



I. ANALISIS KESELURUHAN TES

A. Validitas

Validitas tes adalah tingkat keabsahan atau ketepatan suatu tes.

Tes dikatakan valid apabila mengukur sesuai dengan apa yang hendak diukur.

❖ Validitas Permukaan

Analisis rasional Pertimbangan rasio (bahasa, jawaban, kaidah penulisan)

- Apakah bahasa dan redaksi soal jelas dan sesuai untuk mengukur kemampuan siswa?
- Apakah isi jawaban tidak membingungkan siswa?
- Apakah tes penyusunan tes sudah mengikuti kaidah penulisan butir soal ?

❖ Validitas Isi

Kesesuaian dengan kisi-kisi ; kesesuaian dengan GBPP

❖ Validitas Kriteria

- Korelasi dengan tes lain yang sudah valid; menghitung koefisien korelasi (misal menghitung koefisien korelasi produk momen Pearson) sering disebut sebagai korelasi produk momen:

$$\text{➤ } r_{xy} = \frac{N(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[N(\sum X^2) - (\sum X)^2][N(\sum Y^2) - (\sum Y)^2]}}$$

$$\text{➤ } r_{xy} = \frac{\sum(\Delta x \cdot \Delta y)}{\sqrt{(\sum \Delta x^2)(\sum \Delta y^2)}}$$

$$\Delta x : X - \bar{X}$$

$$\Delta y : Y - \bar{Y}$$

Kriteria:

0,800 – 1,00 : sangat tinggi

0,600 – 0,79 : tinggi

0,400 – 0,59 : cukup

0,200 – 0,39 : rendah

0,000 – 0,19 : sangat rendah

❖ Validitas Ramalan

Sejauhmana hasil tes dapat meramalkan keberhasilan siswa dimasa datang?

B. Reliabilitas

❖ **Test-Retest (Tes Ulang)**

Tes dilakukan dua kali pada sekelompok siswa dengan selang waktu tertentu (agak lama), kemudian hasil keduanya dikorelasikan

❖ **Tes Paralel**

Dua tes (hanya berbeda sedikit: redaksi dan kalimat) diberikan secara paralel pada sekelompok siswa, hasilnya dikorelasikan.

❖ **Tes belah Dua (Splithalf)**

Pelaksanaan tes hanya dilakukan satu kali pada peserta tes yang sama.

(1). Ganjil-Genap

Perangkat tes dibelah dua: yang bernomor ganjil dan yang bernomor genap, kemudian dihitung korelasinya diantaranya dengan menggunakan persamaan *Spearman-Brown*:

$$r_{tt} = \frac{2r_{gg}}{1 + r_{gg}}$$

r_{tt} : koefisien reliabilitas

r_{gg} : koefisien korelasi produk momen ganjil-genap.

(2). Awal – Akhir

Perangkat tes dibelah dua berdasarkan nomor urutnya (misal ada 30 butir soal: dibelah menjadi dua bagian yaitu yang bernomor 1-15 dan yang bernomor 16- 30), selanjutnya dihitung korelasinya.

(3). Beberapa persamaan lain untuk mencari koefisien korelasi**a. Flanagan**

$$r_{11} = 2\left(1 - \frac{s_1^2 + s_2^2}{s_t^2}\right)$$

r_{11} : reliabilitas tes

s_1^2 : varians belahan pertama

s_2^2 : varians belahan kedua

s_t^2 : varians total

b. Bulon

$$r_{11} = 1 - \frac{s_d^2}{s_t^2}$$

s_d^2 : varians dari beda skor antara dua belahan

s_t^2 : varians total

c. Kuder-Richardson (K-R. 20 dan K-R.21)

➤ **K-R.20**

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

n : banyak item

p : proporsi yang menjawab benar

q : proporsi yang menjawab salah

s : varians

➤ **K - R. 21**

$$r_{11} = \left[\frac{n}{n-1}\right] \left[1 - \frac{M(n-M)}{ns_t^2}\right]$$

M : rerata skor total

RELIABILITAS TES URAIAN

Jika penskoran pada tes pilihan ganda menggunakan angka 1 (jawaban benar) dan angka 0 (jawaban salah), maka penskoran untuk tes uraian bervariasi bergantung bobot soalnya. Dengan demikian perhitungan untuk mencari korelasinya juga berbeda. Perhitungan yang biasa dilakukan yaitu dengan menggunakan persamaan *Alpha* :

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

n : banyak butir soal

$\sum s_i^2$: jumlah varians skor tiap-tiap butir soal

s_t^2 : varians total

II. ANALISIS TIAP BUTIR SOAL**A. Daya Pembeda (DP)**

Apakah dapat membedakan siswa kelompok tinggi (U) dan rendah (L)?

$$DP = \frac{B_U - B_L}{N_x}$$

B_U : jumlah jawaban benar kelompok atas

B_L : jumlah jawaban benar kelompok bawah

N_x : jumlah siswa kelompok atas atau bawah

Kriteria:

< 9% : sangat buruk (soal dibuang)

10% - 19% : Buruk (dibuang)

20% - 29% : Baik, revisi

30% - 49% : baik

> 50% : sangat baik

B. Tingkat Kesukaran (TK)

$$TK = \frac{B_U + B_L}{N_U + N_L}$$

N_U : jumlah siswa kelompok atas

N_L : jumlah siswa kelompok bawah

Kriteria:

0 - 15% : sangat sukar (dibuang)

16% - 30% : sukar

31% - 70% : sedang

71% - 85% : mudah

86% - 100% : sangat mudah (dibuang)

cara lain : $TK = \frac{N_B}{N} \times 100\%$

N_B : jumlah siswa yang menjawab benar

N : jumlah siswa seluruhnya

C. Pengecoh (Distraktor)

Indeks Pengecoh (I_p):

$$I_p = N_p \left(\frac{n-1}{N-n_B} \right)$$

N_p : jumlah siswa yang memilih pengecoh

n : banyak option (pilihan)

N : jumlah siswa yang ikut tes

N_B : jumlah siswa yang menjawab benar butir soal ybs.

Kriteria

$> 200\%$: sangat buruk

$0 - 25\%$ atau $176-200\%$: buruk

$26\%-50\%$ atau $151-175\%$: kurang baik

$51\%-75\%$ atau $126-150\%$: baik

$76\%-125\%$: sangat baik

D. Homogenitas

Korelasi antara skor (yang benar) tiap butir soal dengan skor total. Dihitung dengan “Produk momen Pearson” atau dengan “Koefisien Biserial (γ_{bis})”:

$$\gamma_{bis} = \frac{M_p - M_t}{s_t} \sqrt{\frac{p}{q}}$$

M_p : Rerata skor dari subyek yang menjawab benar untuk “butir soal” yang akan dicari korelasinya.

M_t : Rerata skor total

s_t : Standar deviasi skor total

p : proporsi siswa yang menjawab benar

q : proporsi siswa yang menjawab salah

CONTOH