left( \frac{n}{n-1} \right) \left( 1 - \frac{\sum s^2_i}{s^2_t} \right)$

$r_{11}$  = Reliabilitas Tes

$n$  = Banyak Butir Soal/Tes

$\sum s^2_i$  = Jumlah Varians Butir Tes

$s^2_t$  = Varians Skor Total

## TABEL ANALISIS ITEM

NO	NAMA	NOMOR ITEM / BUTIR TES						SKOR TOTAL	KUADRAT SKOR TOTAL
		1	2	3	4	5	6		
1	A	10	6	8	8	10	10	52	2704
2	B	6	4	4	6	6	5	31	961
3	C	8	2	6	8	7	8	39	1521
4	D	7	3	7	7	6	6	36	1296
5	E	0	5	3	2	4	4	18	324
6	F	2	4	2	8	6	8	30	900
7	G	4	3	6	6	6	6	31	961
8	H	5	5	5	7	7	7	36	1296
9	I	5	5	4	6	8	5	33	1089
10	J	3	6	3	4	6	6	28	784
$\Sigma X$		50	43	48	62	66	65	334	11836
$\Sigma X^2$		328	201	264	418	458	451	$\Sigma s^2_i = 23.58$ $s^2_t = 75.60$	

$$\text{KR-21: } r_{11} = \left( \frac{n}{n-1} \right) \left( \frac{s^2t - \sum s^2i}{s^2t} \right) = \left( \frac{6}{5} \right) \left( \frac{75.60 - 23.58}{75.60} \right) = 0.83$$

$$\text{ALPHA: } r_{11} = \left( \frac{n}{n-1} \right) \left( 1 - \frac{\sum s^2i}{s^2t} \right) = \left( \frac{6}{5} \right) \left( 1 - \frac{23.58}{75.60} \right) = 0.83$$

Keempat jenis reliabilitas yang telah diuraikan di atas harus dipilih mana yang paling tepat digunakan. Pemilihannya mempertimbangkan sifat-sifat variabel yang hendak diukur, jenis tes, jumlah subyek (*testee*), serta hasil-hasil yang diharapkan sesuai dengan tujuan pembelajaran.

### 3. Objektivitas

Obyektivitas adalah derajat kesamaan hasil dari dua orang testor atau lebih terhadap obyek dan subyek yang sama. Dalam pengertian sehari-hari telah diketahui bahwa obyektif berarti tidak ada unsur pribadi yang mempengaruhi. Kebalikan dari obyektif adalah subyektif, yang berarti terdapat unsur pribadi yang masuk mempengaruhi.

Untuk meningkatkan obyektivitas dalam pengukuran, hal-hal berikut seharusnya terus diupayakan, yaitu:

- Petunjuk atau prosedur pengukuran harus dirumuskan dengan kata-kata yang tepat dan terperinci.
- Prosedur pengukuran diupayakan agar mudah dan bersifat operasional.
- Apabila mungkin, dapat dipergunakan alat pengukur mekanik.
- Memilih penguji yang telah berpengalaman (*qualified*).
- Para penguji harus menjunjung tinggi sikap-sikap atau kode etik ilmiah.

Selain hal-hal tersebut di atas, untuk melihat tingkat obyektivitas dari tester, dapat menggunakan rumus korelasi product moment.

Tinggi rendahnya derajat validitas, reliabilitas, dan objektivitas suatu tes dinyatakan dengan koefisien korelasi, yaitu ;  $r = \pm 1$ .

Acuan koefisien korelasi suatu tes (Mathews : 1963) adalah sbb :

$r = 0.90 - 0.99$  berarti Sempurna / Sangat tinggi

$r = 0,80 - 0.89$  berarti Tinggi

$r = 0.70 - 0.79$  berarti Sedang / Cukup

$r = 0.60 - 0.69$  berarti Kurang

$r =$  di bawah 0.59 berarti Sangat Kurang

#### **4. Praktikabilitas**

Meskipun kriteria validitas dan reliabilitas tes merupakan hal yang terpenting dari kriteria yang lainnya, namun sejumlah pertimbangan yang bersifat praktis dan dapat mempengaruhi tes perlu dipertimbangkan juga. Pertimbangan-pertimbangan tersebut antara lain meliputi kemudahan dalam administrasi dan interpretasi, waktu, tenaga, peralatan/fasilitas, serta biaya.

