

Pertemuan 7

VALIDITAS DAN RELIABILITAS INSTRUMEN

Tujuan

Setelah perkuliahan ini diharapkan dapat:

- Menjelaskan tentang pengertian 'validitas' dan penerapannya dalam menguji instrument penelitian pendidikan.
 - Menjelaskan tentang 'koefisien korelasi' dan mendeskripsikan perbedaan antara korelasi positif dan negatif.
 - Menjelaskan tentang 'koefisien validitas' dan 'koefisien reliabilitas'
 - Menjelaskan pengertian 'reliabilitas' dan penerapannya untuk menguji instrument penelitian pendidikan
 - Menjelaskan tentang 'error pengukuran'.
-

- ✓ Validitas (kesahihan) berkaitan dengan kelayakan, kebermaknaan, dan kegunaan dari kesimpulan hasil penelitian
- ✓ Reliabilitas berkaitan dengan konsistensi (keajegan) skor atau jawaban dari masing-masing soal dalam instrumen bila digunakan untuk kelompok lain.

VALIDITAS

Validitas merupakan hal yang sangat penting dalam pengujian instrumen penelitian, karena penarikan kesimpulan yang tepat didasarkan pada data yang yang diperoleh peneliti dengan menggunakan instrumen. Akhir-akhir ini, validitas didefinisikan sebagai kesahihan, kelayakan, kebermaknaan, dan kegunaan inferensi spesifik peneliti berdasarkan data yang diperolehnya. Jadi validasi instrumen adalah proses mengumpulkan bukti untuk mendukung inferensi.

Inferensi yang layak harus relevan, yaitu berkaitan dengan tujuan penelitian. Jika tujuan penelitian adalah untuk mengetahui sejauh mana wawasan siswa terhadap budaya Sunda misalnya, maka inferensinya didasarkan pada nilai tes tentang geografi fisik Jawa Barat.

Inferensi yang bermakna, berkaitan dengan makna, misalnya skor tes yang diperoleh dengan menggunakan instrumen. Apa makna dari skor tinggi, sedang, dan rendah yang diperoleh siswa melalui instrumen tes? Tujuan penelitian adalah mengumpulkan informasi dari subjek penelitian. Jika informasi yang diperoleh hanya tentang nama, alamat, gender, tanggal lahir, nomor KTP, maka informasi ini kurang bermakna karena peneliti tidak dapat menarik inferensi berdasarkan data tersebut.

Perlu diingat bahwa tujuan penelitian bukan hanya mengumpulkan data, melainkan menggunakannya untuk menarik kesimpulan tentang data yang terkumpul dari subjek penelitian.

Inferensi yang berguna, adalah inferensi yang dapat membantu peneliti membuat keputusan tentang jawaban yang akan dicarinya. Seorang peneliti yang berminat menyelidiki tentang pengaruh pengajaran berbasis inkuiri terhadap hasil belajar siswa, misalnya, memerlukan informasi yang memungkinkan dia menarik kesimpulan. Seandainya hasil belajar siswa terbukti dipengaruhi oleh pengajaran berbasis inkuiri, lalu bagaimana kelanjutannya?

Validitas sangat tergantung pada jumlah dan jenis pembuktian yang akan mendukung interpretasi peneliti terhadap data yang telah dikumpulkannya. Pembuktian seperti apa? Ada tiga jenis pembuktian.

1. Pembuktian validitas isi (*Content-related evidence of validity*)
2. Pembuktian validitas kriteria (*Criterion-related evidence of validity*)
3. Pembuktian validitas konstruk (*Construct-related evidence of validity*)

1. Pembuktian validitas isi (*Content-related evidence of validity*)

Misalnya, seorang peneliti tertarik untuk meneliti pengaruh program matematika realistik terhadap kemampuan matematika siswa kelas 5. Untuk melengkapi program ini, peneliti berharap siswa dapat memecahkan sejumlah soal ceritera yang berbeda-beda secara benar. Untuk mengases kemampuan matematika, maka peneliti memberikan tes berupa 15 soal matematika yang mengandung 15 masalah berbeda.

