

# Modul 7

## VALIDITAS DAN RELIABILITAS INSTRUMEN

Untuk membantu Anda menguasai hal tersebut di atas dalam modul ini akan disajikan pembahasan dan latihan dalam butir-butir uraian sebagai berikut :

1. Validitas mencakup: Validitas Internal (Validitas isi, Validitas kriteria, Validitas konstruk) dan Validitas eksternal termasuk cara pengujiannya.
2. Reliabilitas mencakup: Pengertian Reliabilitas dan Cara-Cara Mengukur Reliabilitas termasuk Pengukuran Konsistensi eksternal melalui Metode test-retest dan Metode Bentuk-Bentuk Equivalent serta Pengukuran Konsistensi Internal Prosedur Bagi Dua (split half) dan Pendekatan-Pendekatan Kuder-Richardson ((Kuder-Richardson Approaches).

Agar Anda lebih berhasil dalam mempelajari modul ini ikuti petunjuk belajar berikut ini :

1. Bacalah dengan cermat bagian Pendahuluan modul ini sampai Anda memahami betul apa, dan bagaimana mempelajari modul ini.
2. Bacalah sepintas seluruh bagian modul ini dan carilah konsep-konsep yang bersifat prinsip. Bila ada kata - kata baru, pahami terlebih dahulu dengan melihat kamus atau daftar glosari pada modul ini
3. Pelajari pengertian demi pengertian dari isi modul ini melalui pemahaman sendiri atau bertukar pikiran dengan teman guru
4. Terapkan prinsip-prinsip yang telah Anda peroleh dalam situasi yang mungkin Anda temukan dalam kehidupan sehari – hari.
5. Mantapkan pemahaman Anda melalui diskusi kelompok

## BAB I : VALIDITAS

Indikator yang ingin dicapai setelah mempelajari modul ini adalah, siswa mampu:

1. menjelaskan peranan validitas suatu instrumen.
2. membedakan penggunaan validitas isi, criteria, dan konstruk.
3. membedakan dua macam validitas berdasarkan pengujiannya.
4. mengaplikasikan cara menguji validitas eksternal suatu instrumen.
5. mengaplikasikan cara menguji validitas internal suatu instrumen.

### A. Validitas Internal

Dalam penelitian data memiliki kedudukan sangat penting karena dari data itulah variabel penelitian dapat digambarkan. Kesalahan dalam pengambilan data, maka sudah dapat dipastikan akan terjadi pula kesalahan dalam pengambilan kesimpulan. Oleh karena itu, pengambilan data hendaknya dilakukan oleh orang yang benar-benar memahami permasalahannya dan mengikuti prosedur yang telah ditetapkan. Selain itu untuk menjamin kualitas data yang dikumpulkannya, seorang peneliti harus terlebih dahulu memperoleh keyakinan bahwa instrumennya (alat pengambil data) memiliki **validitas** (kesahihan) dan **reliabilitas** (keterandalan/keterpercayaan) yang memadai. Mengenai bagaimana caranya menguji validitas dan reliabilitas akan dijelaskan pada uraian berikutnya. Apabila peneliti tinggal menggunakan instrumen yang telah diakui validitas dan reliabilitasnya, maka peneliti tetap harus menginformasikan taraf validitas dan reliabilitasnya berdasarkan penelitian-penelitian sebelumnya atau berdasarkan konvensi-konvensi tertentu.

Validitas suatu instrumen menunjuk kepada apakah instrumen yang digunakan **mengukur apa yang seharusnya diukur**. Fraenkel (1990), membagi validitas menjadi tiga jenis, sebagaimana yang tertera pada tabel berikut ini:

Jenis Validitas	Metode	Pemakaian umum	Contoh
Isi	Secara logis memberikan kesimpulan, apakah isi tes tersebut memuat definisi-definisi yang memadai atau tidak terhadap tuntutan-tuntutan yang dikehendaki	Tes prestasi (achievement)	Tes mengenai kesiapan sekolah yang diuji untuk melihat apakah isinya berhubungan dengan pengetahuan dan kecakapan yang

	oleh pengukuran tersebut.		diharapkan pada anak-anak tingkat pertama yang mengikuti program Taman Kanak-kanak
Kriteria, yang berhubungan dengan validitas  Keterangan: Dahulu disebut validitas concurrent dan validitas prediktif	Memberikan tes dan membandingkannya dengan variabel kriteria. Variabel kriteria dapat diperoleh secara bersama-sama atau diramalkan dan kemudian akan dicapai di masa yang akan datang, tergantung pada rekomendasi pada penggunaan tes yang asli.	Tes yang digunakan pada tempat yang memiliki prosedur yang kompleks, atau untuk menyeleksi dan mengelompokkan seseorang	Diagnosis mengenai ketidakstabilan mental yang didasarkan pada card-sorting test dibandingkan dengan pengujian psychiatric yang lebih teliti, skor tes intelegensia yang digunakan untuk meramalkan performan masa yang akan datang
Konstrukt	Berdasarkan pada teori yang mendasari tes, menetapkan hipotesis dengan memperhatikan tingkah laku seseorang yang memiliki skor tinggi atau rendah. Uji hipotesis.	Tes yang digunakan untuk pendeskripsian atau di dalam penelitian ilmiah.	Jenis tes kepribadian (personality) yang dipelajari untuk memberikan kesimpulan secara teoritis mengenai faktor-faktor penyebab terhadap hasil-hasil yang diperoleh.

**1. Validitas isi**, merujuk pada sifat-sifat isi termasuk di dalamnya instrumen dan spesifikasi-spesifikasi yang digunakan peneliti untuk merumuskan isi. Bagaimana kelayakan isi? Bagaimana kekomprehensivannya? Apakah secara logis memiliki variabel yang diperlukan? Apakah sample dari item-item atau pertanyaan-pertanyaan yang mewakili isi yang dinilai cukup memadai?