## Beberapa Jenis Skala Penilaian

Dalam melakukan evaluasi hasil belajar siswa, guru dapat menggunakan beberapa bentuk skala penilaian sebagai berikut:

### **a. Skala 1-10**

Ada suatu anggapan bahwa angka 10 sebagai angka tertinggi, karena guru-guru di Indonesia pada umumnya memiliki kebiasaan menggunakan skala 1-10 untuk membuat laporan prestasi belajar siswanya dalam rapor. Bahkan ada kalanya juga digunakan skala 1-100 sehingga memungkinkan guru dalam memberikan evaluasi yang lebih luas. Dalam skala 1-10, guru jarang memberikan angka pecahan seperti 5,5. Angka-angka pecahan tersebut kemudian dibulatkan menjadi 6. Padahal angka 6,4 pun angka dibulatkan menjadi 6. Dengan demikian, maka rentangan angka 5,5 sampai dengan 6,4 (selisihnya hampir 1), tetap keluar dengan satu angka yaitu: 6.

### **b. Skala 1-100**

Biasanya dengan menggunakan skala 1-10, bilangan bulat yang ada masih menunjukkan evaluasi yang agak kasar. Oleh karena itulah, dengan menggunakan skala 1-100 dimungkinkan melakukan evaluasi yang lebih halus, karena terdapat 100 bilangan bulat. Sebagai contoh 5,5 dan 6,4 dalam skala 1-10 biasanya dibulatkan menjadi angka 6, sedangkan dalam skala 1-100 kedua angka tersebut tetap dituliskan dengan angka 55 dan 64.

### **c. Skala Huruf**

Selain menggunakan angka, maka pemberian nilai dapat dilakukan dengan menggunakan skala huruf, misalnya: A, B, C, D dan E.

Sebenarnya sebutan skala di atas masih diperdebatkan oleh beberapa pihak. Ada yang mengatakan, bahwa jarak antara huruf A dan B tidak dapat menggambarkan interval yang sama dengan jarak antara B dan C, atau C dan D. Sedangkan dalam bentuk angka, hal ini dapat dibuktikan dengan garis bilangan bahwa jarak antara 1 dan 2 sama dengan jarak antara 2 dan 3, dst.

Huruf terdapat dalam urutan abjad, penggunaannya dalam evaluasi akan terasa lebih tepat karena tidak menafsirkan arti perbandingan. Huruf tidak menunjukkan kuantitas, akan tetapi dapat dipergunakan sebagai simbol untuk menggambarkan kualitas. Akan tetapi, penggunaan huruf dalam pengisian rapor akan menjumpai kesulitan dalam mengambil jumlah atau rata-rata. Padahal guru tidak dapat terlepas dari upaya mengambil rata-rata.

Salah satu cara yang dapat dipergunakan untuk menentukan rata-rata dari huruf, adalah dengan cara mentransfer skala huruf tersebut menjadi skala angka terlebih dahulu. Yang sering digunakan adalah satu huruf akan mewakili satu rentangan nilai angka. Biasa dikonversikan ke dalam skala 1 – 4 atau 1 – 5. Dengan cara mengkonversikan nilai skala huruf menjadi skala angka, maka dengan mudah dapat dihitung rata-ratanya.

## Beberapa Bentuk Penentuan Nilai

Beberapa bentuk nilai yang biasanya diberikan oleh para guru penjas antara lain adalah:

- ✓ Dalam bentuk angka, misal merentang dari 1 hingga 10
- ✓ Dalam bentuk huruf, misal dari A, B, C, D, E
- ✓ Dalam bentuk kata, misal istimewa, bagus, cukup, dst.
- ✓ Dalam bentuk prosentase, misal 80%, 70%, 60%, dst.
- ✓ Dalam bentuk dua kelas dikhotomi, misal berhasil dan gagal.

Bentuk penentuan nilai dapat dilakukan melalui beberapa cara. Adang Suherman (2001) menyarankan prosedur sebagai berikut:

- a. Menentukan skor alternatif jawaban, dengan menggunakan berbagai skala tersebut di atas.
- b. Menentukan skor total ideal. Skor ideal adalah skor tertinggi yang mungkin dapat diperoleh siswa. Skor total ideal adalah jumlah total dikali skor ideal masing-masing alternatif jawaban. Misal apabila jumlah soal 10 dan skor ideal dari masing-masing alternatif jawaban 4, maka skor total ideal adalah  $10 \times 4 = 40$ .

