# Konsep Instrumen Penelitian Pendidikan

Oleh: Baso Intang Sappaile\*)

strumen merupakan suatu alat yang memenuhi persyaratan Mademis sehingga dapat dipergunakan sebagai alat untuk mengukur objek ukur atau mengumpulkan data mengenai suatu variabel. Instrumen dikatakan baik bila valid dan reliabel, baik validitas bonstruk, validitas empirik, reliabilitas konsistensi tanggapan, mangun reliabilitas konsistensi gabungan butir. Validitas internal butir dikotomi dan validitas internal skor butir politomi berturutand digunakan korelasi biserial dan korelasi poduct moment. Kriteria butir valid atau tidak valid didasarkan pada nilai r-tabel. Belighilitas konsistensi gabungan butir untuk skor butir dikotomi de skor butir politomi berturut-turut digunakan KR-20 dan koefisien Interpretasi terhadap koefisien reliabilitas merupakan morepretasi relatif dalam artian bahwa tidak ada batasan mutlak menunjukkan berapa angka koefisien minimal yang harus dapat agar suatu pengukuran dapat disebut reliabel. Namun, dapat memberikan informasi tentang hubungan varians skor teramati dengan skor sejati kelompok individu. Tujuan mengestimasi reliabilitas lain adalah untuk menentukan seberapa besar variabilitas yang akibat adanya kesalahan pengukuran dan seberapa besar mariabilitas skor tes sebenarnya.

Im kunci: instrumen, penelitian pendidikan.

#### Pendahuluan

mentan merupakan cara mencari menaran melalui metode ilmiah, mengungkapkan dalam mengungkapkan menaran penelitian menggunakan metade ilmiah yang meliputi:

literatur, merumuskan pradugapraduga, pengumpulan data, mengolah data, menganalisis data, dan mengambil kesimpulan.

Dalam hal pengumpulan data, instrumen sangat penting dalam penelitian, karena instrumen meru-

<sup>🌫 8220</sup> Intang Sappaile, M.Pd. adalah Dosen Jurusan Pascasarjana Universitas Negeri

pakan alat ukur dan akan memberikan informasi tentang apa yang kita teliti. Informasi yang akurat diperoleh melalui instrumen yang valid dan reliabel.

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, Menurut Diaali, dkk. (2000: 9), secara umum yang dimaksud dengan instrumen adalah suatu alat yang karena memenuhi persyaratan akademis maka dapat dipergunakan sebagai alat untuk mengukur suatu objek 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, selanjutnya, 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. Sedangkan pendapat Nurkancana (1992: 141) bahwa suatu alat pengukur dapat dikatakan valid apabila alat pengukur tersebut dapat mengukur apa yang hendak diukur secara tepat dan reliabel apabila alat ukur tersebut diteskan pada kelompok yang sama pada waktu atau kesempatan yang berbeda, maka akan memberikan hasil yang relatif sama.

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

was kurang benar akan memberikan mismasi yang tidak benar, sehingga was yang diambil juga tidak Mengutip pendapat Steven Nur, 1987: 1) menyatakan pengukuran adalah pemberian maka atas objek atau kejadian sesuai dengan aturan. Dengan menitikpada alat ukurnya, maka hal ini akan dibahas instrumen Sverius (1991: 9) berpendapat mana tujuan utama dalam melakukan maliasi dalam proses belajaradalah untuk mendapatkan mformasi yang akurat mengenai pencapaian tujuan instruksimal oleh peserta didik, sehingga and dadakan tes yang diberi nama: penempatan, (2) tes formatif. Eagnostik, dan (4) tes sumatif.

## Tinjauan Kepustakaan

dapat dipergunakan sedapat dipergunakan sedapat dipergunakan sedapat dipergunakan sedapat dipergunakan sedapat dipergunakan sedapat dipergunakan data
dapat dapat dipergunakan data
dapat suatu variabel. Instrumen
dapat berbentuk tes dan juga dapat
dapat dapat dipergunakan dapat
dapat dipergunakan dapat
dapat dipergunakan dapat
dapat dipergunakan dapat
dapat dipergunakan dapat
dapat dipergunakan dapat
dapat dipergunakan dapat
dapat dipergunakan dapat
dapat dapat dipergunakan dapat
dapat dapat dapat dapat dapat dipergunakan dapat
dapat dipergunakan dapat dapat

### 2.1 Pengertian Tes

Tes adalah prosedur sistematik 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) berpendapat bahwa tes adalah suatu prosedur sistematis untuk mengamati dan mencandrakan satu atau lebih karakteristik seseorang dengan menggunakan skala numerik atau sistem kategori.