Kinerja siswa pada tes ini sangat penting untuk melihat tingkat kemampuan siswa dalam memecahkan masalah. Untuk itu soal tes yang diberikan harus memiliki validitas untuk mengukur kemampuan siswa yang diharapkan oleh program. Jika soal tes hanya terdiri dari soal yang mudah saja, atau yang sukar saja, atau soal dengan kalimat yang panjang saja, maka tes ini tidak representatif, dan tidak memberi informasi untuk membuat inferensi yang valid.

Salah satu unsur pembuktian validitas isi, adalah kecukupan sampel. Kebanyakan instrumen (terutama tes

hasil belajar) hanya memberikan satu sampel masalah untuk dipecahkan atau satu bentuk pertanyaan untuk dijawab, sehingga kurang menggambarkan keutuhan penelitian.

Aspek lain yang perlu diperhatikan dalam validasi isi adalah format instrumen, yang meliputi kejelasan cetakan, ukuran huruf, ruang yang cukup untuk menjawab (bila diperlukan), dan keterbacaan soal. Untuk itu maka soal yang disusun berdasarkan kisi-kisi, perlu diuji coba kepada sejumlah sampel di luar sampel penelitian, kemudian dianalisis.

2. Pembuktian validitas kriteria (*Criterion-related evidence of validity*)

Untuk mendapatkan pembuktian validitas kriteria, peneliti biasanya membandingkan bentuk instrumen yang satu dengan instrumen lainnya, kriteria bebas. Kriteria adalah ujian kedua bagi instrumen agar dapat diukur. Misalnya suatu instrumen dirancang untuk mengukur kemampuan akademis, maka skor siswa berdasarkan instrumen tersebut harus dibandingkan dengan IPK mereka (kriteria eksternal). Jika instrumen benar-benar mengukur kemampuan akademis, maka siswa yang mendapat nilai tinggi untuk tes tersebut, seharusnya juga tinggi IPK-nya.

Ada dua bentuk validitas kriteria yaitu: *predictive* dan *concurrent*. Untuk memperoleh pembuktian *predictive validity*, peneliti harus menyediakan selang waktu antara pencatatan instrumen dan perolehan skor kriteria. Contohnya: seorang peneliti harus membukukan hasil (skor) tes keterampilan sains siswa SMA yang diberikan pada awal semester untuk melakukan prediksi. Kemudian skor keterampilan sains ini dibandingkan dengan hasil tes akhir semester untuk mata pelajaran sains, inilah pembuktian *predictive validity*.

Apabila data instrumen dan data kriteria telah terkumpul pada waktu yang hampir bersamaan, lalu dibandingkan hasilnya, maka peneliti akan memperoleh pembuktian *concurrent validity*. Contohnya: Bila seorang peneliti melakukan tes inventory pada sekelompok siswa kelas 8 dan kemudian membandingkannya dengan hasil penilaian 'self esteem' yang dilakukan oleh gurunya, di mana kedua penilaian ini dilakukan pada waktu yang hampir bersamaan.

Pengukuran indeks dari kedua validitas kriteria tersebut dilakukan dengan koefisien korelasi (r), yang menunjukkan

derajat hubungan antara skor individual yang diperoleh melalui dua instrumen tersebut. Hubungan positif menunjukkan bila seorang siswa mencapai skor tinggi pada instrumen dan ia mendapat skor tinggi pula untuk instrumen lain. Atau bila ia mendapat skor rendah pada instrumen, dan mendapat skor rendah pula pada instrumen lain. Hubungan negatif menunjukkan bila seorang siswa mendapat nilai tinggi untuk satu instrumen, tetapi mendapat skor rendah untuk instrumen lainnya, atau sebaliknya. Koefisien korelasi berada pada kisaran angka +1,00 dan -1,00; $r=0$ menunjukkan tidak ada hubungan. Koefisien korelasi (r) disebut pula koefisien validitas

3. Pembuktian validitas konstruk (*Construct-related evidence of validity*)

Pembuktian validitas konstruk merupakan yang terluas penggunaannya di antara tiga kategori validitas. Ada tiga tahap untuk memperoleh pembuktian validitas konstruk:

- (1) variabel yang diukur didefinisikan dengan jelas
- (2) hipotesis, berdasarkan teori yang melandasi variabel
- (3) hipotesis diuji secara logis dan empiris

