

## Validitas Tes

Nurul Mukhlisa

Universitas Negeri Makassar, Indonesia

Email: [nurullmukhlisaa@unm.ac.id](mailto:nurullmukhlisaa@unm.ac.id)

**Abstrak:** penelitian ini bertujuan untuk (1) mengetahui pengertian validitas dan (2) mengetahui jeni-jenis validitas. Penelitian ini termasuk studi kepustakaan dengan pendekatan kualitatif. Pengumpulan data dilakukan melalui pengumpulan kajian teori dari berbagai buku. Kemudian setiap hasil kajian teori dalam buku tersebut direview dan dianalisis secara seksama. Berdasarkan hasil penelitian, maka disimpulkan bahwa suatu tes dikatakan valid jika mengukur apa yang hendak diukur. Dengan kata lain suatu tes dikatakan valid jika sesuai dengan tujuan tes tersebut diadakan. Validitas terbagi tiga, yaitu validitas isi, validitas kriteria, dan validitas konstruk. Validasi dilakukan sebelum dan setelah tes diujicobakan. Sebelum diujicobakan, kita dapat mengetahui validitas isi tes berdasarkan hasil validasi oleh pakar. Validitas kriteria dan validitas konstruk dapat diketahui setelah tes diujicobakan.

**Kata kunci:** tes, validitas isi, validitas kriteria, validitas konstruk

### PENDAHULUAN

Capaian pembelajaran terlihat dari hasil belajar peserta didik. Hasil belajar peserta didik diperoleh dari hasil tes peserta didik. Namun, muncul pertanyaan, apakah benar hasil tes tersebut benar-benar menggambarkan capaian pembelajaran siswa? Jawabannya adalah benar, namun jika tes tersebut telah teruji validitasnya. Dengan kata lain, tes tersebut valid untuk mengukur capaian pembelajaran peserta didik.

Hasil belajar peserta didik tidak akan menggambarkan capaian pembelajaran peserta didik jika alat ukur yang digunakan tidak valid. Jika tes yang tidak valid digunakan sebagai alat ukur, maka tinggi atau rendahnya hasil belajar peserta didik tidak dapat dijadikan acuan untuk menggambarkan kompetensi yang dimiliki oleh peserta didik.

Perangkat tes yang baik adalah alat tes yang hasil pengukurannya dapat memberikan gambaran yang sesungguhnya mengenai kemampuan peserta tes dalam bidang tertentu yang menjadi sasaran ukurnya. Perangkat ukur yang baik tidak akan memberikan data yang menyesatkan pada sesuatu yang diukur.

Oleh karena itu, dalam penyusunan perangkat ukur sebelum digunakan perlu dilakukan pengujian validitas pengukuran dan reliabilitas terlebih dahulu.

Suatu tes dikatakan valid apabila tes tersebut mengukur apa yang hendak diukur (Allen, 1979). Validitas didefinisikan sebagai ukuran seberapa cermat suatu alat ukur melakukan fungsi ukurnya. Misalnya, tes hanya dapat melakukan fungsinya dengan cermat, jika ada sesuatu yang diukur. Jadi untuk dikatakan valid, tes harus mengukur sesuatu dan dilakukan dengan cermat (Mardapi, 2008). Jadi suatu tes dinyatakan valid apabila tepat mengukur sesuatu yang hendak diukur sesuai dengan tujuan tes.

Pengertian validitas menurut Aiken Lewis R (1997), "*validity of a test has been defined as the extent to which the test measures what it was designed to measure.*" Selanjutnya, Hadi (1991) "kesahihan dibatasi sebagai tingkat kemampuan suatu instrumen untuk mengungkap sesuatu yang menjadi sasaran pokok pengukuran." Berdasarkan batasan tersebut, validitas dapat diartikan sejauh mana hasil pengukuran dapat diinterpretasikan sebagai cerminan sasaran ukur yang berupa

kemampuan, karakteristik, atau tingkah laku yang diukur melalui alat ukur yang tepat (Susetyo, 2015). Jadi, validitas menunjukkan ketepatan suatu instrumen dalam mengukur.

