

## VALIDITAS DAN RELIABILITAS SUATU INSTRUMEN PENELITIAN

Zulkifli Matondang \*

### Abstrak

*Instrumen merupakan suatu alat yang dipergunakan sebagai alat untuk mengukur suatu obyek ukur atau mengumpulkan data dari suatu variabel. Suatu instrumen dikatakan baik bila valid dan reliabel. Validitas terdiri atas validitas isi, konstruk, empirik. Validitas internal skor butir dikotomi dan skor butir politomi berturut-turut digunakan korelasi biserial dan korelasi product moment. Kriteria suatu butir valid atau tidak valid didasarkan pada nilai r-tabel. Reliabilitas konsistensi gabungan butir untuk skor butir dikotomi dan skor butir politomi berturut-turut digunakan KR-20 dan koefisien Alpha. Interpretasi terhadap koefisien reliabilitas merupakan interpretasi relatif dalam artian bahwa tidak ada batasan mutlak yang menunjukkan berapa angka koefisien minimal yang harus dicapai agar suatu pengukuran dapat disebut reliabel. Namun, memberikan informasi tentang hubungan varians skor teramati dengan skor sejati kelompok individu.*

**Kata kunci: Validitas, Reliabilitas, Instrumen Penelitian.**

### A. PENDAHULUAN

Dalam penelitian pendidikan khususnya penelitian kuantitatif dikenal dengan nama variabel, misalnya variabel laten, variabel manifes dan sebagainya. Variabel inilah yang pada umumnya ingin diketahui karakteristik yang dimilikinya, misalnya rata-rata, median, modus, standar deviasi dan lain-lain.

Untuk mengukur suatu variabel diperlukan alat ukur yang biasa disebut instrumen. Djaali (2000: 9) menyatakan bahwa secara umum yang dimaksud dengan instrumen adalah suatu alat yang karena memenuhi persyaratan akademis maka dapat dipergunakan sebagai alat untuk mengukur suatu obyek ukur atau mengumpulkan data mengenai suatu variabel. Selanjutnya dinyatakan bahwa pada dasarnya instrumen dapat dibagi menjadi dua macam, yakni tes dan

non-tes. Yang termasuk kelompok tes, misalnya tes prestasi belajar, tes inteligensi, tes bakat; sedangkan yang termasuk non-tes misalnya pedoman wawancara, angket atau kuesioner, lembar observasi, daftar cocok (*check list*), skala sikap, skala penilaian, dan sebagainya. Dalam hal pengukuran, Weitzenhoffer (dalam Nur, 1987: 1) menyatakan bahwa pengukuran sebagai suatu operasi yang dilakukan terhadap alam fisik oleh pengamat. Misalnya, ingin mengukur hasil belajar, intelegensi, sikap, motivasi berprestasi, dan sebagainya. Sekarang muncul suatu pertanyaan, yaitu apakah suatu alat ukur benar-benar mengukur apa yang hendak dan seharusnya diukur serta sejauh mana alat ukur tersebut dapat diandalkan dan berguna, sebenarnya menunjuk pada dua hal yang pokok, yaitu validitas dan reliabilitas.

Nurkencana (1992: 141) menyatakan bahwa suatu alat pengukur dapat dikatakan alat pengukur yang valid apabila alat pengukur tersebut dapat mengukur apa yang hendak diukur secara tepat. Dalam hal validitas dan reliabilitas, tentunya dipengaruhi oleh (1) instrumen, (2) subjek yang diukur, dan (3) petugas yang melakukan pengukuran. Dalam hal pengukuran, khususnya dalam pendidikan tentunya yang terpenting adalah informasi hasil ukur yang benar. Sebab dengan hasil ukur yang tidak atau kurang tepat maka akan memberikan informasi yang tidak benar, sehingga kesimpulan yang diambil juga tidak benar.