- c. Menentukan skor jawaban siswa. Jumlah skor jawaban siswa dari masing – masing item soal. Misalnya, seorang siswa dari 10 soal yang diajukan terdapat :
- 5 jawaban memperoleh skor 4 = 20
- 3 jawaban memperoleh skor 3 = 9
- 2 jawaban memperoleh skor 2 = 4
- dengan demikian skor siswa tersebut adalah  $20 + 9 + 4 = 33$
- d. Membuat konversi nilai. Lihat contoh di bawah ini!

Tingkat Penguasaan (%)	Tingkat Penguasaan (Skor)	Katagori	Nilai
80 % ke atas	$80 \% \times 40 = 32$	Istimewa	10
70 % - 79 %	$70 \% \times 40 = 28$	Baik sekali	9
60 % - 69 %	$60 \% \times 40 = 24$	Baik	8
50 % - 59 %	$50 \% \times 40 = 20$	Cukup	7
dibawah 50 %	dibawah 20	Kurang	6

(\*Sumber: Adang Suherman, 2001)

**Pemberian Makna** (pendekatan evaluasi) dapat dilakukan dengan :

1. Kriteria Absolut atau Criterion-Referenced Standard, sering juga disebut Penilaian Acuan Patokan (PAP). *Pendekatan acuan patokan (PAP) ini merupakan pendekatan evaluasi yang membandingkan proses dan hasil belajar siswa dengan suatu patokan atau kriteria tertentu yang biasanya telah ditetapkan sebelumnya. Apabila siswa berhasil mencapai atau melewati patokan tersebut, maka ia dianggap berhasil atau lulus.*

2. Kriteria Kelompok atau Criterion-Referenced Norm, sering juga disebut Penilaian Acuan Norma (PAN). *Penilaian menggunakan acuan normatif ini dilakukan yaitu membandingkan skor siswa dengan rerata skor kelompoknya sebagai norma. Pendekatan ini pada dasarnya bertitik tolak dari penggunaan kurva normal, rerata (Mean) kelompok dan simpangan baku yang menjadi acuannya.*

3. Gabungan antara PAN dan PAP. *Terlebih dahulu ditetapkan passing grade, kemudian siswa yang lulus saja ditentukan kategori nilainya.*

## Penilaian Acuan Patokan (PAP)

Misal skor maksimum 80. Batas penguasaan minimumnya = 50% - 60% dengan nilai 6.  
Norma penilaian 1 - 10 dapat disusun sebagai berikut :

Prosentase Penguasaan	Rentang Skor	Nilai
91 % - 100 %	73 – 80	10
81 % - 90 %	65 – 72	9
71 % - 80 %	57 – 64	8
61 % - 70 %	49 – 56	7
50 % - 60 %	40 – 48	6
40 % - 49%	32 – 39	5
30 % - 39 %	24 – 31	4
20 % - 29 %	16 – 23	3
10 % - 19 %	8 – 15	2
0 % - 9 %	0 – 7	1

### Penilaian Acuan Patokan (PAP) Menggunakan 5 Standar (A,B,C,D,E)

Tingkat Penguasaan	Rentang Skor	Nilai	Kategori
80% - 100%	64 ke atas	A	Baik Sekali
60% - 79%	48 - 63	B	Baik
40% - 59%	32 - 47	C	Cukup
20% - 39%	16 - 31	D	Kurang
19% Ke bawah	0 - 15	E	Sangat Kurang

## Penilaian Acuan Norma (PAN)

### a. Penggunaan Kurve Normal dengan 5 Kategori Nilai (A - E)

Batas daerah dalam Kurve	Nilai	Kategori
M + 1.8 S atau lebih	A	Sangat Baik
Antara M + 0.6 S dan M +1.8 S	B	Baik
Antara M – 0.6 S dan M + 0.6 S	C	Cukup
Antara M – 1.8 S dan M – 0.6 S	D	Kurang
Kurang dari M – 1.8 S	E	Sangat Kurang