Validitas adalah dukungan bukti dan teori bahwa penafsiran skor tes sesuai dengan tujuan penggunaan tes. Oleh karena itu, validitas adalah fundamen paling dasar dalam mengembangkan suatu tes. Proses validasi mencakup pengumpulan bukti-bukti untuk menunjukkan dasar ilmiah penafsiran skor seperti yang telah direncanakan (Mardapi, 2008).

Apabila kita ingin mengukur berat sebuah kalung emas, maka kita harus menggunakan timbangan emas agar hasil ukur tersebut valid. Sebuah timbangan beras memang mengukur "berat", tetapi tidak cukup cermat untuk mengukur berat emas. Oleh karena itu, sebuah timbangan beras tidak valid untuk mengukur berat emas. Demikian pula, apabila kita ingin menghitung waktu tempuh yang kita perlukan dari suatu kota ke kota lainnya dengan mengendarai mobil, sebuah jam tangan adalah valid untuk digunakan. Akan tetapi, jam tangan yang sama tidak cukup valid untuk mengukur waktu yang diperlukan seorang atlet pelari cepat dalam menempuh jarak 100 m karena diperlukan unit waktu terkecil sampai pada pecahan-detik (Mardapi, 2008).

Istilah validitas banyak dipakai dalam penelitian, yaitu validitas eksperimen, validitas pengukuran, dan validitas butir. Oleh karena itu, kita perlu hati-hati mempergunakan istilah validitas yang memungkinkan terjadinya perbedaan makna pada istilah tersebut. Sebelum digunakan sebagai alat pengumpul data, instrumen tes diujicobakan terlebih dahulu sampai memenuhi persyaratan yakni salah satunya, yaitu valid. Suatu tes dinyatakan valid jika perangkat tes butir-butirnya benar-benar mengukur tujuan tes berupa kemampuan dalam hal tertentu, bukan kemampuan yang lainnya. Validitas suatu tes menggunakan indeks angka yang menunjukkan sejauh mana hasil pengukuran dapat mencerminkan secara cermat kemampuan seseorang yang diperoleh dari berbagai proses. Validitas pengukuran memiliki nilai dari yang tinggi sampai rendah, jika semakin tinggi nilai validitas, maka semakin baik pula

tingkat validitasnya. Demikianlah istilah validitas yang dikaitkan dengan pengukuran, sedangkan validitas butir hanya berbicara tentang hubungan di antara skor satuan pada butir dengan skor responden (Susetyo, 2015).

Pemeriksaan validitas pengukuran dilakukan sebelum alat ukur digunakan untuk sesungguhnya. Pemeriksaan validitas pengukuran dapat dilakukan setelah alat ukur baru dibuat atau disusun. Pemeriksaan validitas pengukuran juga dapat dilakukan pada saat uji coba alat ukur. Apabila hasil pemeriksaan menunjukkan tingkat validitas rendah, maka alat ukur dapat diperbaiki. Pemeriksaan validitas dan perbaikan alat ukur dilakukan berulang kali sampai alat ukur mencapai validitas pengukuran yang cukup tinggi (Arafah, 2017).

Sebenarnya validitas pengukuran merupakan satu kesatuan kecocokan antara pengukuran dan sasaran ukur. Berdasarkan sifat pencocokan, validitas dapat dibagi ke dalam beberapa jenis. Pada sejarahnya, nomenklatur jenis validitas pengukuran mengalami beberapa kali perubahan. Pada saat ini, nomenklatur jenis validitas pengukuran yang kita gunakan adalah validitas isi, validitas kriteria, dan validitas konstruk (Arafah, 2017).

Validitas isi dilihat dari kisi-kisi tes, yaitu matriks yang menunjukkan bahan tes serta level kognitif yang terlibat dalam mengerjakan tes, validitas ini ditelaah sebelum tes digunakan. Validitas konstruk diperoleh dari hasil analisis faktor, yaitu jumlah faktor yang diukur suatu tes. Bukti validitas konstruk diperoleh dari hasil penggunaan tes, yaitu data empirik. Validitas prediktif juga memerlukan data empirik untuk dapat dihitung (Mardapi, 2008).

## **METODE PENELITIAN**

Penelitian ini merupakan studi kepustakaan dengan pendekatan kualitatif. Pengumpulan data dilakukan melalui pengumpulan kajian teori dari berbagai buku. Kemudian setiap hasil kajian teori dalam buku tersebut direview dan dianalisis secara seksama.