Steven (dalam Nur, 1987: 1) menyatakan bahwa pengukuran adalah pemberian angka atas objek atau kejadian sesuai dengan aturan. Dengan menitikberatkan pada alat ukurnya, maka dalam hal ini yang akan dibahas instrumen tes.

## **B. PEMBAHASAN**

### **1. Pengertian Tes**

Tes adalah prosedur sistematis yang dibuat dalam bentuk tugas-tugas yang distandardisasikan dan diberikan kepada individu atau kelompok untuk dikerjakan, dijawab, atau direspon, baik dalam bentuk tertulis, lisan maupun perbuatan. Silvirius (1991: 5) menyatakan bahwa tes adalah suatu prosedur sistematis untuk mengamati dan mencandracikan satu atau lebih karakteristik seseorang dengan menggunakan skala numerik atau sistem kategori.

Tes juga dapat diartikan sebagai alat pengukur yang mempunyai standar obyektif sehingga dapat dipergunakan untuk

mengukur dan membandingkan keadaan psikis atau tingkah laku individu. Azwar (1987: 3) menyatakan bahwa tes adalah prosedur yang sistematis, maksudnya (a) butir-butir dalam tes disusun menurut cara dan aturan tertentu, (b) prosedur administrasi tes dan pemberian angka (*scoring*) terhadap hasilnya harus jelas dan dispesifikasi secara terperinci, dan (c) setiap orang yang mengambil tes itu harus mendapat butir-butir yang sama dalam kondisi yang sebanding.

## 2. Validitas

Azwar (1987: 173) menyatakan bahwa validitas berasal dari kata *validity* yang mempunyai arti sejauh mana ketepatan dan kecermatan suatu instrumen pengukur (tes) dalam melakukan fungsi ukurnya. Suatu tes dikatakan memiliki validitas yang tinggi apabila alat tersebut menjalankan fungsi ukur secara tepat atau memberikan hasil ukur yang sesuai dengan maksud dilakukannya pengukuran tersebut. Artinya hasil ukur dari pengukuran tersebut merupakan besaran yang mencerminkan secara tepat fakta atau keadaan sesungguhnya dari apa yang diukur.

Suryabrata (2000: 41) menyatakan bahwa validitas tes pada dasarnya menunjuk kepada derajat fungsi pengukurnya suatu tes, atau derajat kecermatan ukurnya sesuatu tes. Validitas suatu tes mempermasalahkan apakah tes tersebut benar-benar mengukur apa yang hendak diukur. Maksudnya adalah seberapa jauh suatu tes mampu mengungkapkan dengan tepat ciri atau keadaan yang sesungguhnya dari obyek ukur, akan tergantung dari tingkat validitas tes yang bersangkutan. Sudjana (2004: 12) menyatakan bahwa validitas berkenaan dengan ketepatan alat penilaian terhadap konsep yang dinilai sehingga betul-betul menilai apa yang seharusnya dinilai.

Suatu tes yang valid untuk tujuan tertentu atau pengambilan keputusan tertentu, mungkin tidak valid untuk tujuan atau pengambilan keputusan lain. Jadi validitas suatu tes, harus selalu dikaitkan dengan tujuan atau pengambilan keputusan tertentu. Tes masuk di SMA misalnya harus selalu dikaitkan dengan seberapa jauh tes masuk tersebut dapat mencerminkan prestasi atau hasil belajar para calon peserta didik baru setelah belajar nanti.

Konsep validitas tes dapat dibedakan atas tiga macam yaitu validitas isi (*content validity*), validitas konstruk (*construct validity*), dan validitas empiris atau validitas kriteria. Validitas isi suatu tes mempermasalahkan seberapa jauh suatu tes mengukur tingkat

penguasaan terhadap isi atau konten atau materi tertentu yang seharusnya dikuasai sesuai dengan tujuan pengajaran. Dengan kata lain tes yang mempunyai validitas isi yang baik ialah tes yang benar-benar mengukur penguasaan materi yang seharusnya dikuasai sesuai dengan konten pengajaran yang tercantum dalam Garis-garis Besar Program Pengajaran (GBPP).