Tes juga dapat diartikan sebagai alat pengukur yang mempunyai standar objektif sehingga dapat dipergunakan untuk mengukur dan membandingkan keadaan psikis atau tingkah laku individu. Menurut Azwar (1987: 3) tes adalah prosedur yang sistematis, karena: (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.

## 2.2 Validitas

Azwar (1987: 173) menyatakan bah-wa 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) berpendapat 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 objek ukur, akan tergantung dari tingkat validitas tes yang bersangkutan. Pendapat lain, Sudjana (2004: 12) 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. Sedangkan Djaali, dkk. (2000: 70) membedakar konsep validitas tes menjadi 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 memoroporsional perilaku sampel memoroporsional perilaku sam

Untuk mengetahui apakah tes m walid atau tidak, harus dilakukan melalui penelaahan kisi kisi tes untuk memastikan bahwa soal soal tes itu member mewakili atau mencerminkan weekeruhan konten atau materi yang dikuasai secara proporonel Oleh karena itu, validitas isi tes tidak mempunyai besaran mentu yang dihitung secara and tetapi dipahami bahwa tes m sudah valid berdasarkan telaah listi tes. Oleh karena itu, validitas sebenarriya mendasarkan pada malisis logika, tidak merupakan suatu wefisien validitas yang dihitung statistika.

variabel 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 suatu instrumen 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 atas 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.

Selanjutnya, Djaali, dkk. (2000: 76) menyatakan bahwa 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 mempermasalahkan 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.

Menurut Djaali, dkk. (2000: 77) 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 biserial antara skor butir soal nomor i dengan skor total.

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

soal nomor i.

T = rata rata skor total semua responden.

standar deviasi skor total semua responden.

proporsi jawaban yang benar untuk butir soal nomor i.

proporsi jawaban yang salah untuk butir soal nomor i.

Selanjutnya, dikatakan bahwa mtuk menghitung koefisien validitas memal untuk skor butir politomi figunakan korelasi product moment dengan rumus:

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

mana:

koefisien korelasi antara skor butir soal dengan skor total.

x = jumlah kuadrat deviasi skor dari X<sub>i</sub>.

 $x_t = \text{jumlah kuadrat deviasi}$ skor dari  $X_t$ .

Nilai koefisien korelasi baik skor mair dikotomi maupun skor butir mair dikotomi maupun skor butir mair dikotomi untuk masing masing butir mair dengan nilai koefisien melasi yang ada di tabel-r pada alpha metentu misalnya  $\alpha = 0.05$ . Jika mefisien korelasi skor butir dengan skor total lebih besar dari koefisien korelasi dari r-tabel, 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 basil 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.

## 2.3 Reliabilitas

Reliabilitas berasal dari kata reliability berarti sejauh mana hasil suatu pengukuran dapat dipercaya. Nunnaly (dalam Surapranata, 2004: 89) menyatakan bahwa reliabilitas kestabilan skor yang diperoleh orang yang sama ketika diuji ulang dengan tes yang sama pada situasi yang berbeda atau dari satu pengukuran ke pengukuran lainnya. Suatu hasil pengukuran hanya 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. Sedangkan Nur (1987: 47) menyatakan bahwa reliabilitas ukuran menyangkut tentang seberapa jauh skor deviasi individu, atau skor-z, relatif konsisten apabila dilakukan pengulangan pengadministrasian dengan tes yang sama atau dengan tes yang ekivalen.