Dengan demikian validitas ini dimaksudkan untuk menguji apakah bahan atau materi yang diujikan sesuai dengan pengetahuan, pelajaran, kemampuan, pengalaman, atau latar belakang orang yang diuji. Sebagai contoh: apabila isi tes tersebut ke luar dari apa yang diberikan, tidak mencakup keseluruhan bahan (kurang komprehensif), dan sebagainya sesuai dengan hal-hal yang telah dikemukakan di atas, maka tes tersebut tidak valid. Oleh karena itu untuk memperoleh validitas isi, diperlukan adanya uji coba instrumen dengan cara mengadakan sampling yang baik, melalui pemilihan item-item yang representatif dari keseluruhan bahan.

Validitas yang berkenaan dengan bahan atau materi mungkin relatif mudah dicapai, akan tetapi untuk validitas isi yang terkait dengan sikap atau sifat, seperti motivasi, rasa tanggung jawab, dan sebagainya yang terdapat dalam sejumlah kelakuan manusia tidaklah mudah. Hal

ini dikarenakan peneliti akan kesulitan menentukan seberapa besar dan sampai manakah sikap atau sifat itu benar-benar terkandung dalam bentuk kelakuan yang kita pilih.

Dalam hal pembelajaran juga seringkali terdapat perbedaan pendapat antara guru dan siswa mengenai validitas isi tersebut, khususnya yang terkait dengan kedalaman materi yang diujikan. Guru yang menyusun bahan tes, senantiasa berusaha agar instrumen yang dibuatnya memiliki validitas yang tinggi termasuk komprehensivitasnya, namun ada kemungkinan siswa merasa bahwa dalam instrumen tersebut banyak bahan yang belum pernah diajarkan, sehingga dianggap bahwa validitas isi dari instrumen tersebut rendah. Dengan demikian perlu adanya kesesuaian mengenai keseluruhan materi dan pilihan item-item yang representatif.

- 2. Validitas Kriteria,** pada awalnya validitas kriteria ini disebut **Validitas Concurrent** dan **validitas prediktif**. Validitas ini merujuk pada hubungan antara skor yang diperoleh dengan menggunakan instrumen dan skor yang diperoleh dengan satu atau lebih instrumen lain atau pengukuran-pengukuran (sering disebut kriteria). Seberapa tinggi hubungan tersebut? Seberapa baik skor-skor yang ada, memperkirakan keadaan sekarang atau meramalkan penampilan di masa yang akan datang tentang hal-hal tertentu? Jadi dengan validitas prediktif dimaksudkan adanya kesesuaian antara prediksi tentang kelakuan seseorang dengan kelakuan yang nyata.

Sebagai contoh, apabila skor suatu tes secara serentak dikumpulkan dengan informasi-informasi tentang kelakuan seseorang pada waktu yang sama, maka dengan membandingkan kedua jenis informasi tersebut kita dapat mengetahui validitas prediktifnya. Melalui perhitungan statistik korelasional, kita dapat pula menentukan tinggi rendahnya hubungan antara kedua jenis informasi tersebut.

- 3. Validitas konstrukt,** mengacu pada sifat konstruksi atau karakteristik yang diukur oleh suatu instrumen secara psikologi. Seberapa baik konstruksi tersebut menjelaskan perbedaan-perbedaan tingkah laku individu-individu atau penampilan mereka pada tugas-tugas tertentu. Sebagaimana kita ketahui bahwa banyak sifat-sifat yang tidak dapat secara langsung dilihat perwujudannya dalam dalam kelakuan manusia, misalnya kepribadian (personality) seseorang.

Untuk mengetahui aspek-aspek kepribadian mana sebenarnya yang ingin diukur oleh peneliti, dia dapat menggunakan tes kepribadian. Selain dengan tes kepribadian, teknik statistik pun yang dikenal dengan analisis faktor (anafak) dapat digunakan untuk menyelidiki berbagai komponen kepribadian seseorang, sehingga tes yang akan digunakan untuk

memperoleh validitas konstruk dapat disusun berdasarkan komponen-komponen tersebut. Validitas konstruk digunakan apabila kita meragukan apakah gejala yang dites hanya mengandung satu dimensi. Apabila ternyata gejala yang dites itu mengandung lebih dari satu dimensi, maka validitas tes tersebut dapat diragukan.

Apabila dilihat dari cara pengujiannya validitas dibedakan menjadi:

### **1. Validitas Internal**

Validitas internal umumnya merupakan **tujuan pertama** dalam metode eksperimental. Pertanyaan yang perlu dijawab adalah: *Apakah treatment/perlakuan eksperimental pada studi ini betul-betul dapat menimbulkan suatu perbedaan yang spesifik?* Kualitas validitas internal adalah yakin bahwa variable terikat benar-benar ditentukan oleh variabel bebasnya. Misalnya kita akan meneliti pengaruh pendekatan lingkungan terhadap hasil belajar mahasiswa. Maka kualitas variabel internalnya tinggi apabila hasil belajar tersebut yakin disebabkan oleh pendekatan lingkungan.