### b. Penggunaan Kurve Normal dengan skala nilai 1 - 10

Misal: Rerata Skor = 30 dan simpangan bakunya = 5

Skala	Batas Skor	Rentang Skor	Nilai
M + 2.4 S	$30 + 2.4 (5) = 42$	42 - ke atas	10
M + 1.8 S	$30 + 1.8 (5) = 39$	39 – 41	9
M + 1.2 S	$30 + 1.2 (5) = 36$	36 – 38	8
M + 0.6 S	$30 + 0.6 (5) = 33$	33 – 35	7
M + 0.0 S	$30 + 0.0 (5) = 30$	30 – 32	6
M – 0.6 S	$30 - 0.6 (5) = 27$	27 – 29	5
M – 1.2 S	$30 - 1.2 (5) = 24$	24 – 26	4
M – 1.8 S	$30 - 1.8 (5) = 21$	21 – 23	3
M – 2.4 S	$30 - 2.4 (5) = 18$	18 – 20	2
		17 - ke bawah	1

$$\text{Rumus Rata-rata } (\bar{X}) = \frac{\sum X}{N}$$

$$\text{Rumus Simpangan Baku (S)} = \sqrt{\frac{\sum(X - \bar{X})^2}{n-1}}$$

s = Simpangan baku

X = Skor yang dicapai seseorang

$\bar{X}$  = Nilai rata-rata

n = Banyaknya jumlah orang

Subyek	Skor (X)	(X - $\bar{X}$ )	(X - $\bar{X}$ ) <sup>2</sup>
<b>A</b>	<b>11</b>	<b>3</b>	<b>9</b>
<b>B</b>	<b>10</b>	<b>2</b>	<b>4</b>
<b>C</b>	<b>8</b>	<b>0</b>	<b>0</b>
<b>D</b>	<b>7</b>	<b>-1</b>	<b>1</b>
<b>E</b>	<b>4</b>	<b>-4</b>	<b>16</b>
$\Sigma$	<b>40</b>	<b>-</b>	<b>30</b>
	<b>8</b>	<b>-</b>	<b>-</b>

$$s = \sqrt{\frac{\sum(X - \bar{X})^2}{n-1}} = \sqrt{\frac{30}{5-1}} = \sqrt{\frac{30}{4}} = \sqrt{7.5} = 2.74$$

## ➤ Analisis Butir Soal Tes Obyektif

### a. Indeks Kesukaran (difficulty index) Butir Soal.

Soal yang baik adalah soal yang tidak terlalu mudah atau terlalu sukar.

Soal yang terlalu mudah tidak merangsang siswa untuk mempertinggi usaha untuk memecahkannya, sebaliknya soal yang sangat sukar akan menyebabkan siswa menjadi putus asa dan tidak semangat untuk mencoba lagi karena di luar jangkauannya.

Indeks kesukaran butir soal berkisar antara 0.00 sampai 1.00. Adapun rumus indeks kesukaran butir soal adalah sebagai berikut:

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

Keterangan:

P = Indeks kesukaran butir soal

B = Banyaknya siswa yang menjawab soal dengan benar

JS = Jumlah siswa peserta tes

Contoh : Ada 10 siswa mengerjakan soal-soal tes yang terdiri dari 10 butir soal.

Cara Mengerjakannya:

1. Setelah soal selesai dikoreksi, hasilnya dimasukkan dalam tabel persiapan analisis indeks kesukaran butir soal sbb:

Tabel Persiapan Analisis Indeks Kesukaran Butir Soal  
(Jawaban benar = 1 ; Jawaban salah = 0)

No	Siswa	Butir Soal										Skor Siswa
		1	2	3	4	5	6	7	8	9	10	
1	A	1	1	0	0	1	0	1	1	0	1	6
2	B	0	1	0	0	1	0	0	1	0	0	3
3	C	1	1	0	0	1	1	1	1	0	1	7
4	D	0	1	0	0	1	1	0	1	0	0	4
5	E	1	1	0	0	1	0	1	1	0	1	6
6	F	0	0	0	1	1	1	0	1	0	0	4
7	G	1	0	0	1	0	0	1	1	0	1	5
8	H	0	0	0	1	0	0	1	1	0	0	3
9	I	1	1	1	1	1	0	1	1	0	0	8
10	J	0	1	1	1	1	0	1	1	0	0	6
$\Sigma$		5	7	2	5	8	2	7	10	0	4	52

2. Dari tabel tsb di atas dilakukan analisis indeks kesukaran butir soal sbb:

✓ Untuk butir soal no 1.