Agar prosesnya jelas, sebagai contoh: Seandainya peneliti berniat mengembangkan tes tertulis untuk mengukur “kejujuran” dan bermaksud menggunakan pembuktian validitas konstruk. Maka yang harus dilakukannya adalah:

- Mendefinisikan tentang “kejujuran”
- Merumuskan teori tentang kejujuran seseorang dibandingkan dengan yang tidak jujur, sehingga jelas perbedaan antara jujur dan tidak jujur.
- Berdasarkan teori ini peneliti menyusun hipotesis, bahwa individu yang mendapatkan skor tinggi dalam ‘tes kejujuran’ akan lebih tergerak untuk mengembalikan barang milik orang lain yang ditemukannya dibandingkan dengan individu yang mendapatkan skor rendah.

Untuk memenuhi tahap tersebut, maka peneliti mencatat hasil tes ‘kejujuran’, memisahkan siswa yang mendapat skor tinggi dan skor rendah. Kemudian ia melakukan tes lainnya, misalnya ia meninggalkan sebuah dompet berisi uang Rp 50.000,- di kelas. Di dalam dompet terdapat kartu nama pemiliknya lengkap dengan alamat dan no HP. Jika hipotesis penelitiannya benar, maka individu yang menemukan dompet

dan mengembalikannya adalah individu yang mendapat skor tinggi untuk tes 'kejujuran'

Perlu ditekankan di sini bahwa peneliti hendaknya melaksanakan serangkaian studi untuk memperoleh berbagai bukti bahwa skor instrumen tertentu dapat menghasilkan inferensi yang benar mengenai variabel yang diukur oleh instrumen tersebut

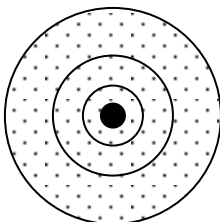
RELIABILITAS

Reliabilitas berkaitan dengan keajegan atau konsistensi dari skor yang diperoleh, yaitu bagaimana konsistensinya antara setiap individu yang dites oleh instrumen tersebut. Contohnya, suatu tes dirancang untuk mengukur logika berpikir. Bila tes ini reliable, maka dapat diharapkan bila siswa mendapat skor tinggi pada tes pertama, iapun akan mendapatkan skor tinggi pula pada kesempatan lain bila ia mengambil tes tersebut. Mungkin skornya tidak identik, tetapi hampir sama.

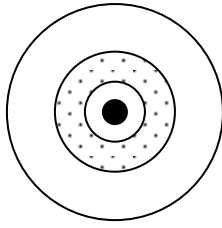
Skor yang diperoleh dari suatu instrumen dapat reliable, tetapi belum tentu valid. Andaikata seorang peneliti memberikan dua macam tes (A dan B) kepada sekelompok siswa kelas 10 untuk mengukur pemahaman mereka tentang "UUD 1945 dan Pancasila" lalu mendapatkan bahwa skor tes mereka konsisten: artinya siswa yang mendapat skor tinggi pada Tes A juga mendapat skor tinggi pada Tes B; dan siswa yang mendapat skor rendah pada tes A juga mendapat skor rendah pada tes B, Maka dapat dikatakan bahwa skor tersebut reliabel. Tetapi bila peneliti menggunakan skor tes yang sama untuk melakukan prediksi keberhasilan siswa dalam pendidikan jasmani, maka ia akan menemukan hasil yang sangat berbeda. Setiap inferensi tentang keberhasilan siswa dalam pendidikan jasmani yang didasarkan pada tes "UUD 1945 dan Pancasila" tidak memiliki validitas.

Bagaimana bila sebaliknya? Apakah instrumen yang menghasilkan skor yang tak reliabel dapat menghasilkan inferensi yang valid? Jawabnya : tidak! Jika skor tidak konsisten untuk seseorang, maka tidak dapat memberikan informasi yang berguna.