## HASIL DAN PEMBAHASAN

### Validitas Isi

Validitas isi pada pengukuran adalah kecocokan antara isi alat ukur dengan isi sasaran ukur. Agar dapat dicocokkan, sasaran atribut perlu dideskripsikan dengan cukup jelas. Berdasarkan keperluan pencocokan, biasanya, isi sasaran ukur disusun dalam bentuk spesifikasi, meliputi bahan atau materi pembelajaran dan tujuan pembelajaran. Pencocokan dilakukan butir demi butir melalui pencocokan dengan spesifikasi. Butir yang dinilai tidak baik atau tidak penting dapat dibuang, diperbaiki, atau diganti (Arafah, 2017).

Sebuah butir tes dikatakan memiliki validitas isi yang baik apabila mengukur tujuan pembelajaran khusus tertentu yang sejajar dengan materi atau bahan pelajaran yang diberikan. Validitas isi dapat diusahakan tercapai sejak saat penyusunan dengan cara memerinci materi kurikulum atau materi buku pelajaran dengan tepat (Arikunto, 2011).

Berdasarkan segi penyusunannya, maka telah dipertimbangkan secara rasional dan logis bahwa tes tersebut akan mengukur apa yang dimaksudkan akan diukur. Cara ini akan menghasilkan validitas isi (*content validity*). Isi tes tersebut merupakan sampel materi dari bahan uji secara keseluruhan dan dapat dikembangkan melalui tabel kisi-kisi (Slameto, 1999).

Bukti validitas berdasarkan isi dapat diperoleh dari suatu analisis hubungan antara isi tes dan konstruk yang ingin diukur. Isi tes mengacu pada tema, kata-kata, format butir, tugas, atau pertanyaan pada tes, seperti juga prosedur administrasi dan penskoran. Bukti validitas isi dapat diperoleh dari analisis hubungan antara isi tes dengan konstruk yang ingin diukur (Mardapi, 2008).

Pengembang tes sering menggunakan spesifikasi domain isi tes. Spesifikasi ini menjelaskan isi secara rinci, sering dengan spesifikasi cakupan isi dan tipe butir. Bukti berdasarkan isi bisa berupa analisis ketepatan isi tes secara logika atau empirik untuk membuat penafsiran skor hasil tes. Bukti validitas isi dapat berupa *judgement* pakar untuk menyatakan hubungan antara isi dan konstruk tes. Pakar yang terlibat adalah yang ahli dalam bidang yang diukur (Mardapi,

2008).

Validasi isi juga berkaitan dengan sejauh mana suatu butir tes mencakup keseluruhan materi atau bahan yang ingin diukur. Sejahter mana suatu tes memiliki bukti validitas ini ditetapkan menurut analisis rasional terhadap isi tes, yang penilaiannya didasarkan atas pertimbangan subjektif individual. Walaupun subjektif, namun yang terlibat adalah beberapa pakar pada bidang yang diukur sehingga hasilnya dapat dipertanggungjawabkan (Mardapi, 2008).

Validitas isi adalah validitas yang akan mengecek kecocokan di antara butir-butir tes yang dibuat dengan indikator, materi atau tujuan pembelajaran yang telah ditetapkan. Butir tes dinyatakan valid, jika butir-butir tes yang dibuat secara tepat dapat mengukur indikator (Djaali dan Puji, 2004). Suatu tes dikatakan memiliki validitas isi, apabila butir-butir yang disusun sesuai dengan materi pelajaran dan indikator yang telah ditetapkan (Susetyo, 2015).

Validitas isi pada tes hasil belajar diketahui dengan melakukan pencocokan antara materi dan butir tes dengan jalan mencocokkan antara isi butir tes hasil belajar dan indikator-indikator yang telah ditetapkan pada setiap topik pembelajaran. Suatu tes telah memenuhi validitas isi, apabila butir tes tersebut telah mencerminkan isi indikator yang telah ditetapkan (Susetyo, 2015).

