

# **KARAKTERISTIK TES YANG BAIK**

**Oleh**

**Farida Agus Setiawati, M.Si**

**faridaagus@yahoo.co.id**

# Karakteristik Tes

- **Validitas alat tes**
- **Reliabilitas Hasil Pengukuran**

# Konteks Validitas

Dalam penelitian validitas digunakan dalam tiga konteks :

- Validitas Penelitian ? Bukan validitas
- Validitas Soal/Item ? Bukan validitas
- Validitas Tes/Alat ukur/Instrumen

Validitas alat ukur/tes menunjukkan bagaimana suatu alat ukur benar-benar dapat mengukur apa yang akan diukur atau diketahui

- Validitas isi membutuhkan analisis rasional dari seorang yang ahli dalam bidang yang dikembangkan alat ukur tersebut atau *profesional judgment*.
- Analisis validitas isi dapat dilakukan secara kuantitatif maupun kualitatif. Analisis secara kualitatif dilakukan berdasar masukan dari para ahli melalui diskusi atau *focus group discussion*.

# Analisis Validitas Isi secara kuantitatif

- validitas isi dari Lawshe yang dikenal dengan CVR (Lawshe, 1975: 567, Shultz & Whitney, 2005:89, Cohen&Swerdlik, 2009:178),
- validitas isi dari Aiken yang dikenal dengan V Aiken (Aiken, 1985: 132, Azwar, 2012 :134 )
- validitas isi dari Gregory (2007: 122-123)

# Formula CVR

$$\text{CVR} = \frac{n_e - (N/2)}{N/2}$$

$n_e$  = Jumlah ahli atau *subject matter experts* (SMEs)  
yang memberi respons essential pada suatu butir

$N$  = Jumlah ahli atau SMEs

Perhitungan CVR dilakukan pada tiap item. Skor CVR tiap item ini selanjutnya dibandingkan dengan skor minimal CVR dengan taraf penerimaan 0.05 sebagaimana diteliti oleh Lawshe (1975:568) sebagaimana terdapat pada tabel

**Number of Panelists****Minimum Value**

---

5	.99
6	.99
7	.99
8	.75
9	.78
10	.62
11	.59
12	.56
13	.54
14	.51
15	.49
20	.42
25	.37
30	.33
35	.31
40	.29

---

# V Aiken

- $V = S / [n * (c-1)]$ , dimana  $S = \sum ni (r-lo)$

Keterangan :

V = indeks validitas dari Aiken

ni = jumlah expert yang memilih kriteria i,

r = kriteria ke i,

lo = rating terendah,

n = jumlah semua expert,

c = banyaknya rating/kriteria

[Contoh](#)



# Formula Validitas Isi dr Gregory

		penilai 1	
		relevan	tdk relevan
penilai 2	relevan	A	B
	tdk relevan	C	D

---

$$A+B+C+D$$

Contoh :

# Validitas Konstruk

- Validitas konstruk menunjukkan sejauhmana tes mengungkap suatu konstruk teori yang diukurnya atau suatu trait yang dikembangkan dalam penyusunan instrumen (Mardapi, 2008:21).
- Konstruk merupakan suatu teori yang penting yang mendasari instrumen. Pada tes yang sama dapat dibuat beberapa instrumen dengan konstruk yang berbeda tergantung dari teori dan ahli yang mengembangkannya.
- Nunnaly (1970:141) menjelaskan ada beberapa aspek penting dalam proses mengembangkan validitas konstruk:
  - menspesifikasikan domain yang diukur dengan jelas dan dapat dilihat,
  - menentukan seberapa luas dan keterkaitan antar aspek pada variabel yang diukur,
  - menentukan keberadaan dan seberapa pentingnya variabel yang dijelaskan dari konstruk.

- Metode yang diakui para ahli untuk mengukur validitas konstruk adalah analisis faktor dan multi trait multi method. Analisis faktor dibedakan :
  - Analisis faktor *explanatory* adalah teknik untuk mengetahui atau mencari konstruk suatu instrumen dengan cara mereduksi banyaknya data atau variabel.
  - Analisis faktor konfirmatori adalah teknik digunakan untuk mengeksplorasi data empiris untuk mengetahui hubungan antar variabel di data lapangan. Analisis ini digunakan apabila konstruk suatu instrumen sudah jelas dan peneliti ingin menguji hipotesis (kebenaran) konstruk yang dibuat.