Jumlah siswa yang menjawab benar,  $B = 5$

Jumlah peserta tes,  $JS = 10$

Masukkan ke dalam rumus :  $P = \frac{B}{JS} = \frac{5}{10} = 0.50$

✓ Untuk butir soal no 2.

Jumlah siswa yang menjawab benar,  $B = 7$

Jumlah peserta tes,  $JS = 10$

Masukkan ke dalam rumus :  $P = \frac{B}{JS} = \frac{7}{10} = 0.70$

3. Konsultasikan angka-angka indeks kesukaran butir-butir soal dengan tabel klasifikasi indeks kesukaran untuk menarik kesimpulannya sbb:

$P = 0.00 - 0.30$  adalah sukar

$P = 0.31 - 0.70$  adalah sedang

$P = 0.71 - 1.00$  adalah mudah

### **b. Indeks Diskriminasi (daya beda) butir soal.**

Daya beda adalah kemampuan suatu soal untuk membedakan antara peserta tes yang berkemampuan tinggi dengan peserta tes yang berkemampuan rendah. Indeks diskriminasi butir soal berkisar antara  $\pm 1$ . Makin tinggi indeks diskriminasinya, maka butir soal tersebut makin baik. Jika suatu soal indeks daya bedanya menunjukkan angka negatif sebaiknya soal tersebut dibuang saja, sebab angka negatif menunjukkan bahwa soal-soal tersebut memiliki daya beda yang terbalik, yaitu kelompok atas (pandai) malah tidak dapat mengerjakan soal tersebut, sedangkan kelompok bawah dapat mengerjakan dengan benar.

Rumus indeks daya beda butir soal adalah sbb:

$$D = \frac{BA}{JA} - \frac{BB}{JB} = PA - PB$$

Keterangan:

D = Indeks diskriminasi atau indeks daya beda butir soal

JA = Jumlah peserta kelompok atas

JB = Jumlah peserta kelompok bawah.

BA = Banyaknya peserta tes kelompok atas yang menjawab benar.

BB = Banyaknya peserta tes kelompok bawah yang menjawab benar.

PA = Proporsi peserta tes kelompok atas yang menjawab benar.

PB = Proporsi peserta tes kelompok bawah yang menjawab benar.

Langkah-langkah menghitung indeks daya beda butir soal adalah sbb:

1. Menyusun ranking hasil tes secara keseluruhan, mulai dari skor tertinggi sampai skor terendah.
2. Menentukan kelompok atas dan bawah dengan ketentuan sbb:
  - Sampel kecil: Seluruh peserta tes di deretkan mulai dari yang memperoleh skor teratas hingga yang memperoleh skor paling bawah. Kemudian seluruh peserta tes dibagi dua sama besar, yaitu 50% kelompok atas dan 50% kelompok bawah.
  - Sampel besar: Seluruh peserta tes di deretkan mulai dari peserta yang memperoleh skor teratas hingga peserta yang memperoleh skor paling bawah. Kemudian peserta tes dibagi, yaitu 27% kelompok atas dan 27% kelompok bawah.

Contoh : Dari hasil tes terhadap 10 orang peserta tes diperoleh hasil sbb:

SISWA	SKOR BUTIR SOAL										SKOR SISWA	KELOMPOK
	1	2	3	4	5	6	7	8	9	10		
A	1	1	1	1	1	0	1	1	0	1	8	50% KELOMPOK ATAS (JA = 5)
B	1	1	0	0	1	1	1	1	0	1	7	
C	0	1	1	0	1	1	1	1	0	1	7	
D	1	1	0	1	1	0	1	1	1	0	7	
E	1	1	0	0	1	0	1	1	0	1	6	
<b>Σ</b>	<b>4</b>	<b>5</b>	<b>2</b>	<b>2</b>	<b>5</b>	<b>2</b>	<b>5</b>	<b>5</b>	<b>1</b>	<b>4</b>	<b>35</b>	
F	0	1	0	1	1	0	1	1	0	0	5	50% KELOMPOK ATAS (JB = 5)
G	0	1	0	0	1	0	0	1	0	0	3	
H	0	0	0	1	0	0	1	1	0	0	3	
I	1	0	0	0	1	0	0	0	1	0	3	
J	0	0	1	0	0	1	0	0	0	0	2	
<b>Σ</b>	<b>1</b>	<b>2</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>1</b>	<b>0</b>	<b>16</b>	
<b>TOTAL</b>	<b>5</b>	<b>7</b>	<b>3</b>	<b>4</b>	<b>8</b>	<b>3</b>	<b>7</b>	<b>8</b>	<b>2</b>	<b>4</b>	<b>51</b>	