Terdapat delapan faktor pengganggu/berpengaruh terhadap validitas internal, yaitu:

- a. Sejarah (history), ada kemungkinan terdapat peristiwa-peristiwa khusus yang terjadi di antara pengukuran yang pertama dan kedua dalam melengkapi variable eksperimental.
- b. Kematangan (maturation), proses-proses di dalam suatu penelitian merupakan fungsi waktu, misalnya (pertambahan usia, rasa lapar, kelelahan, atau kurangnya minat dan perhatian, dll). Oleh karena itu jangan terlalu lama apabila penelitiannya hanya sebentar karena individu senantiasa berkembang.
- c. Testing, efek testing terhadap terhadap test berikutnya, misalnya pretest.
- d. Instrumen, kesalahan dalam pengukuran mungkin disebabkan oleh kesalahan dalam pengkaliberasian instrumen, atau kesalahan di dalam pengamatan atau penimbangan (judge).
- e. Regresi statistik, kemungkinan gejala yang terjadi pada kelompok yang telah diseleksi terdapat suatu skore yang ekstrim.
- f. Pemilihan sampel (selection), kesalahan pemilihan subjek yang akan dibandingkan dapat menghasilkan sesuatu yang bias.
- g. Kematian sampel (Experimental Mortality), berkurangnya subjek atau sampel
- h. Pemilihan-kematangan interaksi, misalnya efek interaksi di antara variabel-variabel tersebut dapat menyebabkan kesalahan atau gangguan terhadap variabel-variabel eksperimen.

Validitas internal merujuk pada adanya kesesuaian antara keseluruhan instrumen yang dibuat peneliti dengan bagian-bagian dari instrumen tersebut. Bagian dari instrumen tersebut dapat berupa **butir-butir** soal atau dapat pula berupa **faktor-faktornya** (biasanya merupakan kumpulan dari butir-butir soal. Karena pada pengujian ini hanya menganalisis keseluruhan instrumen dengan butir-butir dan faktor-faktor penyusun instrumen tersebut, dengan kata lain, tidak melibatkan sumber lain yang letak atau keberadaannya di luar instrumen yang digunakan, maka validitas yang diukur atau yang dihasilkan adalah **validitas internal**.

Dengan demikian, validitas internal suatu instrumen dikatakan tinggi apabila butir-butir soal penyusun instrumen dan juga faktor-faktornya yang merupakan bagian dari instrumen tersebut sejalan atau tidak menyimpang dari fungsi instrumen itu sendiri. Analisis butir dan analisis faktor biasa dilakukan dengan mengkorelasikan nilai-nilai butir soal atau nilai-nilai faktor dengan nilai total instrumen. Dengan diperolehnya nilai koefisien korelasi setiap butir, maka peneliti akan dengan mudah menentukan butir soal mana yang validitas internalnya tinggi atau kurang, demikian pula untuk faktor-faktornya.

## **B. Validitas External**

**Tujuan kedua** dari metode eksperimental adalah validitas eksternal yang menanyakan: *Seberapa representatifkah penemuan-penemuan penelitian dan seberapa besarkah hasilnya dapat digeneralisasikan terhadap subjek-subjek atau kondisi-kondisi yang sama?*

Dari contoh penelitian di atas, apabila perlakuan tersebut diterapkan pada kelas lain yang memiliki subjek dan kondisi yang sama dengan hasil yang sama maka validitas eksternalnya tinggi. Oleh karena itu seorang peneliti haruslah memiliki teknik sampling dan populasi yang baik. Kesalahan dalam menentukan populasi dan sampling akan menyebabkan kesalahan di dalam penarikan kesimpulan.

Terdapat empat faktor yang berpengaruh terhadap eksternal validity, yaitu

1. Pengaruh interaksi seleksi yang bias dan variabel eksperimen
2. Pengaruh interaksi pretest . Subjek yang diberi pretes akan memberikan respon yang berbeda dengan subjek yang tidak diberi pretes.
3. Pengaruh reaktif, dari prosedur eksperimental, pengaruh yang muncul dari setting eksperimental yang tidak akan terjadi pada setting noneksperimental

4. Pengaruh interferensi perlakuan yang berulang-ulang, menggunakan perlakuan yang berulang-ulang terhadap subjek yang sama akan berpengaruh terhadap perlakuan berikutnya karena pengaruh yang terdahulu tidak dapat dihilangkan.

Validitas eksternal merujuk pada data yang dihasilkan oleh suatu instrumen sesuai dengan informasi atau keterangan dari sumber lain yang terkait dengan variable penelitian yang dimaksud. Karena melibatkan sumber lain yang letak atau keberadaannya di luar instrumen yang digunakan maka validitas yang diukur atau yang dihasilkan adalah **validitas eksternal**.

Sebagai contoh, dalam suatu penelitian pendidikan seorang peneliti ingin mengetahui bagaimana validitas eksternal instrumen yang dibuatnya? Pada penelitian tersebut dia membuat instrumen berupa soal tes sebanyak 40 buah. Kemudian dia mengujicobakan soal tersebut kepada sejumlah siswa yang diperkirakan sesuai dengan subyek penelitian. Hasil ujicoba tersebut selanjutnya dikorelasikan dengan nilai-nilai siswa tersebut yang diambil dari nilai rapor. Nilai koefisien korelasi yang diperoleh menunjukkan derajat validitas eksternal instrumen tersebut. Rumus korelasi yang biasa digunakan adalah **korelasi product moment** dari **Pearson** sebagai berikut:

$$\text{Rumus 1 : } r_{XY} = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}}$$

Keterangan:  $x = X - \bar{X}$

$$y = Y - \bar{Y}$$

$\bar{X}$  = nilai rata-rata dari X

$\bar{Y}$  = nilai rata-rata dari Y

$$\text{Rumus 2 : } r_{XY} = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{\{(N\sum X^2) - (\sum X)^2\} \{(N\sum Y^2) - (\sum Y)^2\}}}$$

## KEGIATAN/TUGAS

Untuk memperdalam pemahaman mengenai materi di atas serta dapat mengaplikasikannya dalam kehidupan sehari-hari, silahkan Anda kerjakan kegiatan/tugas berikut ini! Tugas ini harus diserahkan satu minggu setelah Anda menerima dan mempelajari modul ini!

