

# **MENGUJI VALIDITAS DAN RELIABILITAS**

Oleh: Ali Muhson, M.Pd.

1

## **Validitas Instrumen**

- ▣ Kevalidan sebuah alat ukur ditunjukkan dari kemampuan alat ukur tersebut mampu mengukur apa yang seharusnya diukur.

2

## Jenis Validitas

- ▣ Content Validity
  - Apakah alat ukur telah cukup mengukur sebuah konsep?
- ▣ Face Validity
  - Apakah “ahli” mengesahkan bahwa instrumen telah mengukur apa yang seharusnya diukur?
- ▣ Concurrent Validity
  - Apakah alat ukur dapat membantu dalam memprediksi variabel kriteria saat ini?
- ▣ Predictive Validity
  - Apakah alat ukur dapat membantu dalam memprediksi variabel kriteria masa depan?

3

## Jenis Validitas

- ▣ Construct Validity
  - Apakah alat ukur telah sesuai dengan konsep teori?
- ▣ Convergent Validity
  - Apakah kedua instrumen dalam mengukur konsep berkorelasi tinggi?
- ▣ Discriminant Validity
  - Apakah alat ukur memiliki korelasi yang rendah dengan variabel yang seharusnya tidak berhubungan dengannya?

4

## Langkah dalam pengembangan instrumen:

1. *Menentukan variabel yang diteliti*
2. *Menentukan definisi operasional variabel berdasarkan teori*
3. *Menentukan konstruk dan indikator*
4. *Membuat kisi-kisi*
5. *Menyusun butir pertanyaan*

5

## Contoh Pengembangan Instrumen

Variabel Penelitian	Konstruk
Minat jadi guru	<ul style="list-style-type: none"><li>• Pengetahuan dan informasi mengenai profesi guru</li></ul>
	<ul style="list-style-type: none"><li>• Perasaan senang dan ketertarikan terhadap profesi guru</li></ul>
	<ul style="list-style-type: none"><li>• Perhatian yang lebih besar terhadap profesi guru</li></ul>
	<ul style="list-style-type: none"><li>• Kemauan dan hasrat untuk menjadi guru</li></ul>

6

## Uji Validitas Instrumen

Construct Validity (Allen & Yen, 1979)

- ▣ *Exploratory Factor Analysis (EFA)*
  - ▣ Valid jika factor loading melebihi 0,3
  - ▣ Nilai cumulative percentage minimal 50%
- ▣ *Confirmatory Factor Analysis (CFA)*
  - Valid jika factor loading melebihi 0,3
  - Valid jika Convergent Validity melebihi 0,5

**Content Validity** (Shultz & Whitney, 2005; Coaley, 2010) :

- ▣ *Expert Judgment*
- ▣ *CVI (Content Validity Index)*

7

## Analisis Butir Soal

- ▣ Melalui Analisis Butir
- ▣ Hal-hal yang dianalisis:
  - Daya beda butir
  - Tingkat kesukaran butir
  - Efektivitas distraktor

8

## Analisis Tingkat Kesukaran

- ▣ Apakah soal terlalu mudah atau terlalu sulit?
- ▣ Soal yang baik adalah yang memiliki tingkat kesukaran sedang
- ▣ Cara menghitungnya: proporsi siswa yang dapat menjawab dengan benar
- ▣ Soal yang baik memiliki indeks tingkat kesukaran antara 0,3 - 0,7

9

## Indeks Tingkat Kesukaran

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

Keterangan:

P = Taraf kesukaran butir

B = Jumlah peserta yang menjawab benar

N = Jumlah seluruh peserta

10

## ANALISIS DAYA BEDA

- ▣ Kemampuan butir untuk membedakan antara siswa yang pandai dan tidak pandai.
- ▣ Apakah siswa yang pandai sama benarnya dalam menjawab soal dengan siswa yang tidak pandai?
- ▣ Butir memiliki daya beda yang baik jika indeks daya bedanya lebih dari 0,3

11

## Indeks Daya Beda Sederhana

$$D = \frac{B_t - B_r}{\frac{1}{2}(n_t + n_r)}$$

$$D = \frac{B_t}{B_t + B_r}$$

Keterangan:

D = Daya beda butir

$B_t$  = Jumlah jawaban benar dari kelompok tinggi

$B_r$  = Jumlah jawaban benar dari kelompok rendah

$n_t$  = Jumlah peserta kelompok tinggi

$n_r$  = Jumlah peserta kelompok rendah

12

## Indeks daya beda dengan korelasi Biserial

$$r_b = \left( \frac{\bar{X}_p - \bar{X}_q}{\sigma_x} \right) \left( \frac{pq}{Y} \right)$$

Keterangan:

- $\bar{X}_p$  = Skor rata-rata yang menjawab benar
- $\bar{X}_q$  = Skor rata-rata yang menjawab salah
- $\sigma_x$  = Standar deviasi skor
- p = Proporsi peserta yang menjawab benar
- q = Proporsi peserta yang menjawab salah
- Y = Ordinat kurve normal yang membagi menjadi p dan q

13

## Indeks daya beda dengan korelasi Point Biserial

$$r_{pb} = \frac{\bar{X}_p - \bar{X}_q}{\sigma_x} \sqrt{pq}$$

- ▣ Koefisien korelasi *point biserial* hasilnya sama dengan koefisien korelasi *product moment*

14

## Efektivitas Distraktor

- ▣ Kemampuan faktor pengecoh untuk menjebak peserta tes.
- ▣ Soal yang baik adalah yang mampu mengecoh peserta tes yang kurang pandai.
- ▣ Jangan membuat pengecoh yang terlalu jauh dari alternatif jawaban yang benar
- ▣ Pengecoh efektif jika proporsi peserta tes yang terkecoh lebih dari 5%

15

## Contoh Item Analysis Sederhana

Nama Siswa	Nomor Butir										Jml
	1	2	3	4	5	6	7	8	9	10	
Jumlah	10	7	7	7	7	5	5	7	4	6	60
Persen	100	70	70	20	70	50	50	70	40	60	60

Butir paling mudah

Butir paling sulit

Butir dengan daya beda baik

Butir dengan daya beda jelek

16



## Reliabilitas Instrumen

- ▣ Pengertian reliabilitas pada dasarnya adalah sejauh mana hasil suatu pengukuran dapat dipercaya atau menghasilkan sesuatu yang bersifat ajeg/ tetap.

17

## Uji Reliabilitas Instrumen

- ▣ Teknik Paralel (*parallel form*)  
Pada teknik ini kita membagi kuesioner kepada responden yang intinya sama akan tetapi menggunakan kalimat yang berbeda:  
Misalnya:
  - Apakah menurut saudara harga tiket di kereta ini tidak mahal ?
  - Apakah harga di kereta ini telah sesuai dengan pelayanan yang saudara terima ?

18

## Uji Reliabilitas Instrumen

- ▣ Teknik Ulang (*double test / test retest*)

Pada teknik ini kita membagi kuesioner yang sama pada waktu yang berbeda.

Misalnya:

- Pada minggu I ditanyakan:
  - ▣ Bagaimana tanggapan saudara terhadap kualitas dosen di Universitas ini?
- Pada minggu III ditanyakan:
  - ▣ Ditanyakan lagi pada responden yang sama dengan pertanyaan yang sama.

19

- ▣ Reliabilitas Internal (*Internal Consistency*)

- ▣ Uji reliabilitas internal digunakan untuk menghilangkan kelemahan-kelemahan pada uji reliabilitas eksternal.

1. Dengan rumus *Spearman-Brown*
2. Dengan rumus *Flanagant*
3. Dengan rumus *Rulon*
4. Dengan rumus  $K - R_{20}$
5. Dengan rumus  $K - R_{21}$
6. Dengan rumus *Hoyt*
7. Dengan rumus Alpha Cronbach

20

## **Kriteria Reliabilitas**

- ▣ The generally accepted minimum standard is 0,65 (Ebel & Frisbie, 1991: 86)
- ▣ The common standard of a reliability estimate of at least .70 or higher holds for alpha as well (Shultz & Whitney, 2005: 74).
- ▣ Above 0.7 is usually considered to offer reasonable reliability for research purposes (Muijs, 2004: 73)

21

## **Langkah dalam membuat instrumen yang valid dan reliabel**

- ▣ Buatlah instrumen dengan mengikuti prosedur yang benar secara teoretik, yaitu
  - Membuat definisi operasional variabel berdasarkan teori
  - Menentukan konstruk dan indikatornya
  - Membuat kisi-kisi instrumen dan butir pertanyaan
- ▣ Menguji instrumen tersebut secara empirik dengan cara:
  - Uji coba kepada paling sedikit 30 orang responden (batas sampel besar dalam statistik)
  - Tabulasi data yang telah masuk
  - Ujilah validitas dan reliabilitasnya

22