Validitas isi menunjukkan sejauhmana pertanyaan, tugas atau butir dalam suatu tes atau instrumen mampu mewakili secara keseluruhan dan proporsional perilaku sampel yang dikenai tes tersebut. Artinya tes itu valid apabila butir-butir tes itu mencerminkan keseluruhan konten atau materi yang diujikan atau yang seharusnya dikuasai secara proporsional.

Untuk mengetahui apakah tes itu valid atau tidak, harus dilakukan melalui penelaahan kisi-kisi tes untuk memastikan bahwa soal-soal tes itu sudah mewakili atau mencerminkan keseluruhan konten atau materi yang seharusnya dikuasai secara proporsional. Oleh karena itu validitas isi suatu tes tidak mempunyai besaran tertentu yang dihitung secara statistika tetapi dipahami bahwa tes itu sudah valid berdasarkan telaah kisi-kisi tes. Oleh karena itu, validitas isi sebenarnya mendasarkan pada analisis logika, tidak merupakan suatu koefisien validitas yang dihitung secara statistika.

Validitas konstruk (*construct validity*) adalah validitas yang mempermasalahkan seberapa jauh butir-butir tes mampu mengukur apa yang benar-benar hendak diukur sesuai dengan konsep khusus atau definisi konseptual yang telah ditetapkan. Validitas konstruk biasa digunakan untuk instrumen yang dimaksudkan mengukur variabel konsep, baik yang sifatnya performansi tipikal seperti instrumen untuk mengukur sikap, minat konsep diri, lokus kontrol, gaya kepemimpinan, motivasi berprestasi, dan lain-lain, maupun yang sifatnya performansi maksimum seperti instrumen untuk mengukur bakat (tes bakat), inteligansi (kecerdasan intelektual), kecerdasan, emosional dan lain-lain.

Untuk menentukan validitas konstruk dilakukan proses penelaahan teoretik dari suatu konsep dari variabel yang hendak diukur, mulai dari perumusan konstruk, penentuan dimensi dan indikator, sampai kepada penjabaran dan penulisan butir-butir instrumen. Perumusan, konstruk harus dilakukan berdasarkan sintesis dari teori-teori mengenai konsep variabel yang hendak diukur melalui proses analisis dan komparasi yang logik dan cermat.

Menyimak proses telaah teoretik seperti telah dikemukakan, maka proses validasi konstruk sebuah instrumen dilakukan melalui penelaahan atau justifikasi pakar atau melalui penilaian sekelompok panel yang terdiri dari orang-orang yang menguasai substansi atau konten dari variabel yang hendak diukur.

Validitas empiris sama dengan validitas kriteria yang berarti bahwa validitas ditentukan berdasarkan kriteria, baik kriteria internal maupun kriteria eksternal. Validitas empiris diperoleh melalui hasil uji coba tes kepada responden yang setara dengan responden yang akan dievaluasi atau diteliti. Kriteria internal adalah tes atau instrumen itu sendiri yang menjadi kriteria, sedang kriteria eksternal adalah hasil ukur instrumen atau tes lain di luar instrumen itu sendiri yang menjadi kriteria. Ukuran lain yang sudah dianggap baku atau dapat dipercaya dapat pula dijadikan sebagai kriteria eksternal. Validitas yang ditentukan berdasarkan kriteria internal disebut validitas internal sedangkan validitas yang ditentukan berdasarkan kriteria eksternal disebut validitas eksternal.

Validitas internal (validitas butir) termasuk kelompok validitas kriteria yang merupakan validitas yang diukur dengan besaran yang menggunakan tes sebagai suatu kesatuan (keseluruhan butir) sebagai kriteria untuk menentukan validitas butir dari tes itu. Dengan demikian validitas internal memperlakukan validitas butir dengan menggunakan hasil ukur tes tersebut sebagai suatu kesatuan sebagai kriteria, sehingga biasa juga disebut validitas butir. Validitas internal diperlihatkan oleh seberapa jauh hasil ukur butir tersebut konsisten dengan hasil ukur tes secara keseluruhan. Oleh karena itu validitas butir tercermin pada besaran koefisien korelasi antara skor butir dengan skor total tes. Jika koefisien korelasi skor butir dengan skor total tes positif dan signifikan maka butir tersebut valid berdasarkan ukuran validitas internal.