Langkah-langkah yang ditempuh adalah sbb:

1. Menderetkan perolehan skor siswa peserta tes, dari yang memperoleh skor tertinggi hingga yang memperoleh skor terendah.
2. Membagi 50% kelompok atas dan 50% kelompok bawah

Contoh melakukan analisis indeks daya beda butir soal sbb:

- ✓ untuk butir soal no.1

Kelompok atas yang menjawab benar (BA) ada 4 orang

Kelompok bawah yang menjawab benar (BB) ada 1 orang

Jumlah peserta tes kelompok atas (JA) = kelompok bawah (JB) = 5 orang

3. Masukkan dalam rumus:  $D = \frac{BA}{JA} - \frac{BB}{JB} = PA - PB$

$$D = \frac{4}{5} - \frac{1}{5} = PA - PB = 0.80 - 0.20 = 0.60$$

✓ Untuk butir soal no. 5

Kelompok atas) yang menjawab benar (BA ada 5 orang

Kelompok bawah yang menjawab benar (BB) ada 3 orang

Jumlah peserta tes kelompok atas (JA) = kelompok bawah (JB) = 5 orang

$$\text{Masukkan dalam rumus: } D = \frac{5}{5} - \frac{3}{5} = PA - PB = 1 - 0.60 = 0.40$$

4. Konsultasikan angka-angka indeks daya beda butir-butir soal dengan tabel klasifikasi indeks kesukaran untuk menarik kesimpulannya sbb:

D = Negatif : Sangat Jelek

D = 0.00 – 0.20 : Jelek

D = 0.21 – 0.40 : Cukup

D = 0.41 – 0.70 : Baik

D = 0.71 – 1.00 : Sangat Baik

### c. Pola Jawaban Butir Soal / Keberfungsian Option.

Yang dimaksud dengan pola jawaban butir soal adalah distribusi peserta tes dalam hal menentukan alternatif/option jawaban pada tes bentuk pilhan ganda. Pola jawaban butir soal diperoleh dengan menghitung banyaknya peserta tes yang memilih pilihan/alternatif jawaban (a, b, c, d) atau yang tidak memilih pilihan manapun (disebut Omit = O). Idealnya setiap option harus terisi.

Dari pola jawaban butir soal dapat ditentukan apakah pengecoh (distraktor) berfungsi sebagai pengecoh yang baik atau tidak? Pengecoh yang tidak dipilih sama sekali berarti bahwa pengecoh tersebut jelek, hal ini mungkin disebabkan karena terlalu mencolok menyesatkan.

Misal, peserta tes terdiri dari 0 orang disuruh memilih 4 butir soal dengan 4 option (a, b, c, d)

Tabel Persiapan Analisis Pola Jawaban Butir Soal  
(Yang benar = 1, Salah = 0)

SISWA	NOMOR DAN PILIHAN BUTIR SOAL																				KELOMPOK
	1					2					3					4					
	a	b*	c	d	o	a*	b	c	d	o	a	b	c	d*	o	a	b	c*	d	o	
A	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	
B	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	
C	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	
D	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	
E	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	
F	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	
G	0	0	1	0	0	1	0	0	0	0	0	0	0	1	1	0	0	1	0	0	
H	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	0	0	0	
I	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
J	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	
<b>Σ</b>	<b>0</b>	<b>7</b>	<b>2</b>	<b>1</b>	<b>0</b>	<b>9</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>5</b>	<b>0</b>	<b>3</b>	<b>1</b>	<b>3</b>	<b>1</b>	<b>4</b>	<b>1</b>	<b>1</b>	
P	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	
Q	0	1	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	
R	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	1	0	1	0	0	
S	1	0	0	0	0	1	0	1	0	0	0	0	0	0	1	1	0	1	0	0	
T	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	
U	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	
V	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1	
W	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	
X	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	
Y	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	
<b>Σ</b>	<b>2</b>	<b>3</b>	<b>3</b>	<b>1</b>	<b>1</b>	<b>5</b>	<b>1</b>	<b>3</b>	<b>1</b>	<b>0</b>	<b>2</b>	<b>2</b>	<b>3</b>	<b>1</b>	<b>2</b>	<b>2</b>	<b>0</b>	<b>5</b>	<b>1</b>	<b>2</b>	