1. Cobalah Anda aplikasikan Rumus **korelasi product moment** dari **Pearson (rumus 1 atau rumus 2)**:

$$\text{Rumus 1 : } r_{XY} = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}}$$

Keterangan:  $x = X - \bar{X}$

$$y = Y - \bar{Y}$$

$\bar{X}$  = nilai rata-rata dari X

$\bar{Y}$  = nilai rata-rata dari Y

$$\text{Rumus 2 : } r_{XY} = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{\{(N\sum X^2) - (\sum X)^2\} \{(N\sum Y^2) - (\sum Y)^2\}}}$$

Gunakan data yang ada di sekolah atau cari data baru!

### *Petunjuk Jawaban Latihan*

Jika Anda menemui kesulitan dalam menjawab soal latihan tersebut di atas, gunakanlah petunjuk berikut ini !

Cobalah pelajari kembali uraian mengenai validitas eksternal, kemudian terapkan pemakaiannya sesuai dengan tugas yang diberikan!

### **Soal Evaluasi 1**

Pilihlah salah satu jawaban yang paling tepat!

1. Validitas suatu instrumen menunjuk kepada apakah instrumen yang digunakan mengukur apa yang seharusnya diukur. Pernyataan tersebut ...
  - a. salah, sebab validitas untuk mengukur keajegan suatu instrumen

- b. benar, sebab validitas untuk mengukur kesahihan suatu instrumen
  - c. salah, sebab untuk mengukur hal tersebut dengan reliabilitas
  - d. benar, sebab validitas hanya berperan dalam menguji soal yang terkait dengan isi
2. Tes mengenai kesiapan sekolah yang diuji untuk melihat apakah isinya berhubungan dengan pengetahuan dan kecakapan yang diharapkan pada anak-anak kelas 1 SD yang mengikuti program Taman Kanak-kanak, maka validitas yang dikaji adalah validitas ....
- a. isi            b. konstruk            c. prediksi            d. kriteria
3. Experimental Mortality dapat mengganggu validitas eksternal suatu instrmen penelitian. Pernyataan tersebut ...
- a. salah, sebab experimental mortality merupakan faktor pengganggu validitas internal
  - b. benar, sebab experimental mortality dapat mengurangi populasi yang sedang diteliti
  - c. salah, sebab experimental mortality hanya bersifat sementara
  - d. benar, sebab experimental mortality sangat erat kaitannya dengan sampel penelitian
4. Tes mengenai kesiapan sekolah untuk melihat apakah isinya berhubungan dengan pengetahuan dan kecakapan yang diharapkan pada anak-anak tingkat pertama yang mengikuti program Taman Kanak-kanak, sebaiknya menggunakan tes ...
- a. psikologi            b. personality            c. achievement            d. kepribadian
5. Apakah suatu treatment eksperimental pada studi ini betul-betul dapat menimbulkan suatu perbedaan yang spesifik? Pertanyaan tersebut merupakan ...
- a. tujuan kedua metode eksperimental            c. pengujian untuk reliabilitas
  - b. hipotesis yang perlu diuji kebenarannya            d. tujuan pertama metode eksperimental
6. Pilihan di bawah ini merupakan factor-faktor yang berpengaruh terhadap validitas eksternal, **kecuali** ...
- a. pengaruh interaksi seleksi yang bias dan variabel eksperimen
  - b. pengaruh interaksi pretest
  - c. pemilihan sampel
  - d. menggunakan perlakuan yang berulang-ulang terhadap subjek yang sama
7. Validitas internal suatu instrumen dikatakan tinggi apabila ...
- a. hipotesisnya dapat diuji secara empiris
  - b. butir-butir soal penyusun instrumen merupakan bagian dari instrumen tersebut
  - c. faktor-faktor penyusun soal tersebut dapat berdiri sendiri

- d. diujikan pada siapa saja hasilnya akan ama
8. Peristiwa-peristiwa khusus yang terjadi di antara pengukuran yang pertama dan kedua dalam melengkapi variable eksperimental. Hal tersebut termasuk faktor ...
- a. kematangan    b. pemilihan sampel    c. histori    d. instrumen
9. Seandainya pengujian instrument hanya menganalisis keseluruhan instrumen dan tidak melibatkan sumber lain yang letak atau keberadaannya di luar instrumen yang digunakan, maka yang diukur adalah ...
- a. konsistensi internal                      c. konsistensi eksternal  
b. validitas internal                          d. validitas eksternal
10. Kesalahan dalam pengukuran disebabkan oleh kesalahan dalam pengkaliberasian instrumen, atau kesalahan di dalam pengamatan atau penimbangan (judge) dapat menyebabkan gangguan terhadap ...
- a. validitas internal                          c. konsistensi internal  
b. konsistensi eksternal                      d. validitas eksternal

### **Ketuntasan belajar dan umpan balik**

Cocokkanlah jawaban Anda dengan Kunci Jawaban Evaluasi 1 yang terdapat di bagian akhir Modul ini. Hitunglah jawaban Anda yang benar, kemudian gunakan rumus di bawah ini untuk mengetahui tingkat penguasaan Anda terhadap materi Kegiatan Belajar 1.

#### **Rumus :**

**Setiap soal memiliki nilai 10**

$$\text{Tingkat penguasaan} = \frac{\text{jumlah jawaban Anda yang benar}}{10} \times 100\%$$

Arti singkat penguasaan yang Anda capai :

- 90 – 100 %    = baik sekali  
80 – 89 %     = baik  
70 – 79 %     = cukup  
< 70 %        = kurang

Bila Anda mencapai tingkat penguasaan 80 % atau lebih, Anda dapat meneruskan dengan kegiatan belajar selanjutnya. **Bagus !** Akan tetapi apabila tingkat penguasaan Anda masih di bawah 80 %. Anda harus mengulangi materi Bab I, terutama bagian yang belum Anda kuasai.

## BAB II RELIABILITAS

Indikator yang ingin dicapai setelah mempelajari modul ini adalah, siswa mampu:

1. menjelaskan peranan reliabilitas suatu instrumen.
2. mengaplikasikan dua cara menguji konsistensi eksternal reliabilitas suatu instrumen.
3. mengaplikasikan dua cara menguji konsistensi internal reliabilitas suatu instrumen.

### A. Pengertian Reliabilitas

Reliabilitas suatu instrumen menunjukkan **keajegan** (konsistensi) hasil pengukurannya seandainya instrumen tersebut digunakan oleh orang yang sama dalam waktu yang berlainan atau digunakan oleh orang yang berlainan dalam waktu yang sama. Reliabilitas secara implisit juga mengandung obyektivitas, karena hasil pengukurannya tidak terpengaruh oleh siapa pengukurnya. Fraenkel (1990), menyatakan bahwa reliabilitas merujuk terhadap konsistensi skor yang diperoleh. Bagaimana konsistensi siswa dari setiap individu dari satu pengadministrasian (administration) ke pengadministrasian lainnya dan dari satu set item ke set lainnya. Sebagai contoh, suatu tes dikatakan reliabel, apabila seorang siswa memperoleh nilai tinggi pada tes yang pertama akan memperoleh nilai yang tinggi pula pada tes berikut. Skor tersebut mungkin saja tidak persis identik, akan tetapi nilai tersebut harus tidak jauh berbeda.

Skor yang diperoleh dapat benar-benar reliabel, tetapi tidak valid. Hal ini diperkirakan, peneliti memberikan dua bentuk tes terhadap satu kelompok kelas yang dirancang untuk mengukur pengetahuan (knowledge) mereka. Mereka yang nilainya tinggi pada bentuk A juga tinggi pada form B, yang nilainya rendah pada form A juga rendah pada form B, dan seterusnya. Kita akan mengatakan bahwa skor-skor tersebut reliabel. Akan tetapi jika peneliti kemudian menggunakan tes nilai yang sama untuk meramalkan keberhasilan para siswa pada kelas yang berbeda bidang studinya, misalnya pendidikan fisika, dia mungkin akan keheranan. Beberapa kesimpulan mengenai keberhasilan dalam pendidikan fisika berdasarkan skor-skor tes tersebut tidak memiliki validitas. Sekarang bagaimana mengenai kebalikannya? Dapatkah suatu instrumen yang hasilnya tidak reliabel memperoleh kesimpulan-kesimpulan yang valid? Tidak! Jika nilai-nilai secara menyeluruh tidak konsisten untuk seseorang, nilai-nilai tersebut memberikan informasi yang tidak berguna. Kita tidak

mengetahui dengan pasti nilai mana yang dapat digunakan untuk mengambil kesimpulan mengenai kecakapan, sikap atau karakteristik-karakteristik lainnya yang dimiliki oleh siswa.

Dalam suatu penelitian, peneliti umumnya mengadakan pengukuran sebelum dan setelah perlakuan. Apabila hasil penelitiannya menunjukkan perbedaan, maka disimpulkan bahwa perbedaan tersebut adalah sebagai pengaruh variabel perlakuan (eksperimen). Oleh karena itu agar perbedaan tersebut benar-benar merupakan pengaruh variabel perlakuan, maka diperlukan instrumen yang reliabel sebagai syarat mutlak untuk menentukan pengaruh variabel yang satu terhadap variabel yang satu lagi. Selain itu reliabilitas instrumen juga merupakan syarat bagi validitas. Tes yang tidak reliabel dengan sendirinya tidak valid.

## **B. Cara-Cara Mengukur Reliabilitas**

Sebagaimana halnya validitas, reliabilitas suatu instrumen juga dapat diukur konsistensinya melalui metode-metode tertentu. Konsistensi reliabilitas tes yang diukur antara lain adalah konsistensi eksternal dan internalnya.

### **1. Pengukuran Konsistensi Eksternal**

Pengukuran konsistensi eksternal dapat dilakukan dengan metode-metode sebagai berikut:

#### **a) Metode Test-Retest atau Metode ulang**

Metode ini melibatkan dua kali penggunaan tes yang sama terhadap kelompok yang sama dengan interval waktu tertentu. Suatu hal yang perlu diperhatikan adalah interval waktu untuk memberikan tes kedua (re-tes), sebaiknya jangan terlalu dekat sebab dikhawatirkan siswa masih dapat mengingat soal yang diberikan pada tes pertama. Oleh karena itu peneliti hendaknya membuat soal yang tidak mudah diingat oleh siswa, mungkin tes obyektif mungkin lebih “sulit” diingat dibandingkan dengan tes essay yang jumlah soalnya relatif lebih sedikit dibandingkan dengan tes obyektif. Memang tidak ada patokan berapa lama interval waktu untuk melakukan tes dan re-tes, akan tetapi biasanya antara dua sampai empat minggu.

Setelah diperoleh nilai tes yang dipandang sebagai nilai X dan re-tes yang dipandang sebagai nilai Y, selanjutnya koefisien reliabilitas dihitung dengan menggunakan teknik **korelasi product moment** atau **korelasi Pearson** untuk menunjukkan korelasi (hubungan) antara dua set nilai yang diperoleh tersebut. Tinggi rendahnya indeks

korelasi inilah yang menunjukkan derajat reliabilitas instrumen yang digunakan.

#### **b) Metode Bentuk-Bentuk Equivalent (Equivalent-Forms Method) atau Paralel**

Apabila metode ini digunakan, maka dua buah tes (instrumen) yang terpisah satu sama lain, namun ekuivalen (juga disebut tes pilihan atau tes paralel) dikenakan pada kelompok individu yang sama pada periode waktu yang sama dalam arti tidak harus menunggu waktu beberapa hari atau minggu. Meskipun pertanyaan-pertanyaan antara kedua instrumen tersebut berbeda, akan tetapi keduanya harus memiliki isi (kontent) yang sama dan mengukur sesuatu yang sama pula atau identik.

Setelah diperoleh nilai tes untuk instrumen pertama yang dipandang sebagai nilai X dan instrumen kedua yang dipandang sebagai nilai Y, selanjutnya koefisien reliabilitas dihitung dengan menggunakan teknik **korelasi product moment** atau **korelasi Pearson**. Tingginya nilai koefisien korelasi yang diperoleh, menunjukkan bukti yang kuat mengenai reliabilitas bahwa kedua instrumen tersebut mengukur sesuatu yang sama.

Metode ini memungkinkan untuk menggabungkan antara metode tes-retes dan metode paralel (bentuk ekuivalen). Instrumen yang dibuat untuk tes paralel, diberikan dua kali terhadap kelompok individu yang sama, hanya dengan interval waktu yang berbeda sebagaimana dilakukan pada metode tes-retes. Koefisien reliabilitas yang tinggi akan menunjukkan bahwa dua bentuk instrumen tersebut tidak hanya mengukur dua jenis performance yang sama, akan tetapi juga memiliki konsistensi sepanjang waktu.

## **2. Pengukuran Konsistensi Internal**

Pada pengukuran konsistensi eksternal diperlukan dua kali pengadministrasian instrumen atau soal, sementara untuk mengukur konsistensi internal hanya diperlukan satu kali pengadministrasian instrumen. Jadi pada pengukuran konsistensi internal ini hanya diperlukan satu set soal. Beberapa cara pengukuran konsistensi internal ini adalah sebagai berikut:

#### **a) Prosedur Bagi-Dua (Split-Half Procedure)**

Prosedur bagi dua ini meliputi penilaian terhadap satu set soal yang dibagi dua (biasanya soal dengan nomor ganjil dan nomor genap atau awal dan akhir) dikerjakan

secara terpisah oleh setiap orang. Selanjutnya dihitung koefisien korelasi untuk kedua belahan soal tersebut. Koefisien korelasi yang diperoleh menunjukkan derajat korelasi terhadap kedua belahan soal, dan oleh karena itu menggambarkan konsistensi internal dari tes tersebut.

Dengan teknik belah dua ganjil-genap peneliti harus mengelompokkan nilai butir bernomor ganjil sebagai belahan pertama (X) dan kelompok nilai butir bernomor genap sebagai belahan kedua (Y). Selanjutnya carilah  $r_{XY}$  yaitu korelasi nilai skor belahan pertama dan kedua dengan menggunakan rumus Spearman – Brown sebagai berikut:

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

Keterangan:  $r_{XY}$  = Koefisien korelasi  
 N = Jumlah item soal  
 X = Jumlah nilai soal ganjil yang diperoleh setiap siswa  
 Y = Jumlah nilai soal genap yang diperoleh setiap siswa

Oleh karena itu koefisien korelasi yang diperoleh baru menunjukkan hubungan antara kedua belahan instrumen, dan untuk memperoleh koefisien atau indeks reliabilitas soal harus dihitung lagi dengan menggunakan rumus Spearman-Brown sebagai berikut:

$$\text{Reliabilitas dari nilai total tes} = \frac{2 \times \text{reliabilitas untuk } \frac{1}{2} \text{ tes}}{1 + \text{reliabilitas untuk } \frac{1}{2} \text{ tes}} \quad \text{atau} \quad \frac{2 \times r_{\frac{1}{2} \frac{1}{2}}}{1 + r_{\frac{1}{2} \frac{1}{2}}}$$

Keterangan:  $r_{\frac{1}{2} \frac{1}{2}} = r_{XY}$  yaitu kofisien korelasi antara kedua belahan soal

Sebagai contoh: Misalkan setelah dihitung dengan menggunakan rumus korelasi di atas, kita memperoleh nilai koefisien korelasi ( $r_{XY}$ ) kedua belahan soal sebesar 0,56. maka koefisien reliabilitasnya adalah:

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

Hal ini menggambarkan karakteristik penting mengenai reliabilitas. Reliabilitas suatu tes (atau suatu instrumen) dapat secara umum ditambah dengan menambahkan

jumlahnya apabila item-item yang ditambahkan serupa dengan item-item semula.

**b) Pendekatan-Pendekatan Kuder-Richardson (Kuder-Richardson Approaches)**

Metode yang paling sering digunakan untuk menentukan konsistensi internal adalah Pendekatan Kuder-Richardson, khususnya rumus KR20 dan KR21. Formula ini hanya memerlukan tiga buah informasi yaitu: jumlah item tes, rata-rata (Mean), dan standar deviasi (SD). Akan tetapi sebagai catatan bahwa KR21 dapat digunakan hanya jika diasumsikan bahwa item-item memiliki tingkat kesulitan yang sama (**are of equal difficulty**). Rumus KR21 yang paling sering digunakan adalah:

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

**Keterangan:** K = Jumlah soal, M = Rata-rata nilai tes, dan SD = Standar Deviasi nilai tes

Rumus tersebut sangat sederhana untuk digunakan, sebagai contoh, misalnya:

Jumlah soal (K) = 50, rata-rata nilai siswa (M) = 40, dan setelah dihitung, Standar Deviasi dari nilai yang dicapai siswa = 4, maka:

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

$$\text{Koefisien reliabilitas} = \frac{50}{49} \left\{ 1 - \frac{40(50-40)}{50(4^2)} \right\}$$

$$= 1,02 \left\{ 1 - \frac{40(10)}{50(16)} \right\}$$

$$= 1,02 \left\{ 1 - \frac{400}{800} \right\}$$

$$= (1,02)(1 - 0,5)$$

$$= (1,02)(0,5)$$

$$\mathbf{r_{K21} = 0,51}$$

Dengan demikian, reliabilitas dari nilai tes tersebut adalah 0,51. Apakah nilai tersebut baik atau jelek? Tinggi atau rendah? Ada dua hal yang dapat digunakan untuk menilai koefisien reliabilitas. Pertama, kita dapat membandingkan koefisien yang diperoleh dengan dua nilai ekstrim yang mungkin diperoleh, yaitu: Koefisien 0,00 menunjukkan tidak adanya korelasi, oleh karena itu reliabilitas soal tersebut tidak ada sama sekali, sementara 1,00 koefisien maksimal yang mungkin dicapai. Kedua, kita dapat membandingkan koefisien reliabilitas yang diperoleh dengan jenis-jenis koefisien yang biasa diperoleh untuk pengukuran jenis yang sama.

Rumus KR20 tidak memerlukan asumsi bahwa seluruh item memiliki tingkat kesulitan yang sama, akan tetapi rumus tersebut lebih sulit menghitungnya. Akan tetapi program komputer umumnya dapat digunakan dan bilamana peneliti tidak dapat mengasumsikan bahwa seluruh item tingkat kesulitannya sama.

## KEGIATAN/TUGAS

Untuk memperdalam pemahaman mengenai materi di atas serta dapat mengaplikasikannya dalam kehidupan sehari-hari, silahkan Anda kerjakan kegiatan/tugas berikut ini ! Tugas ini harus diserahkan satu minggu setelah Anda menerima dan mempelajari modul ini!

1. Coba Anda buat cara Pengukuran Konsistensi Internal dengan Prosedur Bagi-Dua (Split-Half Procedure). Gunakanlah rumus:

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

Keterangan:  $r_{XY}$  = Koefisien korelasi  
 N = Jumlah item soal  
 X = Jumlah nilai soal ganjil yang diperoleh setiap siswa  
 Y = Jumlah nilai soal genap yang diperoleh setiap siswa

Data untuk mengerjakan tugas tersebut dapat dilakukan di sekolah atau dimana saja yang Memungkinkan!

2. Jelaskan mengapa dalam Pengukuran Konsistensi Internal, Pendekatan-Pendekatan Kuder-Richardson lebih sering digunakan oleh para peneliti?

### **Petunjuk Jawaban Latihan**

Jika Anda menemui kesulitan dalam menjawab kegiatan/tugas tersebut di atas, gunakanlah petunjuk berikut ini !

1. Cobalah pelajari kembali uraian mengenai Prosedur Bagi-Dua (Split-Half Procedure)
2. Pelajari kembali uraian mengenai Pendekatan-Pendekatan Kuder-Richardso

### **Soal Evaluasi 2**

Pilihlah salah satu jawaban yang paling tepat!

1. Formula KR20 dan KR21 hanya dapat digunakan mengukur konsistensi internal jika diasumsikan bahwa item-item soal memiliki tingkat kesulitan yang berbeda. Pernyataan tersebut ...
  - a. salah, sebab formula KR20 dan KR21 digunakan untuk item-item soal yang memiliki tingkat kesukaran sama
  - b. benar, sebab formula KR20 dan KR21 digunakan membandingkan koefisien yang diperoleh dengan dua nilai ekstrim yang mungkin diperoleh
  - c. salah, sebab konsistensi internal hanya dapat diukur dengan menggunakan prosedur split half
  - d. benar, sebab formula KR20 dan KR21 memiliki tingkat validitas yang cukup tinggi
2. Pendekatan Kuder-Richardson digunakan untuk menentukan ...
  - a. konsistensi eksternal
  - b. konsistensi internal
  - c. validitas eksternal
  - d. validitas internal
3. Pada pengukuran konsistensi internal ini diperlukan ...
  - a. 1 set soal
  - b. 2 set soal ganjil genap
  - c. 3 set soal yang bobot kesulitannya sama
  - d. 4 set soal yang setara
4. Manakah dari pilihan di bawah ini yang dapat digunakan untuk menguji konsistensi eksternal suatu instrumen ...
  - a. prosedur bagi dua
  - b. metode Test-Retest atau Metode ulang
  - c. pendekatan-Pendekatan Kuder-Richardson
  - d. korelasi product moment

5. Metode bentuk equivalent atau parallel digunakan untuk mengukur ...
- reliabilitas secara umum
  - konsistensi internal suatu instrumen
  - koefisien korelasi
  - konsistensi eksternal suatu instrumen
6. Apabila seorang peneliti memberikan tes pada kelompok individu yang sama pada periode waktu yang sama (tidak harus menunggu waktu beberapa hari atau minggu). Maka peneliti tersebut sedang melakukan pengujian instrument dengan cara ...
- bentuk Equivalent
  - test-retest
  - pendekatan Kuder-Richardson
  - split half
7. Pilihan di bawah ini merupakan karakteristik reliabilitas, **kecuali** ...
- menunjukkan keajegan suatu instrumen
  - hasil pengukurannya tidak dipengaruhi oleh siapa pengukurnya
  - menunjukkan kesahihan suatu alat ukur
  - hasil pengukurannya akan sama seandainya instrumen tersebut digunakan oleh orang yang sama dalam waktu yang berlainan atau digunakan oleh orang yang berlainan dalam waktu yang sama
8. Pendekatan Kuder-Richardson merupakan metode yang paling sering digunakan untuk menentukan konsistensi internal. Pilihan di bawah ini merupakan bagian-bagian yang harus diketahui untuk dapat menggunakan Pendekatan Kuder-Richardson, **kecuali** ...
- rata-rata (Mean)
  - standar deviasi (SD)
  - chi kuadrat
  - jumlah item tes
9. Reliabilitas instrumen merupakan syarat bagi validitas. Pernyataan tersebut ...
- salah, sebab soal yang tidak reliabel belum tentu tidak valid
  - benar, sebab instrument yang tidak reliabel dengan sendirinya tidak valid
  - salah, sebab instrument yang reliabel pasti valid
  - benar, sebab instrument yang reliabel dapat digunakan untuk semua disiplin ilmu
10. Apabila seorang siswa memperoleh nilai tinggi pada tes yang pertama dan memperoleh nilai yang tinggi pula pada tes berikutnya, maka soal tersebut dikatakan ...
- validitasnya tinggi
  - konsistensi internal
  - konsistensi internal
  - reliabel