Kegunaan dari validitas isi adalah untuk memperbaiki alat ukur melalui pemeriksaan butir pada alat ukur dan butir yang tidak baik atau tidak memenuhi syarat dibuang, diperbaiki, atau diganti. Cara yang paling umum adalah validasi melalui beberapa orang pakar. Butir dianggap tidak baik atau tidak memenuhi syarat jika dianggap tidak penting atau tidak cocok oleh mayoritas pakar (Arafah, 2017).

Validitas isi dilakukan oleh pakar atau praktisi yang ahli dalam bidang atau materi yang ingin diukur. Validitas isi dilakukan dengan mencocokkan isi butir tes dengan kompetensi dasar, indikator, tujuan pembelajaran, dan level kognitif. Validitas isi selanjutnya dapat dikuantitatifkan dengan membandingkan skor hasil validasi oleh pakar atau praktisi dengan standar validitas butir yang sudah ada. Perhitungan validitas isi dapat dilakukan dengan memilih salah satu

perhitungan validitas yang ada, seperti CVR oleh Lawshe, indeks validitas oleh Aiken, dan validitas isi oleh Gregory. Butir tes yang tidak valid dapat direvisi atau dibuang.

### **Validitas Kriteria**

Fungsi validitas kriteria adalah untuk memperbaiki alat ukur melalui uji coba alat ukur. Terdapat dua sasaran ukur, yakni sasaran ukur prediktor (menghasilkan skor prediktor) dan sasaran ukur kriteria (menghasilkan skor kriteria). Misalnya, sasaran ukur prediktor adalah ujian penerimaan karyawan baru atau mahasiswa baru dan sasaran ukur kriteria adalah keberhasilan mereka sebagai karyawan atau sebagai mahasiswa (Arafah, 2017).

Validitas kriteria merupakan kecocokan di antara prediktor (skor prediktor) dengan kriteria (skor kriteria). Validitas kriteria ditunjukkan kepada baik atau tidak baiknya prediktor (skor prediktor). Jika validitas kriteria baik, maka alat ukur prediktor (skor prediktor) dapat digunakan untuk berbagai keperluan sejenis. Jenis validitas kriteria, yaitu validitas serentak yakni kriteria terdapat pada saat yang sama dengan prediktor dan validitas prediksi yakni kriteria terdapat kemudian setelah prediktor (Arafah, 2017).

Analisis hubungan skor tes dengan variabel eksternal dilakukan untuk melengkapi bukti validitas. Variabel eksternal bisa berupa kriteria bahwa tes diharapkan memprediksi, seperti hubungan dengan tes lain yang mengukur kostrak yang sama dan tes lain yang mengukur hal yang berbeda. Kriteria lain seperti kriteria performans, sering digunakan untuk keperluan seleksi atau penempatan pegawai atau karyawan. Bukti validitas berdasarkan hubungan dengan variabel lain sering disebut dengan validitas kriteria (Mardapi, 2008).

Prediktor dan kriteria terjadi pada waktu yang bersamaan dan menggunakan skor prediktor untuk mengetahui kriteria. Pada kriteria terdapat alat ukur yang sudah biasa dipakai dan memberikan hasil yang baik. Prediktornya adalah alat ukur rakitan baru untuk mengganti atau mendampingi alat ukur yang biasa dipakai, alat ukur rakitan baru yang lebih sederhana, atau alat ukur rakitan baru yang menggunakan cara ukur atau

besaran lain (Arafah, 2017).

Pada pengujian validitas konkuren, data yang mencerminkan pengalaman yang diperoleh pada masa lalu yang telah valid, dibandingkan dengan data hasil tes yang diperoleh sekarang ini. Jika hasil tes yang ada sekarang ini mempunyai hubungan searah dengan hasil tes berdasar pengalaman yang lalu, maka tes yang memiliki karakteristik seperti itu dapat dikatakan telah memiliki validitas konkuren (Sudijono, 2012).

Membandingkan hasil sebuah tes maka diperlukan suatu kriteria atau alat banding. Maka hasil tes merupakan sesuatu yang dibandingkan. Misalnya seorang guru ingin mengetahui apakah tes sumatif yang disusun sudah valid atau belum, jadi diperlukan sebuah kriteria masa lalu yang sekarang datanya dimiliki. Misalnya nilai ulangan sumatif yang telah lalu (Arikunto, 2011).