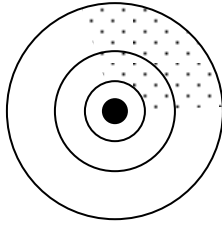
Ada beberapa kemungkinan mengenai ketergantungan antara reliabilitas dan validitas, seperti dijelaskan berikut:



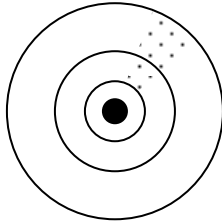
- a. Tidak reliabel dan tidak valid: Jika data tidak reliabel, maka tidak dapat menghasilkan inferensi yang valid



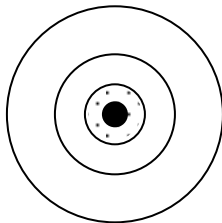
b. Cukup reliabel dan cukup valid: Jika reliabilitas meningkat, maka validitas akan meningkat pula.



c. Cukup reliabel tetapi tidak valid: peningkatan reliabilitas tidak diikuti oleh validitas



d. Reliabilitas tinggi tetapi tak-valid: suatu instrumen dapat sangat reliabel tetapi tak-valid



e. Reliabilitas dan validitas tinggi: sesuai dengan harapan

EROR DALAM PENGUKURAN

Bila seseorang mengikuti tes yang sama dua kali, maka jarang mendapat skor yang tepat sama – artinya skor jawabannya tidak identik. Hal ini disebabkan oleh beberapa faktor misalnya perbedaan motivasi, energi, rasa cemas, perbedaan situasi dll. Faktor-faktor ini menyebabkan error dalam pengukuran. Oleh karena error dalam pengukuran selalu muncul, maka peneliti berharap ada variasi skor tes (jawaban atau rating). Suatu instrumen dapat diberikan kepada kelompok yang sama lebih dari satu kali dengan format yang berbeda, dapat pula digunakan dua bentuk instrumen yang berbeda seluruhnya atau berbeda sebagian. Koefisien reliabilitas juga mengungkapkan hubungan, antara skor dari individu yang sama dengan instrumen yang sama pada dua kesempatan yang berbeda. Salah satu cara memperoleh koefisien reliabilitas adalah dengan metode ‘test-retest; metode bentuk ekuivalen; dan metode konsistensi internal. koefisien reliabilitas memiliki kisaran nilai 0,00 -1,00.

METODE TEST-RETEST

Dalam metode ini tes diberikan dua kali pada kelompok yang sama setelah interval waktu tertentu. Setelah itu, koefisien reliabilitas dihitung untuk mengetahui hubungan antara dua perangkat skor yang diperoleh.

Koefisien reliabilitas akan dipengaruhi oleh panjangnya rentang waktu antara dua tes, semakin panjang rentang interval, semakin rendah koefisien reliabilitasnya karena adanya perubahan pada individu peserta tes. Oleh karena itu untuk pembuktian reliabilitas tes-retest interval waktu ini harus diperhitungkan oleh peneliti.

Peneliti jangan berharap variabel akan selalu stabil. Pengalaman menunjukkan ada kemampuan tertentu seperti mengarang, yang lebih mudah berubah dibandingkan kemampuan lain seperti penalaran abstrak. Karakteristik pribadi seperti '*self esteem*' cenderung lebih stabil dibandingkan dengan minat (pada remaja). Untuk penelitian pendidikan rentang waktu 2-3 bulan umumnya cukup untuk melakukan test-retest terhadap individu atau kelompok.

METODE BENTUK EKIVALEN

Metode ini menggunakan dua bentuk instrumen tes yang ekuivalen, biasa disebut tes paralel atau tes alternatif yang diberikan pada kelompok atau individu yang sama pada rentang waktu tertentu. Sekalipun pertanyaannya berbeda namun isinya sama. Koefisien reliabilitas antara dua skor ini kemudian dihitung, bila koefisien reliabilitasnya tinggi berarti kedua bentuk tes ini mengukur materi yang sama.

Peneliti dapat pula mengombinasikan metode test-retest dan bentuk ekuivalen dengan memberikan dua bentuk berbeda dari tes yang sama pada interval waktu tertentu. Koefisien reliabilitas menunjukkan bahwa kedua bentuk soal tersebut mengukur kinerja yang sama dan dapat diharapkan memiliki konsistensi sepanjang waktu.