## **Ketuntasan belajar dan umpan balik**

Cocokkanlah jawaban Anda dengan Kunci Jawaban evaluasi 2 yang terdapat di bagian akhir Modul ini. Hitunglah jawaban Anda yang benar, kemudian gunakan rumus di bawah ini untuk mengetahui tingkat penguasaan Anda terhadap materi Kegiatan Belajar 2.

### **Rumus :**

#### **Setiap soal memiliki nilai 10**

$$\text{Tingkat penguasaan} = \frac{\text{jumlah jawaban Anda yang benar}}{10} \times 100\%$$

Arti singkat penguasaan yang Anda capai :

90 – 100 % = baik sekali

80 – 89 % = baik

70 – 79 % = cukup

< 70 % = kurang

Bila Anda mencapai tingkat penguasaan 80 % atau lebih, Anda dapat meneruskan dengan kegiatan belajar selanjutnya. **Bagus !** Akan tetapi apabila tingkat penguasaan Anda masih di bawah 80 %. Anda harus mengulangi materi Bab II, terutama bagian yang belum Anda kuasai.

## **KUNCI JAWABAN**

### **Kunci Soal evaluasi 1**

- |      |       |
|------|-------|
| 1. b | 6. d  |
| 2. a | 7. b  |
| 3. a | 8. c  |
| 4. c | 9. b  |
| 5. c | 10. a |

### **Kunci Soal evaluasi 2**

- |      |       |
|------|-------|
| 1. a | 6. a  |
| 2. b | 7. c  |
| 3. a | 8. c  |
| 4. b | 9. b  |
| 5. d | 10. d |

## DAFTAR ISTILAH PENTING

**Korelasi product moment** atau **korelasi Pearson** untuk menunjukkan derajat reliabilitas korelasi antara dua set nilai yang diperoleh. Tinggi rendahnya indeks korelasi tersebut yang menunjukkan derajat reliabilitas instrumen yang digunakan.

**Reliabilitas suatu instrument**, menunjukan **keajegan** (konsistensi) hasil pengukurannya seandainya instrumen tersebut digunakan oleh orang yang sama dalam waktu yang berlainan atau digunakan oleh orang yang berlainan dalam waktu yang sama.

**Validitas eksternal**, merupakan tujuan kedua dari metode eksperimental yang menanyakan: Seberapa representatifkah penemuan-penemuan penelitian dan seberapa besarkah hasil-hasilnya dapat digeneralisasikan terhadap subjek-subjek atau kondisi-kondisi yang sama? Apabila perlakuan tersebut diterapkan pada kelas lain yang memiliki subjek dan kondisi yang sama hasilnya akan yang sama, validitas eksternalnya maka tinggi.

**Validitas internal**, merupakan tujuan pertama dalam metode eksperimental yang meyakini bahwa variable terikat benar-benar ditentukan oleh variabel bebasnya. Pertanyaan yang perlu dijawab adalah: Apakah treatment/perlakuan eksperimental pada studi ini betul-betul dapat menimbulkan suatu perbedaan yang spesifik? Validitas internal merujuk pada adanya kesesuaian antara keseluruhan instrumen yang dibuat peneliti dengan bagian-bagian dari instrumen tersebut.

**Validitas isi**, merujuk pada sifat-sifat isi termasuk di dalamnya instrumen dan spesifikasi-spesifikasi yang digunakan peneliti untuk merumuskan isi. Dengan demikian validitas ini dimaksudkan untuk menguji apakah bahan atau materi yang diujikan sesuai dengan pengetahuan, pelajaran, kemampuan, pengalaman, atau latar belakang orang yang diuji.

**Validitas konstruk**, mengacu pada sifat konstruksi atau karakteristik yang diukur oleh suatu instrumen secara psikologi.

**Validitas kriteria**, pada awalnya validitas kriteria ini disebut **Validitas Concurrent** dan **Validitas prediktif**. Validitas ini merujuk pada hubungan antara skor yang diperoleh dengan menggunakan instrumen dan skor yang diperoleh dengan satu atau lebih instrumen lain atau pengukuran-pengukuran (sering disebut kriteria). validitas prediktif dimaksudkan adanya kesesuaian antara prediksi tentang kelakuan seseorang dengan kelakuan yang nyata.

**Validitas suatu instrument**, menunjuk kepada apakah instrumen yang digunakan mengukur apa yang seharusnya diukur.

#### **DAFTAR PUSTAKA**

Borg, Walter R. and Mredith D. Gall. (1989). *Educational Research*. Fifth Edition. New York: Pitman Publishing Inc.

Creswell, John W. (2008). *Educational Research*. Third Edition. New Jersey: Pearson Education Australia Pty. Limited.

Fraenkel, Jack R., Norman E.Wallen.(1990).*How to Design and Evaluate Research in Education*.  
Second Edition. San Fransisco: Mc Graw - Hill Publishing Company.

Stephen, Isaac and William B. Michael (1982). *Handbook in Research and Evaluation*.  
California: Edits Publisher.