# Validitas Kriteria

- Validitas kriteria adalah kemampuan suatu tes untuk memprediksikan performansi pada pengukuran yang lain (Friedenberg, 1995:225).
- Kriteria tersebut merupakan perilaku yang dapat digunakan untuk memprediksi skor tes. Kriteria yang dimaksudkan dapat berupa alat tes lain yang standar yang mengukur atribut yang sama yang disebut sebagai validitas konkuren
- Kriteria prestasi yang akan datang yang menunjukkan performansi yang diukur, misalnya IP, nilai ujian akhir yang disebut validitas prediktif.
- Validitas kriteria dihitung berdasarkan korelasi antara skor tes yang dibuat dengan skor kriteria yang dihubungkan

# Faktor-faktor yang mengancam validitas

- Identifikasi kawasan ukur yang tidak jelas
- Operasionalisasi konsep yang tidak tepat
- Penulisan item yang tidak mengikuti kaidah
- Administrasi tes yang tidak hati-hati
  - Kondisi penampilan tes
  - Kondisi subjek
  - Kondisi testing
- Pemberian skor yang tidak cermat
- Interpretasi yang keliru

# Reliabilitas

- Konsep reliabilitas muncul dari asumsi dalam teori tes klasik yang terkait dengan skor tampak ( $X$ ), skor murni ( $T$ ), dan skor kesalahan pengukuran ( $E$ ). Konsep utama dalam teori ini adalah  $X = T + E$ , atau skor tampak merupakan gabungan dari skor murni dan kesalahan pengukuran
- Reliabilitas terkait dengan variasi error, data hasil pengukuran yang tidak konsisten berakibat munculnya error yang besar.

Reliabilitas menunjukkan sejauhmana hasil pengukuran dengan alat tes tersebut konsisten, ajeg sehingga hasilnya dapat dipercaya (Suryabrata, 2000)

- Pendekatan perhitungan reliabilitas
- Pendekatan tes ulang
- Pendekatan tes paralel
- Pendekatan satu kali pengukuran



# Pendekatan tes ulang

- Alat tes diberikan pada sekelompok subjek dua kali, dengan selang waktu tertentu, misalnya selang dua minggu.
- Reliabilitas diukur dari korelasi skor pada tes 1 dan tes 2
- Pendekatan ini mengandung kelemahan karena kondisi subjek pada tes 1 dan ke 2 bisa berubah karena unsur belajar, pengalaman, motivasi dll.
- Sesuai digunakan untuk pengukuran ketrampilan, terutama ketrampilan fisik

# Contoh

Subjek	Skor 1	Skor 2
1	50	50
2	40	50
3	40	40
4	50	50
5	60	60
6	70	60
7	50	40
8	60	50
9	60	60
10	70	70

# Pendekatan tes paralel

- Dua perangkat tes yang paralel, misalnya perangkat 1 dan 2 diberikan pada sekelompok subjek.
- Reliabilitas tes diukur dengan menghitung korelasi antara skor perangkat 1 dan 2
- Keterbatasan perangkat ini adalah sulitnya menyusun tes yang paralel

# contoh

Subjek	Perangkat 1	perangkat 2
1	45	50
2	40	50
3	45	40
4	50	55
5	60	60
6	70	70
7	50	40
8	60	60
9	60	60
10	70	75