Koefisien korelasi yang tinggi antara skor butir dengan skor total mencerminkan tingginya konsistensi antara hasil ukur keseluruhan tes dengan hasil ukur butir tes atau dapat dikatakan bahwa butir tes tersebut konvergen dengan butir-butir lain dalam mengukur suatu konsep atau konstruk yang hendak diukur.

Djaali (2000: 77) menyatakan bahwa untuk menghitung validitas internal untuk skor butir dikotomi digunakan koefisien korelasi biserial (rbis) dengan rumus:

$$r_{bis(i)} = \frac{\overline{X_i} - \overline{X_t}}{S_t} \sqrt{\frac{p_i}{q_i}}$$

Keterangan:

$r_{bis(i)}$  = koefisien korelasi antara skor butir ke i dengan skor total.

$\overline{X_i}$  = rata-rata skor total responden yang menjawab benar butir ke i.

$\overline{X_t}$  = rata-rata skor total semua responden.

$S_t$  = standar deviasi skor total semua responden.

$p_i$  = proporsi jawaban yang benar untuk butir ke i.

$q_i$  = proporsi jawaban yang salah untuk butir ke i.

Selanjutnya, dikatakan bahwa untuk menghitung koefisien validitas internal untuk skor butir politomi digunakan korelasi product moment (r) dengan rumus:

$$r_{it} = \frac{\sum x_i x_t}{\sqrt{\sum x_i^2 \sum x_t^2}}$$

dengan:

$r_{it}$  = koefisien korelasi antara skor butir soal dengan skor total.

$\sum x_i$  = jumlah kuadrat deviasi skor dari  $X_i$ .

$\sum x_t$  = jumlah kuadrat deviasi skor dari  $X_t$ .

Nilai koefisien korelasi baik skor butir dikotomi maupun skor butir politomi untuk masing-masing butir dibandingkan dengan nilai koefisien korelasi yang ada di tabel-r pada alpha tertentu misalnya  $\alpha = 0,05$ . Jika koefisien korelasi skor butir dengan skor total lebih besar dari koefisien korelasi dari tabel-r, koefisien korelasi butir signifikan dan butir tersebut valid secara empiris.

Validitas eksternal dapat berupa hasil ukur tes baku atau tes yang dianggap baku dapat pula berupa hasil ukur lain yang sudah tersedia dan dapat dipercaya sebagai ukuran dari suatu konsep atau variabel yang hendak diukur. Validitas eksternal diperlihatkan oleh suatu besaran yang merupakan hasil perhitungan statistika.

Jika kita menggunakan hasil ukur tes yang sudah baku sebagai kriteria eksternal, maka besaran validitas eksternal dari tes yang kita kembangkan didapat dengan jalan mengkorelasikan skor hasil ukur tes yang dikembangkan dengan skor hasil ukur tes baku yang dijadikan kriteria. Makin tinggi koefisien korelasi yang didapat, maka validitas

tes yang dikembangkan juga makin baik. Kriteria yang digunakan untuk menguji validitas eksternal digunakan nilai r-tabel.

Jika koefisien korelasi antara skor hasil ukur tes yang dikembangkan dengan skor hasil ukur tes baku lebih besar daripada r-tabel maka tes yang dikembangkan adalah valid berdasarkan kriteria eksternal yang dipilih (hasil ukur instrumen baku). Jadi keputusan uji-validitas dalam hal ini adalah mengenai valid atau tidaknya tes sebagai suatu kesatuan, bukan valid atau tidaknya butir tes seperti pada validitas internal.