Cara Menganalisis setiap pengecoh (distraktor) pola jawaban butir soal yaitu: menjumlahkan pemilih kelompok atas dengan kelompok bawah, kemudian dibagi dengan jumlah peserta tes kelompok atas dan bawah, selanjutnya dikalikan 100%.

Kesimpulannya :

1. Jika pemilih kelompok atas lebih sedikit dari pada kelompok bawah, maka pengecoh tersebut berfungsi baik. Sebaliknya jika pemilih kelompok atas lebih banyak dari pada kelompok bawah, maka pengecoh tersebut jelek.
2. Jika hasilnya  $5\% >$  maka pengecoh tersebut berfungsi baik, sebaliknya; Jika hasilnya  $<$  dari  $5\%$  maka pengecoh tersebut jelek.

Contoh Pengerjaannya:

➤ Untuk butir soal nomor 1. (Jawaban yang benar adalah b).

$$\text{Pengecoh a:} = \frac{0+2}{10+10} \times 100\% = 0.1 \times 100\% = 10\%$$

$$\text{Pengecoh c:} = \frac{2+3}{10+10} \times 100\% = 0.25 \times 100\% = 25\%$$

$$\text{Pengecoh d:} = \frac{1+1}{10+10} \times 100\% = 0.1 \times 100\% = 10\%$$

$$\text{Omit O} = \frac{0+1}{10+10} \times 100\% = 0.05 \times 100\% = 5\%$$

Kesimpulannya:

1. Semua pengecoh untuk butir soal no.1 ini sudah berfungsi dengan baik, karena tes dipilih oleh lebih 5% dari seluruh peserta tes dan pemilih kelompok atas lebih sedikit dari pada kelompok bawah.
2. Tidak lebih dari 5% peserta tes yang blanko (Omit), kelompok atas lebih sedikit yang blanko dari pada kelompok bawah.

- Untuk butir soal no.3. (Jawaban yang benar adalah d)

$$\text{Pengecoh a:} = \frac{1+2}{10+10} \times 100\% = 0.15 \times 100\% = 15\%$$

$$\text{Pengecoh b:} = \frac{5+2}{10+10} \times 100\% = 0.35 \times 100\% = 35\%$$

$$\text{Pengecoh c:} = \frac{0+3}{10+10} \times 100\% = 0.15 \times 100\% = 15\%$$

$$\text{Omit O} = \frac{1+2}{10+10} \times 100\% = 0.15 \times 100\% = 15\%$$

Kesimpulannya:

1. Pengecoh a dan c untuk butir soal no.3 ini sudah berfungsi dengan baik, karena tes dipilih oleh lebih 5% dari seluruh peserta tes dan pemilih kelompok atas lebih sedikit dari pada kelompok bawah. Namun Pengecoh b jelek, karena pemilih kelompok atas lebih banyak dari pada kelompok bawah.
2. Terlalu banyak peserta tes yang blanko/Omit (> 5%), maka butir soal ini perlu ditinjau kembali.

#### d. Indeks Reliabilitas Tes

Dengan melakukan analisis indeks reliabilitas tes ini akan diketahui apakah alat ukur yang digunakan mempunyai tingkat keajegan/kemantapan/kestabilan yang baik atau tidak. Ukuran tinggi rendahnya derajat keterandalan suatu tes disebut indeks reliabilitas yang digambarkan melalui koefisien korelasi yang besarnya berkisar antara  $\pm 1$ . Jika  $r = 0.00$  berarti tidak ada hubungan dari kedua variabel.