Hubungan antara dua alat ukur sebagai dasar perhitungan, satu alat ukur prediktor merupakan alat ukur baru dan alat ukur lainnya sebagai kriteria, yaitu alat ukur yang biasa dipakai. Perangkat ukur baru diberikan secara bersama-sama dengan alat ukur yang biasa digunakan. Contoh perangkat ukur TPA yang baru dibuat untuk mengganti atau cadangan alat ukur TPA yang biasa dipakai. Soal ujian cadangan untuk menggantikan soal ujian yang biasa dipakai. TPA melalui komputer untuk menggantikan TPA melalui pensil dan kertas. Penentuan validitas serentak berdasarkan kecocokan antara alat ukur baru dan alat ukur yang biasa digunakan dengan menggunakan perhitungan korelasi (Susetyo, 2015).

Beberapa contoh yang lain, yaitu TOEFL melalui komputer untuk menggantikan atau mendampingi TOEFL melalui pensil dan kertas, alat uji olahraga cara baru di dalam ruangan untuk menggantikan atau mendampingi alat uji olahraga lari satu kilometer di luar ruangan, soal ujian cadangan untuk menggantikan soal ujian jika terjadi kebocoran sebelum ujian, dan simulasi menyetir mobil di layar komputer untuk mendampingi menyetir mobil sesungguhnya (Arafah, 2017).

Validitas prediktif, yaitu kriteria terjadi kemudian setelah prediktor dan menggunakan skor prediktor untuk menge-

tahui kriteria. Alat ukur yang dipakai untuk mengetahui kriteria, misalnya ujian penerimaan karyawan untuk menyeleksi calon karyawan atau ujian penerimaan mahasiswa untuk menyeleksi calon mahasiswa. Sasaran terjadi kemudian. Biasanya tidak mudah untuk ditentukan secara tepat. Sukar menentukan kriteria karyawan yang baik atau mahasiswa yang berhasil (Arafah, 2017).

Setiap kali kita menyebutkan istilah ramalan, maka di dalamnya akan mengandung pengertian mengenai sesuatu yang bakal terjadi di masa mendatang atau sesuatu yang sekarang belum terjadi dan baru akan terjadi pada waktu yang akan datang. Apabila istilah ramalan itu dikaitkan dengan validitas tes, maka yang dimaksud dengan validitas prediksi dari suatu tes adalah suatu kondisi yang menunjukkan seberapa jauh sebuah tes dengan tepat menunjukkan kemampuannya untuk meramalkan yang akan terjadi pada masa yang akan datang (Sudijono, 2012).

Contohnya 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 mengikuti kuliah. Jika nilai tesnya tinggi akan menjamin keberhasilannya kelak. Sebaliknya seorang calon dikatakan tidak lulus tes karena memiliki nilai tes yang rendah jadi diperkirakan akan tidak mampu mengikuti perkuliahan yang akan datang (Arikunto, 2011).

Validitas prediktif diperoleh apabila pengambilan skor kriteria tidak bersamaan dengan pengambilan skor tes. Setelah subjek dikenai tes yang akan dicari validitas prediktifnya, lalu diberikan tenggang waktu tertentu sebelum skor kriteria diambil dari subjek yang sama. Umpamanya, untuk melihat validitas prediktif tes masuk pendidikan tinggi, maka skor kriterianya dapat diambil dari indeks prestasi yang dicapai setelah beberapa semester atau beberapa tahun kemudian. Koefisien korelasi antara skor tes masuk pendidikan tinggi dan indeks prestasi tersebut merupakan koefisien validitas prediktif tes masuk. Contoh lain adalah bila kita mengkorelasikan suatu tes bakat untuk penempatan karyawan dengan performansi kerjanya setelah satu atau dua tahun. Dalam

hal ini performansi kerja merupakan kriteria untuk melihat validitas prediktif tes bakat (Mardapi, 2008).