### 3. Reliabilitas

Reliabilitas berasal dari kata *reliability* berarti sejauh mana hasil suatu pengukuran dapat dipercaya. Suatu hasil pengukuran dapat dipercaya apabila dalam beberapa kali pelaksanaan pengukuran terhadap kelompok subyek yang sama, diperoleh hasil pengukuran yang relatif sama, selama aspek yang diukur dalam diri subyek memang belum berubah. Nur (1987: 47) menyatakan bahwa reliabilitas ukuran menyangkut seberapa jauh skor deviasi individu, atau skor-z, relatif konsisten apabila dilakukan pengulangan pengadministrasian dengan tes yang sama atau tes yang ekuivalen.

Azwar (2003 : 176) menyatakan bahwa reliabilitas merupakan salah-satu ciri atau karakter utama instrumen pengukuran yang baik. Arifin (1991: 122) menyatakan bahwa suatu tes dikatakan reliabel jika selalu memberikan hasil yang sama bila diteskan pada kelompok yang sama pada waktu atau kesempatan yang berbeda.

Konsep reliabilitas dalam arti reliabilitas alat ukur berkaitan erat dengan masalah kekeliruan pengukuran. Kekeliruan pengukuran sendiri menunjukkan sejauh mana inkonsistensi hasil pengukuran terjadi apabila dilakukan pengukuran ulang terhadap kelompok subyek yang sama. Sedangkan konsep reliabilitas dalam arti reliabilitas hasil ukur berkaitan erat dengan kekeliruan dalam pengambilan sampel yang mengacu pada inkonsistensi hasil ukur apabila pengukuran dilakukan ulang pada kelompok yang berbeda. Sudjana (2004: 16) menyatakan bahwa reliabilitas alat penilaian adalah ketepatan atau keajegan alat tersebut dalam menilai apa yang dinilainya. Artinya, kapanpun alat penilaian tersebut digunakan akan memberikan hasil yang relatif sama.

Djaali (2000: 81) menyatakan bahwa reliabilitas dibedakan atas dua macam, yaitu reliabilitas konsistensi tanggapan, dan reliabilitas

konsistensi gabungan butir. Reliabilitas konsistensi tanggapan responden mempersoalkan apakah tanggapan responden atau obyek ukur terhadap tes atau instrumen tersebut sudah baik atau konsisten. Dalam hal ini apabila suatu tes atau instrumen digunakan untuk melakukan pengukuran terhadap obyek ukur kemudian dilakukan pengukuran kembali terhadap obyek ukur yang sama, apakah hasilnya masih tetap sama dengan pengukuran sebelumnya. Jika hasil pengukuran kedua menunjukkan ketidakkonsistenan maka jelas hasil pengukuran itu tidak mencerminkan keadaan obyek ukur yang sesungguhnya.

Untuk mengetahui apakah tanggapan terhadap tes atau instrumen itu mantap, konsisten atau tidak plin-plan, dapat dilakukan dengan cara memberikan tes yang sama secara berulang kali (dua kali) kepada obyek ukur atau responden yang sama. Pengetesan dua kali merupakan syarat minimal untuk mengetahui apakah tanggapan obyek ukur terhadap tes tersebut konsisten atau tidak.

Dalam pelaksanaan pengetesan dua kali ini dapat ditempuh berbagai cara yaitu kita melakukan pengetesan dua kali dengan tes sama terhadap obyek ukur yang sama, atau dengan melakukan pengetesan sekali dengan menggunakan dua tes yang butir-butirnya setara. Jika kita menggunakan pengetesan sekali maka kesamaan atau kesetaraan tes yang digunakan merupakan syarat mutlak yang harus dipenuhi, karena kemantapan atau konsistensi tanggapan terhadap butir-butir yang akan diperiksa.

