

VALIDITAS (*KESAHIHAN*)

1. Pengertian Validitas

Pertanyaan-pertanyaan yang paling utama yang harus diajukan terhadap suatu prosedur pengukuran adalah: *sampai di manakah validitasnya?* Dalam hal ini harus dilihat apakah ujian yang dipakai betul-betul mengukur semua yang seharusnya diukur dan tidak lain dari pada itu. **Suatu alat pengukur dikatakan valid jika ia benar-benar cocok untuk mengukur apa yang hendak diukur.** Sebagaimana dikemukakan oleh Scarvia B. Anderson dalam bukunya "Encyclopedia of Educational Evaluation" disebutkan bahwa "*A test is valid it measures what it purpose to measure*" (sebuah tes dikatakan valid apabila tes tersebut mengukur apa yang hendak diukur).

Dalam bahasa Indonesia "*valid*" disebut dengan istilah "*sahih*". Misalnya: Untuk mengukur panjang dipakai meteran, mengukur berat dipakai timbangan, mengukur penguasaan matematika dipakai ujian matematika untuk kelas yang setara, dan sebagainya. Secara lebih jelas, suatu ujian untuk mata kuliah tertentu dikatakan valid jika ia benar-benar cocok dengan tujuan-tujuan yang telah ditetapkan untuk dicapai dengan penyajian mata kuliah tersebut.

Sebenarnya pembicaraan validitas ini *bukan ditekankan pada tes* itu sendiri tetapi pada *hasil pengetesan atau skornya*. Suatu tes hasil belajar dapat dikatakan tes yang valid apabila tes tersebut betul-betul dapat mengukur hasil belajar. Jadi bukan sekedar mengukur daya ingat atau kemampuan bahasa saja misalnya.

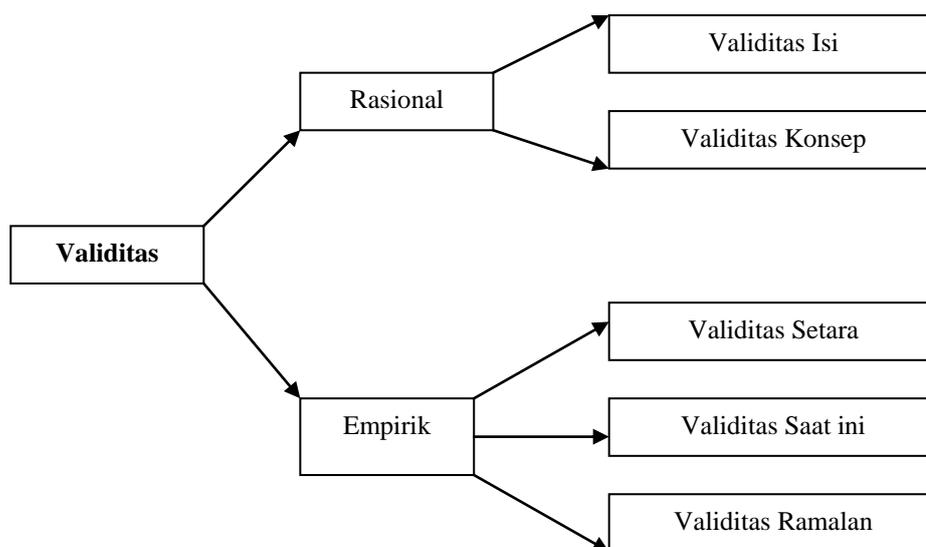
2. Jenis-jenis Validitas

Ada dua kenyataan pokok yang memperlihatkan taraf validitas suatu ujian, yaitu yang dipertimbangkan secara *rasional* dan yang dilihat melalui *prosedur empirik*. Analisis secara rasional dapat dilakukan terhadap topik dan bidang yang diujikan, yaitu isi ujian tersebut. Validitas yang diperoleh melalui analisis seperti ini disebut *validitas isi (contens validity)*. Analisis rasional dapat juga dilakukan terhadap kegiatan-kegiatan dan proses-proses sesuai dengan konsep tertentu yang seharusnya menjadi isi dari ujian itu, dan hasilnya disebut *validitas konsep atau konstruksi (concept / contruct validity)*.