Korelasi antara dua alat ukur sebagai dasar perhitungan yang mana satu alat ukur digunakan sebagai prediksi dan alat ukur lainnya digunakan sebagai kriteria. Perangkat ukur prediksi diberikan terlebih dahulu, kemudian nantinya digunakan untuk mengetahui kriteria yang diinginkan. Penentuan validitas prediktif berdasarkan kecocokan di antara prediksi dengan kriteria yang dilakukan melalui penghitungan koefisien korelasi (Susetyo, 2015).

### Validitas Konstruk

Validitas konstruk mengacu pada sejauh mana suatu tes mengukur konsep dari suatu teori, yaitu yang menjadi dasar penyusunan tes. Pengumpulan bukti validitas konstruk merupakan proses yang terus berlanjut sejalan dengan perkembangan konsep mengenai *trait* yang diukur (Mardapi, 2008).

Syarat validasi, yaitu ada referensi yang telah diketahui untuk digunakan pada perbandingan dengan konstruk dan ada referensi yang layak dijadikan pembanding untuk digunakan pada perbandingan dengan konstruk. Jenis referensi, yaitu referensi cocok yakni referensi yang cocok dengan konstruk dan referensi beda yakni referensi yang beda dengan konstruk (Arafah, 2017).

Terdapat dua metode pencocokan, yaitu secara konvergen dan diskriminan. Pencocokan secara konvergen adalah pencocokan konstruk yang terdapat dalam perangkat ukur dengan rujukan yang cocok. Suatu perangkat ukur tergolong tinggi validitas konstraknya jika terdapat kecocokan butir tes yang dibuat dengan perangkat ukur butir tes yang telah valid menurut prosedur validitas konstruk. Adapun pencocokan diskriminan merupakan pencocokan dengan rujukan yang berbeda atau kontras dengan konstruk dalam suatu perangkat ukur. Suatu perangkat ukur tergolong tinggivaliditas konstraknya jika terdapat ketidakcocokan butir-butir tes yang dibuat dengan perangkat ukur yang butir-butir tes berbeda dan telah memenuhi validitas konstruk. Prosedur perhitungan validitas konstruk, baik konvergen maupun diskriminan menggunakan teknik korelasi atau analisis faktor (Susetyo, 2015).

## **SIMPULAN DAN SARAN**

Simpulan dari penelitian ini adalah suatu tes dikatakan valid jika mengukur apa yang hendak diukur. Dengan kata lain suatu tes dikatakan valid jika sesuai dengan tujuan tes tersebut diadakan. Validitas terbagi tiga, yaitu validitas isi, validitas kriteria, dan validitas konstruk. Validasi dilakukan sebelum dan setelah tes diujicobakan. Sebelum diujicobakan, kita dapat mengetahui validitas isi tes berdasarkan hasil validasi oleh pakar. Validitas kriteria dan validitas konstruk dapat diketahui setelah tes diujicobakan. Jenis validitas yang terbaru adalah validitas isi dan validitas isi telah menjamin bahwa tes dibuat telah valid. Jadi validitas isi adalah validitas yang paling utama harus dilakukan sebelum kita menggunakan suatu tes untuk mengukur capaian pembelajaran peserta didik. Sehingga hasil tes tersebut benar-benar menggambarkan hasil belajar dan capaian pembelajaran pembelajaran peserta didik.

Saran dari penelitian ini adalah terdapat penelitian lanjutan yang dapat menerapkan teori validitas tes ini. Bahkan tidak hanya pada instrumen tes, namun juga pada instrumen non tes.

## **DAFTAR RUJUKAN**

Allen, Mary J. (1979). *Introduction to Measurement Theory*. California:

Brooks/Cole Publishing Company.  
Arafah, Kaharuddin. (2017). *Bahan Kuliah Teori Tes*. Makassar: Universitas Negeri Makassar. Arikunto, Suharsimi. 2011. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: PT Bumi Aksara.  
Mardapi, Djemari. (2008). *Teknik Penyusunan Instrumen Tes dan Nontes*. Yogyakarta: MitraCendikia Press.  
Slameto. 1999. *Evaluasi Pendidikan*. Jakarta: PT Bumi Aksara.  
Sudijono, Anas. 2012. *Pengantar Evaluasi Pendidikan*. Jakarta: PT Rajagrafindo Persada. Susetyo, Budi. 2015. *Prosedur Penyusunan dan Analisis Tes*. Bandung: PT Refika Aditama.