Pada teknik belah dua ini pengukuran dilakukan dengan dua kelompok butir yang setara pada saat yang sama. Karena setiap kelompok butir merupakan separuh dari seluruh tes, maka biasanya kelompok butir pertama diambil dari butir-butir tes yang bernomor ganjil, sedangkan kelompok butir yang kedua diambil dari butir-butir tes yang bernomor genap. Perlu diketahui bahwa reliabilitas dengan teknik ini sangat relatif, karena reliabilitas akan tergantung pada cara penomoran dan pengelompokan butir yang diambil. Di sini pengukuran dilakukan dengan menggunakan dua tes yang dibuat setara kemudian diberikan kepada responden atau obyek tes dalam waktu yang bersamaan. Skor dari kedua kelompok butir tes tersebut dikorelasikan untuk mendapatkan reliabilitas tes.

Djaali (2000: 81) menyatakan bahwa reliabilitas konsistensi gabungan butir berkaitan dengan kemantapan antara butir suatu tes. Hal ini dapat diungkapkan dengan pertanyaan, apakah terhadap obyek

ukur yang sama, butir yang satu menunjukkan hasil ukur yang sama dengan butir yang lainnya? Dengan kata lain bahwa terhadap bagian obyek ukur yang sama, apakah hasil ukur butir yang satu tidak kontradiksi dengan hasil ukur butir yang lain.

Jika terhadap bagian obyek ukur yang sama, hasil ukur melalui butir yang satu kontradiksi atau tidak konsisten dengan hasil ukur melalui butir yang lain maka pengukuran dengan tes (alat ukur) sebagai suatu kesatuan itu tidak dapat dipercaya. Dengan kata lain tidak reliabel dan tidak dapat digunakan untuk mengungkap ciri atau keadaan yang sesungguhnya dari obyek ukur. Kalau hasil pengukuran pada bagian obyek ukur yang sama antara butir yang satu dengan butir yang lain saling kontradiksi atau tidak konsisten maka kita jangan menyalahkan obyek ukur, melainkan alat ukur (tes) yang dipersalahkan dengan mengatakan bahwa tes tersebut tidak reliabel terhadap obyek yang diukur.

Koefisien reliabilitas konsistensi gabungan butir untuk skor butir dikotomi dapat dihitung dengan menggunakan rumus Kuder-Richardson yang dikenal dengan nama KR-20 (Djaali, 2000: 77) dengan rumus:

$$KR-20 = \frac{k}{k-1} \left( 1 - \frac{\sum p_i q_i}{S_t^2} \right)$$

Keterangan:

$k$  = cacah butir.

$p_i q_i$  = varians skor butir.

$p_i$  = proporsi jawaban yang benar untuk butir nomor  $i$ .

$q_i$  = proporsi jawaban yang salah untuk butir nomor  $i$ .

$S_t^2$  = varians skor total responden.

Koefisien reliabilitas gabungan butir untuk skor butir politomi, maka koefisien reliabilitas dihitung menggunakan koefisien Alpha (Djaali, 2000: 122) dengan rumus:

$$r_{ii} = \frac{k}{k-1} \left( 1 - \frac{\sum s_i^2}{s_t^2} \right)$$

Keterangan:

$r_{ii}$  = koefisien reliabilitas.

$k$  = cacah butir.

$s_i^2$  = varians skor butir.

$s_t^2$  = varians skor total responden.

Interpretasi terhadap koefisien reliabilitas merupakan interpretasi relatif, artinya tidak ada batasan mutlak yang menunjukkan berapa angka koefisien minimal yang harus dicapai agar suatu pengukuran dapat disebut reliabel. Namun, memberikan informasi tentang hubungan varians skor teramati dengan varians skor sejati kelompok individu. Misalnya, diperoleh koefisien reliabilitas sama dengan 0,87. Koefisien reliabilitas ini dapat diartikan bahwa: (1) 87% varians skor teramati diakibatkan oleh varians skor sejati kelompok individu, dan (2) korelasi antara skor teramati dan skor sejati sama dengan  $\sqrt{0,87}$  atau 0,93 (Nur, 1987: 61).