Jenis kedua ialah kenyataan validitas yang bersifat empirik dan statistik. Jenis ini diperoleh dengan memperhatikan hubungan yang ada antara alat (ujian) yang

sedang dipelajari dengan pengukuran atau kenyataan-kenyataan yang lain. Bahan-bahan perbandingan ini mungkin diperoleh bersamaan waktunya dengan waktu penyelenggaraan ujian yang dimaksud mungkin juga tidak. Untuk validitas jenis empirik ini akan dibicarakan tiga kenyataan validitas yaitu : *validitas pengukuran setara (congruent validity)*, *validitas pengukuran serentak (concurrent validity)* dan *validitas ramalan (predictive validity)*.

Secara ilustrasi dapat dilihat sebagai berikut:



a. Validitas Isi (contens validity)

Validitas isi artinya ketepatan daripada suatu tes dilihat dari segi isi tersebut. Suatu tes hasil belajar dikatakan valid, apabila materi tes tersebut betul-betul merupakan bahan-bahan yang *representatif* terhadap bahan-bahan pelajaran yang diberikan. Dengan kata lain sebuah tes dikatakan memiliki validitas isi apabila *mengukur tujuan khusus tertentu yang sejajar dengan materi atau isi pelajaran yang diberikan*. Oleh karena materi yang diajarkan tertera dalam kurikulum maka validitas isi ini sering disebut ***validitas kurikuler***. Validitas isi dapat diusahakan tercapainya sejak saat penyusunan dengan cara memerinci materi kurikulum atau materi buku pelajaran.

Misalnya untuk siswa kelas I SMU akan diberikan tes Matematika, maka item-itemnya harus diambil dari materi pelajaran kelas I, apabila kita sisipkan item-item yang diambil dari materi pelajaran kelas III maka tes tersebut sudah tidak valid lagi. Contoh lain, untuk mahasiswa jurusan bahasa Inggris tingkat

pertama disusun ujian tentang "ketepatan ekspresi tertulis". Berkenaan dengan validitas isi untuk ujian yang disusun ini harus ditelaah isi ujian itu sendiri. Jika misalnya ujian itu belum tersusun secara lengkap penelaahan dapat dilakukan terhadap perencanaan ujian tersebut. Tujuan utama penelaahan ini ialah untuk melihat sampai di mana isi ujian mencerminkan hal-hal yang telah diajarkan. Untuk penelaahan seperti ini seringkali diperlukan peninjauan kembali terhadap bahan-bahan yang telah diajarkan, misalnya (untuk pengajaran bahasa Inggris di tahun pertama) peraturan pemakaian huruf besar dan tanda baca, pemilihan kata-kata yang tepat, bentuk-bentuk kata kerja (tenses) dan sebagainya. Jika ujian tersebut mempunyai validitas isi yang tinggi maka ia harus mengandung bahan-bahan seperti yang telah diajarkan itu.

Untuk menilai apakah suatu tes memiliki validitas isi atau tidak dapat kita lakukan dengan jalan membandingkan materi tes tersebut dengan analisa rasional yang kita lakukan terhadap bahan-bahan yang seharusnya dipergunakan dalam menyusun tes tersebut. Apabila materi tes tersebut telah cocok dengan analisa rasional yang kita lakukan, berarti tes yang kita nilai itu mempunyai validitas isi. Sebaliknya apabila materi tes tersebut menyimpang dari analisa rasional kita, berarti tes tersebut tidak valid ditinjau dari validitas isinya.

b. Validitas Konsep atau Konstruksi (*concept/construct validity*)

Sebuah tes dikatakan memiliki validitas konstruksi apabila butir-butir soal yang membangun tes tersebut mengukur setiap aspek berpikir seperti yang disebutkan dalam Tujuan pembelajaran Khusus. Dengan kata lain jika butir-butir soal mengukur aspek berpikir tersebut sudah sesuai dengan aspek berpikir yang menjadi tujuan pembelajaran.

Untuk mengetahui apakah suatu tes memenuhi syarat-syarat validitas konstruksi atau tidak maka kita harus membandingkan susunan tes tersebut dengan syarat-syarat penyusunan tes yang baik. Apabila susunan tes tersebut telah memenuhi syarat-syarat penyusunan tes maka berarti tes tersebut memenuhi syarat validitas konstruksi, apabila tidak memenuhi syarat-syarat penyusunan tes berarti tidak memenuhi validitas konstruksi.