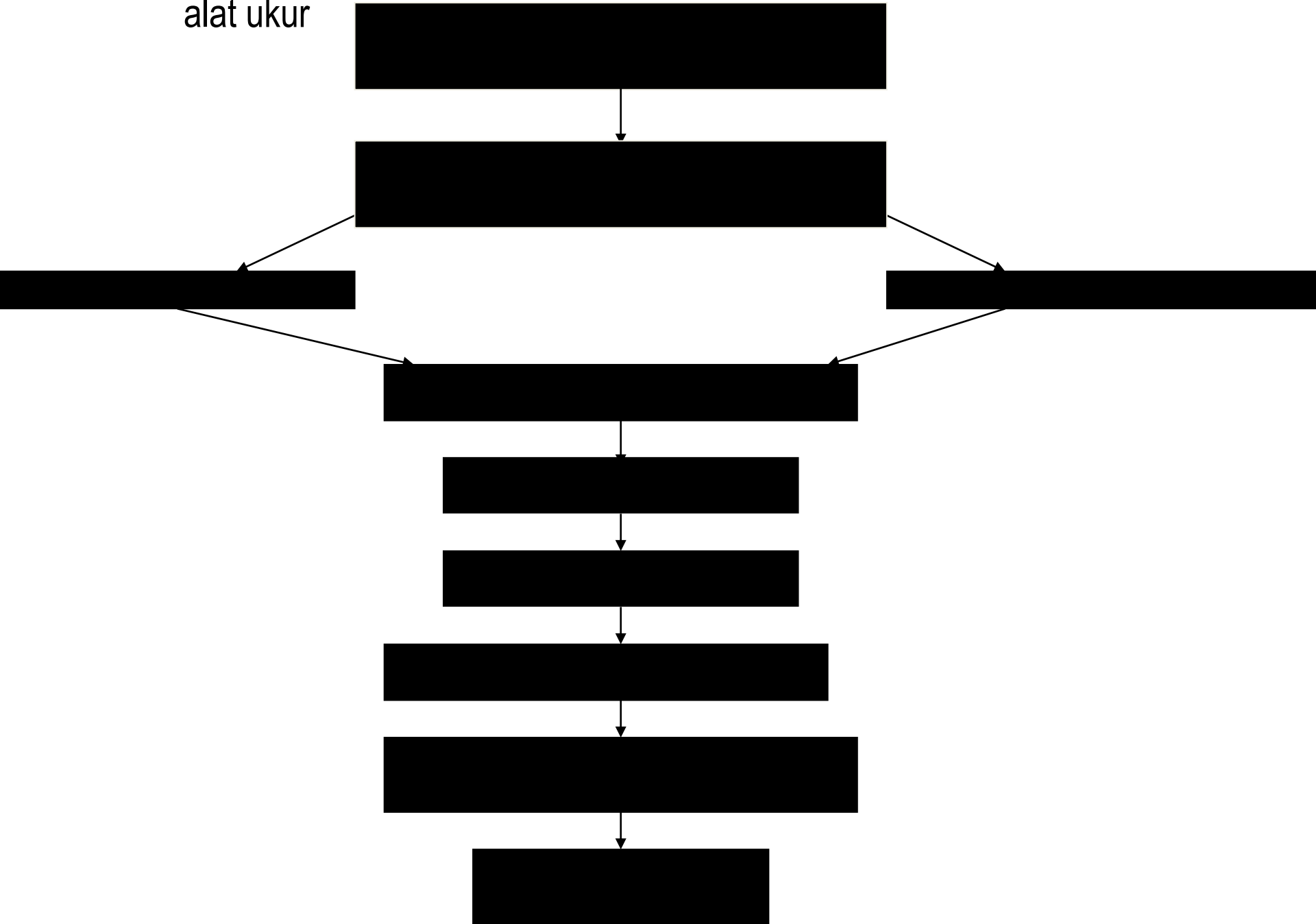
## Pendekatan satu kali pengukuran

- Alat tes diberikan pada sekelompok subjek sekali lalu dengan cara tertentu dihitung reliabilitas tes tersebut . Teknik/Cara yang digunakan :
- Teknik belah dua (rumus rulon, flanagan)
- Teknik kuder
- Teknik analisis varians (rumus Hoyt, rumus Alpha)

# Penyusunan Instrumen

- Merumuskan tujuan pengukuran
- Spesifikasi Instrumen
- Draft instrumen
- Uji Ahli
- Revisi
- Ujicoba
- Format akhir

# Prosedur Pengembangan alat ukur



# Analisis Hasil

- Validitas isi
- Seleksi Item (Hasil uji lapangan)

Instrumen tes :

- Index kesukaran soal
- Daya beda
- kualitas pengecoh/distraktor

Reliabilitas



# Index Kesukaran Soal

- Index kesukaran soal ( $p$ ) adalah jumlah subjek yang menjawab benar pada suatu item/soal ( $B$ ) dibagi dengan jumlah subjek yang mengerjakan soal ( $T$ ),
- $$p = B/T$$
- Nilai  $p$  antara 0 hingga 1, 0 menunjukkan soal tsb sulit, semakin mendekati 1 menunjukkan soal tersebut mudah . Dengan demikian pada  $p$  pada hakekatnya index kemudahan item.

# Daya Beda

- Diartikan sebagai apakah item tersebut dapat membedakan orang yang tinggi kemampuannya dan orang yang rendah kemampuannya.
- Pengujian semula dilakukan dengan membandingkan skor pada kelompok atas dan kelompok bawah, kelompok atas diambil 27% dan kelompok bawah juga diambil 27%.
- Selanjutnya dianalisis secara statistic dengan uji beda. :

- Daya beda ini selanjutnya dianalisis secara statistic dengan rumus rbis dan rpbis
- Pada hakekatnya untuk mengetahui item tersebut dijawab benar pada subjek yang kemampuannya tinggi dan rendah atau membandingkan subjek yang menjawab item tersebut dengan benar dibandingkan dengan subjek yang menjawab item dengan salah.
- Rumus statistik untuk menghitung rbis dan rpbis sebagai berikut : 
$$rbis = \frac{Xb - Xs}{St} \times \frac{y}{x} \quad rpbis = \frac{Xb - Xs}{St} \times \sqrt{pq}$$

$$\frac{2r}{1+r} = \frac{2r}{1+r}$$

# Rumus Reliabilitas

- Rumus Spearman Brown
- Rumus Rulon
- Rumus Flanagan
- Rumus Kuder

KR20

KR21

- Rumus Hoyt
- Rumus Alpha

# Rumus Reliabilitas alpha

$$\frac{n}{n-1} \left( 1 - \frac{\sum \sigma_i^2}{\sigma^2} \right)$$

- $\alpha$  = koefisien reliabilitas alpha
- $\sigma_i^2$  = varian skor subjek bagian i
- $\sigma^2$  = varians skor subjek pada keseluruhan tes.
- $n$  = banyaknya bagian (potongan tes)