### C. KESIMPULAN

Instrumen merupakan suatu alat yang karena memenuhi persyaratan akademis maka dapat dipergunakan sebagai alat untuk mengukur suatu obyek ukur atau mengumpulkan data mengenai suatu variabel. Validitas isi mempermasalahkan sejauh mana suatu tes mengukur tingkat penguasaan terhadap isi atau materi tertentu yang seharusnya dikuasai sesuai dengan tujuan pengajaran dan validitas isi tidak mempunyai besaran. Validitas konstruk mempermasalahkan seberapa jauh butir-butir tes mampu mengukur apa yang benar-benar hendak diukur sesuai dengan definisi konseptual yang telah ditetapkan. Validitas empiris (validitas kriteria) yang berarti bahwa validitas ditentukan berdasarkan kriteria, baik kriteria internal maupun kriteria eksternal.

Reliabilitas mempermasalahkan sejauh mana hasil suatu pengukuran dapat dipercaya. Suatu hasil pengukuran hanya dapat dipercaya apabila dalam beberapa kali pelaksanaan pengukuran terhadap kelompok subyek yang sama, diperoleh hasil pengukuran yang relatif sama. Penentuan koefisien reliabilitas instrumen untuk skor butir dikotomi digunakan rumus KR-20, sedang untuk skor politomi digunakan rumus Alpha. Interpretasi terhadap koefisien reliabilitas merupakan interpretasi relatif, artinya tidak ada batasan mutlak yang menunjukkan berapa angka koefisien minimal yang harus dicapai agar suatu pengukuran dapat disebut reliabel. Namun, memberikan informasi tentang hubungan varians skor teramati dengan varians skor sejati kelompok individu.

Berdasarkan kesimpulan, maka beberapa saran yaitu: *Pertama*. Dalam penelitian pendidikan khususnya penelitian kuantitatif disaran-

kan kepada peneliti pemula bahwa instrumen yang mengukur variabel manifes digunakan instrumen yang valid, baik melalui validitas isi maupun validitas internal. *Kedua*. Bila instrumen yang akan digunakan adalah instrumen yang mengukur variabel laten, disarankan agar instrumen yang akan digunakan dalam penelitian adalah instrumen yang valid berdasarkan konstruk dan valid berdasarkan kriteria. *Ketiga*. Dalam menentukan validitas butir dan reliabilitas instrumen, disarankan menggunakan rumus yang tepat dengan memperhatikan apakah butir tersebut diskor dikotomi atau diskor politomi. *Keempat*. Koefisien reliabilitas yang telah diperoleh, baik instrumen yang diskor dikotomi maupun yang diskor politomi, disarankan agar menginterpretasi koefisien reliabilitasnya.

### DAFTAR PUSTAKA

- Azwar, Saifuddin. *Sikap Manusia Teori dan Pengukurannya*. Liberty: Yogyakarta, 1988.
- Azwar, Saifudidin. *Sikap Manusia Terori dan Pengukurannya*. Yokyakarta: Pustaka Pelajar, 2003.
- Djaali., dkk. *Pengukuran Dalam Pendidikan*. Jakarta: Program Pascasarjana, 2000.
- Nurkancana, Wayan., PPN. Sunartana. *Evaluasi Hasil Belajar*, Surabaya: Usaha Nasional, 1992.
- Nur, Mohamad. *Teori Tes*. Surabaya: IKIP Surabaya, 1987.
- Silverius, Suke. *Evaluasi Hasil Belajar dan Umpan Balik*. Jakarta: PT Gramedia Widiasarana Indonesia, 1991.
- Sudjana, Nana, *Penilaian Hasil Proses Belajar Mengajar*, Bandung: Remaja Rosdakarya, 2004.
- Suryabrata, Sumadi, *Pengembangan Alat Ukur Psikologis*, Yogyakarta: Andi, 2000

---

\* Dr. Zulkifli Matondang, M.Si adalah dosen Jurusan Pendidikan Teknik Bangunan FT dan Pascasarjana Unimed