METODE KONSISTENSI INTERNAL

Metode konsistensi internal terdiri atas beberapa prosedur sebagai berikut:

1. *Split-Half Procedure*

Dalam prosedur ini penskoran dilakukan separuh-separuh (item ganjil dan item genap) dari suatu tes secara terpisah untuk setiap individu, dan kemudian dihitung koefisien korelasi antara dua perangkat skor tersebut.

Koefisien reliabilitas dapat dihitung dengan *rumus Spearman-Brown*, yang disederhanakan sebagai berikut:

$$\text{Reliabilitas skor tes total} = \frac{2 \times \text{reliabilitas untuk } \frac{1}{2} \text{ test}}{1 + \text{reliabilitas untuk } \frac{1}{2} \text{ test}}$$

Jadi bila koefisien korelasi 0,56 dengan membandingkan separuh soal dengan separuh soal lainnya, maka reliabilitas skor untuk keseluruhan tes adalah:

$$\text{Reliabilitas skor tes total} = \frac{2 \times 0,56}{1 + 0,56} = \frac{1,12}{1,56} = \mathbf{0,72}$$

Rumus lainnya adalah dengan *Pendekatan Kuder-Richardson*. Metode yang banyak digunakan untuk menentukan konsistensi internal adalah metode Kuder-Richardson, terutama KR20 dan KR21. Rumus KR21 memerlukan 3 informasi yaitu: (1) jumlah soal tes, (2) rerata (mean), dan (3) simpangan baku (SD). rumus ini digunakan bila soal tes diasumsikan memiliki tingkat kesukaran yang seimbang.

$$\text{Koefisien reliabilitas KR21} = \frac{K}{K-1} \left[1 - \frac{M(K-M)}{K(SD^2)} \right]$$

Contoh, jika $K=50$; $M= 40$; $SD=4$

$$\begin{aligned} \text{Reliabilitas} &= \frac{50}{49} \left[1 - \frac{40(50-40)}{50 \cdot 4^2} \right] \\ &= 1,02 \left[1 - \frac{40(10)}{50(16)} \right] \\ &= 1,02 \left[1 - \frac{400}{800} \right] \\ &= (1,02)(1-0,50) \\ &= (1,02)(0,50) \\ &= 0,51 \qquad \rightarrow \text{Reliabilitas} = 0,51 \end{aligned}$$

Rumus koefisien Alpha (α)

Rumus koefisien Alpha disebut juga *Cronbach Alpha* sesuai dengan nama penemunya, merupakan bentuk umum KR20.

Untuk jelasnya, beda validitas dan reliabilitas dapat dikhtisarkan pada tabel berikut :

Validitas (<i>truthfulness</i> / kesahihan)	
Metode	Prosedur
Pembuktian validitas isi	Pertimbangan ahli
Pembuktian validitas kriteria	Berkaitan dengan pengukuran lain dari variabel yang sama
Pembuktian validitas konstruk	Prediksi berdasarkan teori

Reliabilitas (konsistensi)			
Metode	Isi	Interval waktu	Prosedur
Test-retest	Identik	Bervariasi	Memberikan instrumen yang sama dua kali
Bentuk ekivalen	Berbeda	Tidak ada	Memberikan dua bentuk instrumen
Bentuk retest-ekivalen	Berbeda	Bervariasi	Memberikan dua bentuk instrumen dengan interval waktu tertentu
Konsistensi internal	Berbeda	Tidak ada	Membagi instrumen menjadi dua (nomor ganjil dan genap) dan diuji dengan rumus KR
Kesepakatan dengan observer	Identik	Tidak ada	Membandingkan skor yang diperoleh dari dua observer atau lebih.

Kesimpulan

- Apa yang dimaksud dengan 'validitas' dan bagaimanakah penerapannya ?
- Apa yang dimaksud dengan 'koefisien korelasi' dan apa makna perbedaan antara korelasi positif dan negatif?.
- Apa perbedaan antara validitas dan reliabilitas?
- Beri penjelasan tentang perbedaan antara 'koefisien validitas' dan 'koefisien reliabilitas'
- Bagaimana penerapan 'reliabilitas' dalam pengujian instrumen penelitian pendidikan
- Apa yang dimaksud dengan 'error pengukuran'