Azwar (2003 : 176) menyatakan bahwa reliabilitas merupakan salahsatu ciri atau karakter utama instrumen pengukuran yang baik. Arifin (1991: 122) berpendapat 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. Surapranata (2004: 90) menyatakan bahwa dalam pengukuran perlu diidentifikasi sumber-sumber utama yang menyebebkan terjadinya kesalahan pengukuran. Kekeliruan pengukuran sendiri menunjukkan sejauh mana inkonsistensi hasil pengukuran terjadi apabila dilakukan pengukuran ulang terhadap kelompok subyek yang sama. Nunnaly (dalam Supranata, 2004: 90) menyatakan bahwa sumber kesalahan pengukuran antara lain: (1) variasi dalam tes itu sendiri, (2) struktur sampel yang dipilih, (3) variasi di antara tes yang sedang digunakan. 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 arsebut dalam menilai apa yang dinilainya. Artinya, kapanpun alat penilaian tersebut digunakan akan memberikan hasil yang relatif sama.

Menurut Djaali, dkk. (2000: 81) miliabilitas dibedakan atas dua macam. reliabilitas konsistensi tangmean, dan reliabilitas konsistensi mbungan butir. Reliabilitas konsistensi tanggapan responden mempersoalkan apakah tanggapan responden atau objek ukur terhadap atau instrumen tersebut sudah atau konsisten. Dalam hal ini mabila suatu tes atau instrumen figunakan untuk melakukan pengterhadap objek ukur kemudian makukan pengukuran kembali andap objek ukur yang sama, wakah hasilnya masih tetap sama pengukuran sebelumnya. Jika masil pengukuran kedua menunketidakkonsistenan, maka jelas pengukuran itu tidak mencerminkan keadaan objek ukur yang gguhnya.

Untuk mengetahui apakah megapan terhadap tes atau men itu mantap, atau konsisten, dan dilakukan dengan cara menerikan tes yang sama secara kali (dua kali) kepada objek menerikan tesponden yang sama.

syarat minimal untuk mengetahui apakah tanggapan objek 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 objek 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 pengetesan sekali, 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 objek tes dalam waktu vang bersamaan. Skor dari kedua kelompok butir tes tersebut dikorelasikan untuk mendapatkan reliabilitas tes. Selanjutnya, Djaali, dkk. (2000; 81) menegaskan bahwa reliabilitas konsistensi merupakan gabungan butir berkaitan dengan kemantapan antara butir-butir suatu tes. Hal ini dapat diungkapkan dengan pertanyaan, apakah terhadap objek ukur yang sama, butir yang satu menunjukkan hasil ukur yang sama dengan butir yang lainnya? Dengan kata lain bahwa terhadap bagian objek ukur yang sama, apakah hasil ukur butir yang satu tidak kontradiksi dengan hasil ukur butir yang lain.

Jika terhadap bagian objek 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 objek ukur.

Kalau hasil pengukuran pada bagian objek ukur yang sama antara butir yang satu dengan butir yang lain saling kontradiksi atau tidak konsisten maka kita jangan menyalahkan objek ukur, melainkan alat ukur (tes) yang dipersalahkan dengan mengatakan bahwa tes tersebut tidak reliabel terhadap objek 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, dkk., 2000: 77) dengan rumus sebagai berikut.

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

k = cacah butir.

 $p_i q_i = varians skor butir.$ 

p<sub>i</sub> = proporsi jawaban yang benar untuk butir nomor i.

q<sub>i</sub> = proporsi jawaban yang salah untuk butir nomor i.

S<sub>t</sub><sup>2</sup> = varians skor total responden.

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

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

#### Kenerangan:

- = koefisien reliabilitas.
- = cacah butir.
- == varians skor butir.
- = warians skor total responden.

