

VALIDITAS DAN RELIABILITAS

Penyusun:

Anggraini Magfirah Marti Sandjojo 1813024023

Brigita Dwi indah sari 1813024007

Faya Hirsia Aulia 1853024009

Milenia Nurfitri 1813024011

Nanda Luthfiatul Hasanah 1853024005

Triana Setia Lestari 1853024007

Ummul shafira 1813024039

Mata Kuliah : Evaluasi Pembelajaran Biologi



PENDIDIKAN BIOLOGI

FAKULTAS KEGURUAN DAN ILMU PENDIDIKAN

UNIVERSITAS LAMPUNG

BANDAR LAMPUNG

2020

PEMBAHASAN

1. 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.

Macam- macam Validitas

a. Validitas isi (*content validity*)

Validitas isi suatu tes mempermasalahkan seberapa jauh suatu tes mengukur tingkatpenguasaan 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 (Matondang, 2009: 90).

b. Validitas konstruk (*construct validity*)

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, locus kontrol, gaya kepemimpinan, motivasi berprestasi, dan lain-lain, maupun yang sifatnya performansi maksimum seperti instrumen untuk mengukur bakat (tes bakat), inteligensi (kecerdasan intelektual), kecerdasan, emosional dan lain-lain (Matondang, 2009: 90).

c. Validitas empiris atau validitas kriteria

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 (Matondang, 2009: 91).

Cara menentukan validitas

Untuk menguji validitas setiap butir soal maka skor-skor yang ada pada butir yang dimaksud dikorelasikan dengan skor totalnya. Skor tiap butir soal dinyatakan skor X dan skor total dinyatakan sebagai skor Y, dengan diperolehnya indeks validitas setiap butir soal, dapat diketahui butir-butir soal manakah yang memenuhi syarat dilihat dari indeks validitasnya (Arikunto, 1999:78).

$$r_{xy} = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{\left(n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right) \left(n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2 \right)}}$$

Keterangan :

r_{xy} adalah koefisien korelasi antara variable X dan variable Y

x_i adalah nilai data ke-i untuk kelompok variable X

y_i adalah nilai data ke-i untuk kelompok variable Y

n adalah banyak data

Tentukan kategori dari validitas instrument yang mengacu pada pengklasifikasian validitas yang dikemukakan oleh Guilford (1956, h.145) adalah sebagai berikut:

koefisien korelasi (r_{xy})	Tingkat Validitas
0,80 - 1,00	Sangat tinggi (sangat baik)
0,60 - 0,80	Tinggi (baik)
0,40 - 0,60	Validitas sedang (cukup)
0,20 - 0,40	Validitas rendah (kurang)
0,00 - 0,20	Validitas sangat rendah (jelek)
0,00	Tidak valid

2. Reliabilitas

Kata reliabilitas dalam bahasa Indonesia di ambil dari *reliability* dalam bahasa inggris, berasal dari kata, reliable yang artinya dapat dipercaya. “reliabilitas” merupakan kata benda, sedangkan “reliable” merupakan kata sifat atau keadaan. Reliabilitas merupakan penerjemahan dari kata reliability yang mempunyai asal kata rely dan ability. Pengukuran yang memiliki reliabilitas tinggi disebut sebagai pengukuran yang reliabel (*reliable*). Walaupun reliabilitas mempunyai berbagai arti seperti kepercayaan, keterandalan, keajegan, kestabilan dan konsistensi, namun ide pokok yang terkandung dalam konsep reliabilitas adalah sejauh mana hasil pengukuran dapat dipercaya.

Menurut Azwar (2011:4) Reliabilitas merupakan penerjemah dari kata reliability yang mempunyai asal kata rely dan ability. Pengukuran yang memiliki reliabilitas tinggi disebut sebagai pengukuran yang reliabel. Konsep reliabilitas adalah sejauh mana hasil suatu pengukuran dapat dipercaya seperti pendapat dari Masri Singarimbun, realibilitas adalah indeks yang menunjukkan sejauh mana suatu alat ukur dapat dipercaya atau dapat diandalkan. Bila suatu alat pengukur dipakai dua kali untuk mengukur gejala yang sama dan hasil pengukuran yang diperoleh relative konsisten, maka alat pengukur tersebut reliable. Dengan kata lain, realibitas menunjukkan konsistensi suatu alat pengukur di dalam pengukur gejala yang sama.

Dapat disimpulkan reliabilitas tes merupakan suatu alat ukur yang digunakan untuk mengetahui konsistensi pengukuran tes yang hasilnya menunjukan keajegan (konsisten). Seorang dikatakan dapat di percaya apabila orang tersebut berbicara konsistensi, tidak berubah- ubah pembicaraannya dari waktu ke waktu. Dalam sebuah tes pentingnya diamati keajegan dan kepastian tes tersebut dilihat dari hasil tes yang didapat.

Jenis- Jenis Reliabilitas

Walizer (1987) menyebutkan bahwa ada dua cara umum untuk mengukur reliabilitas, yaitu;

a. Relibilitas stabilitas

Menyangkut usaha memperoleh nilai yang sama atau serupa untuk setiap orang atau setiap unit yang diukur setiap saat anda mengukurnya. Reliabilitas ini menyangkut penggunaan indikator yang sama, definisi operasional, dan prosedur pengumpulan data setiap saat, dan mengukurnya pada waktu yang berbeda. Untuk dapat memperoleh reliabilitas stabilitas setiap kali unit diukur skornya haruslah sama atau hampir sama.

b. Reliabilitas ekivalen

Menyangkut usaha memperoleh nilai relatif yang sama dengan jenis ukuran yang berbeda pada waktu yang sama. Definisi konseptual yang dipakai sama tetapi dengan satu atau lebih indikator yang berbeda, batasan-batasan operasional, peralatan pengumpulan data, atau pengamat-pengamat

Menguji reliabilitas dengan menggunakan ukuran ekivalen pada waktu yang sama bisa menempuh beberapa bentuk. Bentuk yang paling umum disebut teknik belah- tengah. Cara ini seringkali dipakai dalam survai. Apabila satu rangkaian pertanyaan yang mengukur satu variable dimasukkan dalam kuesioner, maka pertanyaan- pertanyaan tersebut dibagi dua bagian persis lewat cara tertentu. (Pengacakan atau pengubahan sering digunakan untuk

teknik belah tengah ini.) Hasil masing-masing bagian pertanyaan diringkas ke dalam skor, lalu skor masing-masing bagian tersebut dibandingkan. Apabila dalam skor kemudian skor masing-masing bagian tersebut dibandingkan. Apabila kedua skor itu relatif sama, dicapailah reliabilitas belah tengah.

Reliabilitas ekivalen dapat juga diukur dengan menggunakan teknik pengukuan yang berbeda. Kecemasan misalnya, telah diukur dengan laporan pulsa. Skor- skor relatif dari satu indikator macam ini haruslah sesuai dengan skor yang lain. Jadi bila seorang subyek nampak cemas pada "ukuran gelisah" orang tersebut haruslah menunjukkan tingkatan kecermatan relatif yang sama bila tekanan darahnya yang diukur.

Faktor-Faktor Yang Mempengaruhi Koefisien Reliabilitas

Ada beberapa faktor yang mempengaruhi koefisien reliabilitas dari instrumen yang berupa tes, diantaranya:

- a) Panjang Tes. Pada umumnya semakin panjang tes semakin tinggi reliabilitasnya. Hal ini disebabkan tes yang cacah butirnya banyak akan memuat cukup banyak tingkah laku yang diukur.
- b) Penyebaran Skor. Koefisien reliabilitas dipengaruhi oleh penyebaran skor. Makin lebar penyebaran skor makin tinggi estimasi koefisien reliabilitasnya. Hal ini disebabkan koefisien reliabilitas akan semakin tinggi apabila individu- individu cenderung tetap pada kedudukan terhadap kelompoknya.
- c) Tingkat Kesukaran Tes. Tes yang terlalu sukar atau terlalu mudah cenderung menurunkan koefisien reliabilitas. Hal ini disebabkan tes yang terlalu sukar atau terlalu mudah menghasilkan sebaran yang terbatas dan terkumpul diujung bawah atau diujung atas.
- d) Objektivitas suatu tes menunjukkan seberapa jauh dua orang yang mempunyai kemampuan yang sama mendapatkan skor yang sama.

Pengujian Realibilitas

Pendekatan dan Teknik pengujian realibilitas antara lain :

a) Teknik Paralel (Paralel Form atau Alternate Form)

Teknik paralel disebut juga teknik "double test double trial". Sejak awal peneliti harus sudah menyusun dua perangkat instrument yang paralel (ekuivalen), yaitu dua buah instrument yang disusun berdasarkan satu buah kisi- kisi. Setiap butir soal dari instrument yang satu selalu harus dapat dicarikan pasangannya dari instrumen kedua. Kedua instrumen tersebut

diuji cobakan semua. Sesudah kedua uji coba terlaksana, maka hasil instrumen tersebut dihitung korelasinya dengan menggunakan rumus product moment (korelasi Pearson).

b) Teknik Ulang (Test Re-test)

Teknik ini disebut juga teknik "*single test double trial*". Menggunakan sebuah instrument, namun dites dua kali. Hasil atau skor pertama dan kedua kemudian dikorelasikan untuk mengetahui besarnya indeks reliabilitas. Teknik perhitungan yang digunakan sama dengan yang digunakan pada teknik pertama yaitu rumus korelasi Pearson.

Menurut Saifuddin Azwar, realibilitas tes-retest adalah seberapa besar derajat skor tes konsisten dari waktu ke waktu. Realibilitas diukur dengan menentukan hubungan antara skor hasil penyajian tes yang sama kepada kelompok yang sama, pada waktu yang berbeda.

Metode pengujian reliabilitas stabilitas yang paling umum dipakai adalah metode pengujian tes-kembali (*test-retest*). Metode test-retest menggunakan ukuran atau "*test*" yang sama untuk variable tertentu pada satu saat pengukuran yang diulang lagi pada saat yang lain. Cara lain untuk menunjukkan reliabilitas stabilitas, bila kita menggunakan survai, adalah memasukkan pertanyaan yang sama di dua bagian yang berbeda dari kuesioner atau wawancara. Misalnya the *Minnesota Multiphasic Personality Inventory* (MPPI) mengecek reliabilitas test-retest dalam satu kuesionernya dengan mengulang pertanyaan tertentu di bagian-bagian yang berbeda dari kuesioner yang panjang.

Kesulitan terbesar untuk menunjukkan reliabilitas stabilitas adalah membuat asumsi bahwa sifat/ variable yang akan diukur memang benar-benar bersifat stabil sepanjang waktu. Karena kemungkinan besar tidak ada ukuran yang andal dan sah yang tersedia. Satu-satunya faktor yang dapat membuat asumsi-asumsi ini adalah pengalaman, teori atau putus dan terbaik. Dalam setiap kejadian, asumsi ini selalu ditantang dan sulit rasanya mempertahankan asumsi tersebut atas dasar pijakan yang obyektif.

c) Teknik Belah Dua

Penentuan reabilitas tes hasil belajar bentuk obyektif dengan menggunakan formula Spearman-Brown dikenal dengan istilah: teknik belah dua (*split half technique*). Disebut "belah dua", sebab dalam penentuan reabilitas tes, penganalisisannya dilakukan dengan jalan membelah dua butir-butir soal tes menjadi dua bagian yang sama, sehingga masing-masing testee memiliki dua macam skor. Salah satu skor merupakan bagian pertama atau belahan pertama dari tes, sedangkan skor yang satunya lagi merupakan bagian kedua atau belahan kedua dari tes hasil belajar bentuk obyektif tersebut.

Dengan demikian, penerapan formula Spearman-Brown akan menghasilkan dua buah distribusi skor belahan pertama drngan distribusi skor belahan keduannya itu dipandang sebagai reabilitas bagian butir-butir soal tes hasil belajar bentuk obyektif tersebut; sedangkan untuk

mengetahui reabilitas tes secara keseluruhan Spearman-Brown menciptakan formula sebagai berikut dimana :

$$R_{tt} = \frac{2 \cdot r_{hh}}{1 + r_{hh}}$$

r_{tt} = Koefisien reabilitas tes secara total (tt = total tes)

r_{hh} = Koefisien korelasi product moment antara separoh (bagian pertama) tes, dengan separoh (bagian kedua) dari tes tersebut

(hh = half-half)

1&2 = Bilangan Konstantan

Rumus lain yang sejenis dengan rumus diatas adalah ;

Dimana :

r_{11} = Koefisien reabilitas tes secara keseluruhan

r = Koefisien korelasi product moment antara separoh (1/2) tes (belahan) dengan separoh (1/2) tes (belahan) dari tes tersebut.

1&2 = Bilangan Konstan

Untuk mengetahui besarnya r_{hh} atau r dapat digunakan salah satu diantara rumus berikut :

Realibilitas Seluruh Tes = $1 + \frac{2 \times \text{reliabilitas separuh tes}}{1 + \text{reliabilitas separuh tes}}$

Atau $r = \frac{2 \times r}{1 + r}$

Misalnya, berdasarkan penghitungan koefisien korelasi separuh soal didapatkan r sebesar 0,713 , maka tingkat kepercayaan seluruh tes adalah:

Realibilitas Seluruh Tes = $2 \times 0,713 = 0,832$

$1 + 0.713$

Koefisien reliabilitas seluruh tes yang diperoleh tersebut termasuk kategori tinggi, maka tes yang diuji itu dinyatakan memiliki tingkat reliabilitas yang tinggi pula. Artinya hasil pengukuran yang dilakukan dengan alat tes yang bersangkutan dinyatakan konsisten. Dalam penerapan formula Spearman-Brown tersebut diatas, Spearman-Brown mempersembahkan dua buah model, yaitu Model Gasal- Genap dan Model Kiri-Kanan.

a) Model Gasal-Genap

Pada model ini skor-skor yang dimiliki oleh testee untuk butir item yang bernomor gasal (misalnya item nomor 1, 3, 5, 7, 9, 11 dan seterusnya) dianggap sebagai separoh bagian pertama dari tes, sedangkan skor-skor yang dimiliki testee untuk butir-butir item yang bernomor genap (misalnya item nomor 2, 4, 6, 8, 10, 12 dan seterusnya) dianggap sebagai separoh bagian kedua dari tes yang bersangkutan.

- Pendekatan Single Test-Single Trial dengan Menggunakan Formula Spearman-Brown Model Gasal Genap. Langkah-langkah yang perlu ditempuh dalam penentuan reliabilitas tes dengan pendekatan single-test dimana digunakan formula Spearman-Brown Model Genap adalah sebagai berikut :
- Menjumlah skor-skor dari butir-butir item yang bernomor gasal yang dimiliki oleh masing-masing, individu testee.
- Menjumlahkan skor-skor dari butir-butir item yang bernomor genap yang dimiliki oleh masing-masing individu testee.
- Mencari (menghitung) koefisien korelasi “r” product moment ($r_{xy} = r_{hh} = r$).

Dalam hal ini jumlah skor-skor dari butir-butir item yang bernomor gasal kita anggap sebagai variable X, sedangkan jumlah skor-skor dari butir-butir item yang bernomor genap kita anggap sebagai variable Y, dengan menggunakan rumus Mencari (menghitung) koefisien reliabilitas tes ($r_{11} = r_{tt}$) dengan menggunakan rumus Memberikan interpretasi terhadap r_{11} .

b) Model Belahan Kiri- Kanan

Pada model belahan kiri-kanan, jumlah butir-butir item yang ada dalam tes, dibelah menjadi dua bagian yang sama besar. Pendekatan Single Test – Single Thrial dengan menggunakan Formula Spearman-Brown Model Belahan Kiri dan Kanan. Langkah-langkah yang perlu ditempuh dalam penentuan reliabilitas tes dengan menggunakan pendekatan Single Test – Single Thrial di mana digunakan formula Sperman-Brown model belahan kiri dan kanan adalah sebagai berikut :

- Menjumlahkan skor-skor dari butir-butir item yang terletak di separoh bagian kiri tes, yang dimiliki oleh masing-masing individu siswa yaitu : butir-butir item dengan nomor 1,2,3, 4,5, 6, 7, 8, 9, 10, 11 dan 12.
- Menjumlahkan skor-skor dari butir-butir item yang terletak di separoh bagian kanan tes, yang dimiliki oleh masing-masing individu siswa yaitu : butir-butir item dengan nomor 13,14,15,16,17,18,19,20,21,22,23 dan 24.
- Mencari (menghitung) angka indeks korelasi “r” product moment, antara variabel X (separoh belahan kiri) dengan variabel Y (separoh belahan kanan) yaitu r_{xy} atau r_{hh} atau r .
- Mencari (menghitung) koefisien reliabilitas tes (r_{11} atau r_{tt}) dengan menggunakan rumus Telah diketahui ; dengan demikian r_{11} .
- Memberikan interpretasi terhadap r_{11}

Penentuan reabilitas tes dengan menggunakan formula Spearman-Brown memiliki beberapa kelemahan yaitu ;

- a) Formula Spearman-Brown menghendaki agar bekahan yang dicarikorelasinya yaitu belahan gasal-genap dan belahan kiri-kanan haruslah sebanding.
- b) Penerapan formula Spearman-Brown juga menuntut persyaratan, agar jumlah butir-butir item yang akan diuji reabilitasnya haruslah merupakan bilangan genap; jadi seandainya jumlah butir-butir item itu berupa bilangan gasal, maka formula ini tidak mungkin untuk diterapkan.

Dengan dua buah model perhitungan tersebut (model gasal-genap dan model kiri-kanan), dapat terjadi bahwa koefisien reabilitas tes menunjukkan bilangan yang tidak sama, sehingga dapat terjadi bahwa dengan menggunakan model gasal genap tes dinyatakan reliable (karena r_{11} atau r_{tt} menunjukkan angka 0,70 atau lebih), tetapi dengan menggunakan model kiri-kanan ternyata tes dinyatakan un-reliabel (karena besarnya r_{11} dibawah 0,70).

DAFTAR PUSTAKA

- Arikunto, Suharsimi. 1999. *Prosedur Penelitian: Suatu Pendekatan Praktek. ed. Rev. IV.* Yogyakarta: Rineka Cipta.
- Azwar, S. 2011. *Metode Penelitian.* Yogyakarta: Pustaka Pelajar.
- Azwar. 1987. *Metodologi Penelitian.* Jakarta: Binarupa Aksara.
- Matondang, Z. 2009. Validitas Dan Reliabilitas Suatu Instrumen Penelitian. *Jurnal Tabularasa PPS UNIMED.* 6 (1): 90-91.
- Sudjana, N. 2004. *Penilaian Hasil Proses Belajar Mengajar.* Bandung: Remaja Rosdakarya.
- Suryabrata, S. 2000. *Pengembangan Alat Ukur Psikologis.* Yogyakarta: Andi.
- Walizer, H.Michael dan Wiener, L. Paul. (1987). *Metode dan Anlisis Penelitian : Mencari Hubungan,* Jilid Dua, Alih Bahasa Arief Sukadi Sadiman. Jakarta: Erlangga.