Contoh: Dari 10 orang peserta tes terhadap 10 butir soal diperoleh hasil sbb:

Siswa	Skor Butir Soal										Jumlah Skor	
	1	2	3	4	5	6	7	8	9	10	Soal Gasal	Soal Genap
A	1	1	0	1	0	0	1	1	1	1	3	4
B	0	1	0	0	1	1	1	0	0	1	2	3
C	1	1	1	1	1	1	1	0	1	0	5	3
D	1	0	0	1	1	0	0	1	1	1	3	3
E	1	1	0	1	0	1	1	0	1	1	3	4
F	0	0	0	0	1	1	0	0	0	1	1	2
G	1	1	1	1	0	0	1	1	1	1	4	4
H	1	0	0	1	0	0	1	1	0	0	2	2
I	1	1	1	1	1	1	1	1	0	1	4	5
J	1	1	0	1	1	0	1	1	1	1	4	4

Tabel Persiapan Analisis Korelasi Skor Butir Soal Gasal dan Genap

N0	Soal Gasal X	Soal Genap Y	X <sup>2</sup>	Y <sup>2</sup>	X.Y
1	3	4	9	16	12
2	2	3	4	9	6
3	5	3	25	9	15
4	3	3	9	9	9
5	3	4	9	16	12
6	1	2	1	4	2
7	4	4	16	16	16
8	2	2	4	4	4
9	4	5	16	25	20
10	4	4	16	16	16
Σ	31 961 = (X) <sup>2</sup>	34 1156 = (Y) <sup>2</sup>	109	124	112

$$r_{xy} = \frac{N \cdot \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \cdot \sum X^2 - (\sum X)^2\} \{N \cdot \sum Y^2 - (\sum Y)^2\}}}$$

$r_{xy}$  = Korelasi antara variabel X (butir soal gasal) dan Y (butir soal genap)

N = Jumlah subyek = 10

$\sum XY$  = Jumlah dari perkalian skor X dengan Y = 112

$\sum X$  = Jumlah skor variabel X = 31

$\sum Y$  = Jumlah skor variabel Y = 34

$\sum X^2$  = Jumlah dari kuadrat skor-skor X = 109

$\sum Y^2$  = Jumlah dari kuadrat skor-skor Y = 124

$\sum (X)^2$  = Jumlah skor variabel X dikuadratkan = 961

$\sum (Y)^2$  = Jumlah skor variabel Y dikuadratkan = 1156

$$r_{xy} = \frac{N \cdot \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \cdot \sum X^2 - (\sum X)^2\} \{N \cdot \sum Y^2 - (\sum Y)^2\}}} = \frac{(10 \times 112) - (31 \times 34)}{\sqrt{\{(10 \times 109) - 961\} \{(10 \times 124) - 1156\}}}$$

$$= \frac{1120 - 1054}{\sqrt{(1090 - 961)(1240 - 1156)}} = \frac{66}{\sqrt{129 \times 84}} = \frac{66}{\sqrt{10836}} = \frac{66}{104.0961}$$

$$= 0.6340$$

Karena metode yang dipakai adalah split half, hasil penghitungannya adalah korelasi separuh tes. Agar diperoleh koefisien korelasi seluruh tes maka masih perlu dikoreksi dengan rumus Spearman Brown sbb:

$$r_{11} = \frac{2 \cdot r_{1/2}}{(1+r_{1/2})} = \frac{2 \times 0.63}{(1+0.63)} = \frac{1.26}{1.63} = 0.77$$

Setelah hasil penghitungan korelasi diperoleh, untuk menetapkan analisisnya dilakukan dengan cara mengkonsultasikan indeks reliabilitas tes dengan tabel klasifikasi korelasi tes. Menurut Barry L. Johnson (1974), koefisien korelasi tes diklasifikasikan sebagai berikut:

r : 0.00	= Tidak ada hubungan
r : 0.01 – 0.20	= Rendah
r : 0.21 – 0.50	= Kurang
r : 0.51 – 0.70	= Cukup
r : 0.71 – 0.90	= Tinggi
r : 0.91 – 1.00	= Sempurna

Jadi Kesimpulannya bahwa soal-soal IPS yang diujikan mempunyai tingkat reliabilitas 0.77 tersebut termasuk kategori tinggi, artinya butir-butir soalnya dapat dipercaya.

**SEKIAN DAN TERIMAKASIH**