Interpretasi terhadap koefisien - Ibbilitas merupakan intrepretasi waif, artinya tidak ada batasan malak yang menunjukkan berapa maka koefisien minimal yang harus dapai agar suatu pengukuran dapat reliabel, Namun, memberikan masi tentang hubungan varians ser teramati dengan varians skor kelompok individu. Misalnya, meroleh koefision roliabilitas sama 3 Noefisien reliabilitas ini meat diartikan bahwa: (1) 87% skor teramati diakibatkan oleh amans skor sejati kelompok individu, (2) korelasi antara skor teramati skor sejati sama dengan 0.87 atau 0,93 (Nur, 1987: 61). Selanjutnya, Surapranata (2004: 88) menyatakan bahwa tujuan utama mengestimasi reliabilitas adalah untuk menentukan seberapa besar varia-Mas yang terjadi akibat adanya lesalahan pengukuran dan seberapa besar variabilitas skor tes sebenamya. Selanjutnya dinyatakan bahwa teori reliabilitas ditentukan berdasarkan yarian antara skor perolehan yang merupakan penjumlahan dari varian skor sebenarnya dengan varian kesalahan pengukuran. Hal ini dapat diartikan bahwa varian kesalahan pengukuran (e) merupakan selisih varian skor amatan (X) dengan varian skor sebenarnya (T) atau e = X - T.

## 3. Simpulan dan Saran

## 3.1 Simpulan

Pertama, instrumen merupakan suatu alat yang memenuhi persyaratan akademis schingga dapat dipergunakan sebagai alat untuk mengukur suatu objek ukur atau mengumpulkan data mengenai suatu variabel. Kedua, 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. Ketiga, validitas konstruk mempermasalahkan seberapa jauh butir-butir tes mampu mengukur apa yang benar-benar hendak diukur sesuai dengan definisi konsep-tual yang telah ditetapkan. Keempat, validitas empiris (validitas kriteria) yang berarti bahwa validitas ditentukan berdasarkan kriteria, baik kriteria internal maupun kriteria eksternal. Kelima, reliabilitas

mempermasalahkan sejauh mana hasil suatu pengukuran dapat dipercaya. Suatu hasil pengukuran hanya dapat dipercaya apabila dalam beberapa kali pelaksanaan pengukuran terhadap kelompok subyek sama, diperoleh hasil pengukuran yang relatif sama. Keenam, penentuan koefisien reliabilitas instrumen untuk skor butir dikotomi digunakan rumus KR-20. sedang untuk skor politomi digunakan rumus Alpha, dan Ketujuh, interpretasi terhadap koefisien reliabilitas merupakan intrepretasi 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.

#### 3.2 Saran

Pertama. Dalam penelitian pendidikan khususnya penelitian kuantitatif disarankan 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, maka instrumen yang digunakan dalam penelitian adalah instrumen yang valid berdasarkan konstruk dan valid berdasarkan kriteria. Ketiga. Dalam menentukan validitas butir dan reliabilitas instrumen, agar menggunakan rumus yang tepat dengan memperhatikan apakah butir tersebut diskor dikotomi atau diskor politomi, dan Keempat. Koefisien reliabilitas yang telah diperoleh, baik instrumen yang diskor dikotomi maupun yang diskor politomi, agar menginterpretasi koefisien reliabilitasnya.

## Pustaka Acuan

Arifin, Zainal. 1991. Evaluasi Instruksioal, Prinsi-Teknik-Prosedur, Remaja Rosdakarya: Bandung.

Azwar, Saifuddin. 1987. Sikap Manusia Teori dan Pengukurannya, Liberty: Yogyakarta.

\_\_\_\_\_\_\_, 2003. Sikap Manusia Terori dan Pengukurannya, Yogyakarta: Pustaka Pelajar.

- Djaali., Pudji Muljono, Ramly. 2000. Pengukuran Dalam Pendidikan, Jakarta: Program Pascasarjana.
- Markancana, Wayan., PPN. Sunartana. 1992. Evaluasi Hasil Belajar, Surabaya: Usaha Nasional.
- Mohamad. 1987. Teori Tes, Surabaya: IKIP Surabaya.
- Swerius, Suke. 1991. Evaluasi Hasil Belajar dan Umpan Balik, Jakarta: Gramedia Widiasarana Indonesia.
- Sudjana, Nana. 2004. Penilaian Hasil Proses Belajar Mengajar, Bandung: Remaja Rosdakarya.
- Supranata, Sumarna. 2004. Analisis, Validitas, Reliabilitas dan Interpretasi Hasil Tes, Bandung: Remaja Rosdakarya.
- Suryabrata, Sumadi. 2000. Pengembangan Alat Ukur Psikologis, Yogyakarta: Andi.

min of growing or complete