"Konstruksi" dalam pengertian ini bukanlah **"susunan"** seperti yang sering dijumpai dalam teknik, tetapi merupakan rekaan psikologis yaitu suatu rekaan yang dibuat oleh para ahli Ilmu Jiwa yang dengan suatu cara tertentu

"memerinci" isi jiwa atas beberapa aspek seperti : ingatan (pengetahuan), pemahaman, aplikasi dan seterusnya. Dalam hal ini mereka menganggap seolah-olah jiwa dapat dibagi-bagi. Tetapi sebenarnya tidak demikian. Pembagian ini hanya merupakan tindakan sementara untuk mempermudah mempelajari.

Misalnya kalau kita akan memberikan tes mata pelajaran IPA, kita harus membuat soal yang ringkas dan jelas yang benar-benar mengukur kecakapan IPA, bukan mengukur kemampuan bahasa karena soal itu ditulis secara berkepanjangan dengan bahasa yang sukar dimengerti.

Pada prakteknya, penelaahan terhadap validitas isi dan validitas konstruksi seringkali bersangkut paut.

c. Validitas pengukuran Setara (*congruent validity*)

Jenis kevalidan ini menunjukkan kenyataan yang diperoleh dengan mengkorelasikan hasil suatu ujian dengan pengukuran yang setara (mengukur fungsi yang sama). Demikianlah, mengkorelasikan hasil sebuah tes intelegensi yang baru dengan hasil tes intelegensi yang sudah ada akan memberikan kenyataan validitas jenis ini.

d. Validitas Pengukuran Serentak (*concurrent validity*)

Validitas ini lebih umum dikenal dengan validitas empiris. Sebuah tes dikatakan memiliki validitas empiris jika hasilnya sesuai dengan pengalaman. Jika istilah "*sesuai*" tentu ada dua hal yang dipasangkan. Dalam hal ini hasil tes dipasangkan dengan hasil pengalaman. Pengalaman selalu mengenai hal yang telah lampau sehingga data pengalaman tersebut sekarang sudah ada (ada sekarang, *concurrent*).

Dalam membandingkan hasil sebuah tes maka diperlukan suatu kriterium atau alat banding. Maka hasil tes merupakan sesuatu yang dibandingkan.

Jenis validitas pengukuran serentak ini menunjukkan kenyataan yang diperhitungkan dengan mengkorelasikan hasil ujian yang dimaksud dengan suatu ukuran lain yang pengukurannya dilaksanakan bersamaan waktu dengan pelaksanaan ujian itu. Jika suatu tes yang memeriksa sosialibilitas misalnya, dikorelasikan dengan kedudukan sosialibilitas mahasiswa yang di tes itu menurut

teman-temannya yang terdekat, maka hasilnya akan memperlihatkan kenyataan validitas jenis ini.

Misalnya seorang guru ingin mengetahui apakah tes sumatif yang disusun sudah valid atau belum. Untuk ini diperlukan sebuah kriterium mas lalu yang sekarang datanya dia memiliki misalnya nilai ulangan harian atau nilai ulangan sumatif yang lalu.

e. Validitas Ramalan (*predictive validity*)

Memprediksi artinya meramal, dan meramal selalu mengenai hal yang akan datang jadi sekarang belum terjadi. Sebuah tes dikatakan memiliki validitas prediksi atau validitas ramalan apabila mempunyai kemampuan untuk meramalkan apa yang akan terjadi pada masa yang akan datang.

Jenis validitas ini menunjukkan kenyataan jika ujian yang dimaksud dihubungkan dengan kriteria-kriteria tentang hasil karya atau kesuksesan di masa depan. Demikianlah jika suatu tes bakat skolastik diberikan pada siswa-siswa SMU dikorelasikan dengan prestasi mereka di perguruan tinggi, maka kenyataan yang diperoleh itu akan menunjukkan validitas ramalan.

Misalnya tes masuk Perguruan Tinggi adalah sebuah tes yang diperkirakan mampu meramalkan keberhasilan peserta tes dalam mengikuti kuliah di masa yang akan datang. Calon yang tersaring berdasarkan hasil tes diharapkan mencerminkan tinggi-rendahnya kemampuan mengetahui kuliah. Jika nilai tesnya tinggi tentu menajmin keberhasilannya kelak. Sebaliknya seorang calon dikatakan tidak lulus tes karena memiliki nilai tes yang rendah diperkirakan akan tidak mampu mengikuti perkuliahan yang akan datang.

Sebagai alat pembanding validitas prediksi adalah nilai-nilai yang diperoleh setelah peserta tes mengikuti pelajaran di Perguruan Tinggi. Jika ternyata siapa yang memiliki nilai tes lebih tinggi gagal dalam ujian semester I dibandingkan dengan yang dahulu nilai tesnya lebih rendah maka tes masuk yang dimaksud tidak memiliki validitas prediksi.

MENGHITUNG VALIDITAS

Validitas tes dapat dilakukan dengan berbagai cara. Satu diantaranya adalah dengan menggunakan teknik korelasi antara tes yang sedang dicari validitasnya dengan tes lain yang sudah memiliki validitas cukup baik.

Teknik Korelasi

Banyak teknik korelasi yang dapat digunakan untuk menghitung validitas . pada bagian ini hanya dibicarakan teknik korelasi product moment yang dikembangkan Karl Pearson, dan yang lain teknik korelasi tat Jenjang atau Rank Order Correlation yang dikembangkan oleh Spearman. Perbedaan antara keduanya terletak pada data yang diolah. Teknik korelasi Product Moment mendasarkan perhitungannya pada angka-angka kasar seperti apa adanya (data apa adanya), sedangkan Tata Jenjang justru mengabaikan angka kasar ini dan mendasarkan perhitungannya pada jenjang atau urutan ranking.

Umpamanya kita akan menguji validitas prediksi tes masuk Akper “Aisyiah dalam bidang studi bahasa Inggris yang nilainya dijadikan criteria dalam seleksi penerimaan siswa baru. Nilai seleksi tes masuk dan hasil belaajr 5 siswa yang diperoleh waktu belajar di Akper, dalam bidang studi yang sama, seperti tertera pada table berikut ini beserta langkah perhitungan untuk rumus spearman (rs).

No	Nama	Nilai Tes Seleksi	Nilai Hasil belajar saat ini	Rank		d	d ²
				X	Y		
1.	Aisyah	6	6	4,5	5	-0,5	0,25
2.	Dahlan	7	7	2,5	3	0,5	0,25
3.	Yusuf	6	7	4,5	3	1,5	2,25
4.	Rahmah	8	8	1	1	0	0
5.	Zahra	7	7	2,5	3	-0,5	0,25
Jumlah							2

Berdasarkan harga pada table 9, maka harga rs dapat dihitung yaitu:

$$\begin{aligned}rs &= 1 - \frac{6 \cdot \sum d^2}{N^3 - N} \\&= 1 - \frac{6 \cdot 2}{5^3 - 5} \\&= 0,1\end{aligned}$$

harga rs yang diperoleh, untuk menguji signifikansinya harus dihitung melalui uji t dari Kendal dengan rumus :

$$\begin{aligned}t &= \sqrt{\frac{N-2}{1-rs^2}} \\t &= \sqrt{\frac{5-2}{1-0,1}} = 1,7407\end{aligned}$$

untuk menafsirkan signifikansi harga t digunakan table F.

Harga t table dengan df N-2 yaitu $5-2 = 3$ pada tingkat kepercayaan 0,95 sebesar 5,841. hal ini membuktikan bahwa kedudukan t hitung dalam t table = 1,7407. 5,842 yang berarti validitas tes masuk studi bahasa Inggris terhadap hasil belajar bahasa Inggris di akper tidak signifikan (dipercaya).

Kemudian untuk menafsirkan besar kecilnya hubungan digunakan criteria berikut:

Kurang dari 0,20 hubungan dapat dianggap tidak ada

Antara 0,20 – 0,40 hubungan ada tapi rendah

Antara 0,40 – 0,70 hubungan cukup

Antara 0,70 – 0,90 hubungan tinggi

Antara 0,90 – 1,00 hubungan sangat tinggi

Mengacu kepada criteria makna hubungan, maka dapat disimpulkan bahwa validitas tes masuk Akper bidang studi bahasa Inggris terhadap hasil belajar, termasuk hubungan yang dianggap tidak ada, oleh karena harga rs 0,1 terletak di antara kurang dari 0,20 dan hubungan tersebut ternyata tidak signifikan pada tingkat kepercayaan 0,95.

Rumus Pearson (r) yang digunakan untuk pengujian validitas ialah :

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

Contoh :

Seorang Dosen ingin menguji validitas tes yang ia konstruksi dalam bidang studi Kesehatan masyarakat dengan jumlah butir soal 10 dalam tipe jawaban singkat. Ia bermaksud menguji validitasnya dengan cara : skor tiap siswa disajikan dalam table perhitungan berikut: