

Validitas dan Reliabilitas Instrumen Penilaian Pada Mata Pelajaran Bahasa Arab

Syaifudin

Institut Agama Islam Sultan Muhammad Syafiuddin Sambas

Email: Usyaf96@yahoo.com

ABSTRACT

The discussion regarding assessment is certainly inseparable from the teaching and learning process for one semester or half semester. Good learning outcomes are also determined by good assessments in the objective sense according to reality, therefore in the assessment absolute validity is needed. This article is studied using library research methods to reveal valid and reliable test instruments. Based on the putaka study carried out, there are at least four characteristics or characteristics that must be possessed in a learning test, so that the test can be declared a good test, namely: (1) valid (sahih = 2); (2) reliable (tsabit = 3); (3) objective (maudu'iy = موضوعي); and (4) practical ('amaliy = عملي). However, the discussion of this paper focuses on two things, namely (1) valid (sahih = صحيح) and (2) reliable (tsabit = ثابت).

Keywords: validity, reliability, Arabic language test instrument

ABSTRAK

Pembahasan mengenai penilaian tentu tidak terpisahkan lagi dari proses belajar mengajar selama satu semester atau setengah semester. Hasil belajar yang baik juga ditentukan dengan penilaian yang baik pula dalam arti objektif sesuai kenyataan, oleh karena itu dalam penilaian diperlukan kesahihan yang mutlak. Artikel ini dikaji menggunakan metode penelitian pustaka (*library research*) guna mengungkap instrument tes yang valid dan reliable. Berdasarkan kajian putaka yang dilakukan, setidaknya ada empat ciri atau karakteristik yang harus dimiliki dalam tes belajar, sehingga tes tersebut dapat dinyatakan sebagai tes yang baik, yaitu: (1) valid (sahih= صحيح); (2) reliabel (tsabit= ثابت); (3) obyektif (maudu'iy= موضوعي); dan (4) praktis ('amaliy= عملي). Namun pembahasan makalah ini fokus pada dua hal yaitu (1) valid (sahih= صحيح) dan (2) reliabel (tsabit= ثابت).

Kata kunci: validitas, reliabilitas, instrument tes bahasa arab

PENDAHULUAN

Pada era revolusi industri 4.0 ini, suatu penilaian adalah hal mutlak yang harus dilakukan oleh seorang guru. Selain untuk memantau perkembangan siswa atau peserta didiknya, juga untuk mengevaluasi kinerja atau pengajaran yang dilakukan. Tentu suatu penilaian yang diharapkan ialah penilaian yang dilakukan secara objektif, bukan subjektif. Atau bahkan menilai dengan cara “suka-suka” atau “kira-kira”. Ketika dikaji firman Allah dalam al-Qur'an Surah al-Hasyr ayat 18 yang menjelaskan bahwa Allah memerintahkan kepada orang yang beriman untuk mengintrospeksi diri atau mengevaluasi diri, menilai-nilai diri

sebagai bentuk takwa kepada Allah, yang menjadi sorotan utama ialah pada kata “*wal tanzur nafsum ma qaddamad li ghad*” yang ditafsirkan oleh Ibnu Katsir sebagai upaya seorang manusia untuk mengintrospeksi dan mempersiapkan diri menjadi lebih baik di masa yang akan datang. Selain itu Umar bin Khattab juga berkata “*hāsibū anfusakum qobla an tuhāsabū*” yang dapat diinterpretasikan bahwa sebagai seorang guru mestilah selalu mawas diri mengevaluasi cara mengajarnya setiap saat sehingga hasil belajar siswa semakin meningkat.

METODE PENELITIAN

Metode yang digunakan dalam kajian ini menggunakan metode atau pendekatan kepustakaan (*library research*). Dewasa ini, studi pustaka atau kepustakaan dapat dimaknai sebagai suatu rangkaian kegiatan yang berkenaan dengan metode pengumpulan data pustaka, membaca dan mencatat serta mengolah bahan penelitian (Zed, 2004). Terdapat empat ciri khas dalam penelitian studi pustaka diantaranya :Pertama, bahwa penulis atau peneliti berhadapan langsung dengan data berupa teks (*nash*) atau angka, bukan dengan terjun langsung ke lapangan. Kedua, data pustaka bersifat instan “siap pakai” maksudnya peneliti langsung mengkaji sumber data penelitian yang memang sudah ada di perpustakaan, baik itu buku atau pun jurnal. Ketiga, bahwa data pustaka umumnya adalah sumbersekunder, artinya data yang dicari oleh peneliti berasal dari data yang sudah diteliti sebelumnya baik itu buku maupun jurnal. Dengan kata lain peneliti memperoleh bahan atau data dari peneliti atau penulis terdahulu dan bukan asli dari hasil penelitian pertama di lapangan. Keempat, bahwa untuk mendapatkan data pustaka tidak dibatasi oleh *setting location* dan *time schedule* penelitian (Zed, 2004).

Berdasarkan penjelasan tersebut, maka pengumpulan data dalam penelitian ini, dilakukan dengan menelaah serta mengeksplorasi beberapa jurnal, buku, dan dokumen-dokumen (baik yang berbentuk cetak maupun elektronik) serta sumber-sumber data dan informasi lainnya yang dianggap relevan dengan penelitian atau kajian yang diteliti.

PEMBAHASAN

A. Validitas

1. Pengertian Validitas

Kata valid berasal dari kata *validity* yang memiliki arti ketepatan, kecermatan, sedangkan dalam bahasa Arab valid disebut dengan (*shahih*= صحيح) yang artinya: tepat, benar, shahih, absah; jadi kata validitas dapat diartikan dengan ketepatan, kebenaran, keshahihan atau keabsahan. Apabila kata valid dikaitkan dengan fungsi tes sebagai alat pengukur, maka sebuah tes dikatakan valid

apabila tes tersebut dengan secara tepat, secara benar, secara shahih, atau secara absah dapat mengukur apa yang seharusnya diukur. Dengan kata lain, sebuah tes dikatakan telah memiliki validitas apabila tes tersebut dengan secara tepat, benar, shahih atau absah telah dapat mengungkap atau mengukur hasil belajar siswa setelah menempuh proses belajar dalam jangka waktu tertentu (Sudijono, 1996).

Menurut Azwar (1986) validitas berasal dari kata *validity* yang mempunyai arti sejauh mana ketepatan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya. Menurut Arikunto (1999) validitas adalah suatu ukuran yang menunjukkan tingkat kesahihan suatu tes. Menurut Nursalam (2003) validitas adalah suatu ukuran yang menunjukkan tingkat kevalidan atau kesahihan suatu instrumen.

Valid menurut Gronlund (1985) dapat diartikan sebagai ketepatan interpretasi yang dihasilkan dari skor atau instrumen evaluasi (butir soal). Menurut Gay (1983) dan Johnson (2002) menerangkan bahwa suatu instrumen evaluasi (butir soal) dikatakan valid apabila instrumen yang digunakan dapat mengukur apa yang hendak diukur (Sukardi, 2010). Pada buku "*encyclopedia of education evaluation*" yang ditulis oleh Scarvia b Anderson dan kawan-kawan disebutkan: "a test is valid if it measures what it purpose to measure" artinya bahwa sebuah tes dikatakan valid apabila tes tersebut mengukur apa yang hendak diukur (Arikunto, 2012). validitas suatu butir instrumen penilaian tercermin dari sejauhmana hasil pengukuran sesuai fakta atau keadaan sebenarnya dari apa yang diukur (Supardi, 2016).

Contoh: untuk mengukur besarnya partisipasi siswa dalam proses belajar-mengajar akan menjadi tidak valid jika diukur melalui nilai yang diperoleh pada waktu ulangan, tetapi dilihat melalui kehadiran, terpusatnya siswa pada pelajaran, ketepatan menjawab pertanyaan-pertanyaan yang diajukan guru dalam arti relevan pada permasalahannya. Nilai yang diperoleh pada waktu ulangan, bukan menggambarkan partisipasi, tetapi menggambarkan prestasi belajar (Arikunto, 2012). Contoh lainnya yaitu: dalam mata pelajaran pendidikan jasmani, olahraga dan kesehatan, misalnya indikator "mempraktikkan gerak dasar jalan", maka penilaian yang valid adalah apabila menggunakan penilaian unjuk kerja (praktik), jika menggunakan tes tulis (soal tertulis) maka penilaian tidak valid (Uno.B.hamzah. Koni, 2014).

Berdasarkan beberapa pendapat tentang pengertian validitas di atas, maka dapat diambil kesimpulan bahwa validitas adalah suatu standar ukuran yang menunjukkan ketepatan dan kesahihan suatu instrumen. Hakikatnya validitas bukan menekankan pada tes itu sendiri melainkan pada hasil pengetesan atau skornya (Uno.B.hamzah. Koni, 2014).

Dalam kehidupan sehari-hari contoh real yang dapat menjelaskan tentang validitas adalah cara kerja alat pengukur berupa barometer dan termometer, barometer akan menunjukkan kevalidan suatu tekanan udara sedangkan termometer akan menunjukkan kevalidan suatu tekanan suhu.

Contoh validitas:

Tujuan pembelajaran	Instrumen soal	Valid	Tidak valid
Mampu mengidentifikasi bunyi ungkapan-ungkapan yang diberikan, melalui pelafalan yang diperdengarkan	-Lafalkan kembali ungkapan yang kalian dengarkan! -Sebutkan ungkapan apa saja yang kalian dengarkan! - tulislah kosa kata yang tidak kamu ketahui! - isilah kalimat yang kosong !	✓ ✓	× ×

2. Jenis-Jenis validitas

Encyclopedia of Education Evaluation sebuah buku yang ditulis oleh oleh Scarvia B. Anderson dan kawan-kawan. Bahwa secara garis besar validitas terbagi menjadi dua yaitu validitas logis dan validitas empiris, validitas logis terdiri dari: validitas isi, validitas muka dan validitas konstruk, sedangkan validitas empiris terdiri dari: validitas banding dan validitas ramal.

a. Validitas logis (rasional)

Validitas rasional adalah validitas yang diperoleh atas dasar hasil pemikiran, validitas yang diperoleh dengan berpikir secara logis (Sudijono, 2011).

1) Validitas Tampak/muka (*Face Validity*)

Validitas tampak adalah sejauhmana penentuan validitas tersebut berdasarkan apa yang nampak. Ini adalah penilaian yang sangat minimal karena melihat dari bungkusnya. Contoh: Seseorang mengaku polisi. Kita melihat orang

tersebut memakai baju seragam polisi, maka dari jenis validitas tampak bahwa pengakuan orang tersebut valid.

- 2) Validitas Isi (*Content Validity*) Validitas isi dari suatu tes hasil belajar adalah validitas yang diperoleh setelah dilakukan penganalisisan, penelusuran atau pengujian terhadap isi yang terkandung dalam tes hasil belajar tersebut. Validitas isi sering pula dinamakan validitas kurikulum, karena yang ditelaah adalah isi materi yang terdapat dalam kurikulum sehingga isi tes dapat dikatakan valid apabila telah dapat mewakili secara representatif terhadap keseluruhan materi atau bahkan pelajaran yang seharusnya diteskan (diujikan) (Sudijono, 2011).

Contoh (1) soal matematika hanya mengukur kemampuan menghitung bukan mengukur kemampuan baerbahasa

Contoh (2) sebuah tes yang dirancang untuk mengukur kemampuan bermain bola basket dalam mata pelajaran penjaskes, tentu yang diukur adalah kemampuan berlari, membawa bola, menembak bola dan mendreble bola.

Contoh (3) sebuah tes yang dirancang untuk mengukur *mahārah kalām* dalam mata pelajaran bahasa Arab, tentu yang diukur adalah kemampuan melafazkan kalimat, intonasi, dan kefasihan.

- 3) Validitas konstruksi (*Construct Validity*) Validitas konstruksi dapat diartikan sebagai validitas yang ditilik dari segi susunan, kerangka atau rekaannya. Adapun secara terminologis suatu tes dikatakan memiliki validitas konstruksi apabila soal-soalnya mengukur setiap aspek berpikir seperti diuraikan dalam standar kompetensi, kompetensi dasar, indikator maupun tujuan pembelajaran dalam kurikulum. Soal yang dikembangkan berkaitan dengan aspek ingatan, pemahaman, aplikasi, analisis, sintesis dan evaluasi. Contoh: tujuan pembelajaran "*siswa mampu membandingkan antara efek biologis dan efek psikologis*", maka intruksi atau perintah pada butir soal tes tersebut ialah: "Jelaskan perbandingan antara efek biologis dan efek psikologis !"

b. Validitas Empirik

Validitas empirik adalah ketepatan mengukur yang didasarkan pada hasil analisis yang bersifat empirik. Dengan kata lain, validitas empirik adalah validitas yang bersumber pada atau diperoleh atas dasar pengamatan di lapangan. untuk dapat

menentukan apakah tes hasil belajar sudah memiliki validitas empirik atau belum, dapat dilakukan penelusuran dari dua segi, yaitu dari segi daya ketepatan ramalnya dan daya ketepatan bandingannya.

- 1) Validitas ramalan (*Predictive validity*) Validitas ramalan adalah suatu kondisi yang menunjukkan seberapa jauhkah sebuah tes telah dapat dengan secara tepat menunjukkan kemampuannya untuk meramalkan apa yang bakal terjadi pada masa mendatang.

Misalnya tes masuk perguruan tinggi adalah sebuah tes yang diperkirakan mampu memprediksi keberhasilan calon mahasiswa dalam mengikuti kuliah di masa yang akan datang. Calon mahasiswa yang tersaring berdasarkan hasil tes yang diharapkan mencerminkan tinggi rendahnya kemampuan mengikuti kuliah. Jika nilainya tinggi maka diprediksi bahwa calon mahasiswa tersebut mampu berhasil kelak. Begitu sebaliknya jika tidak lulus tes karena nilai rendah maka diprediksi bahwa calon mahasiswa tersebut tidak mampu mengikuti perkuliahan di masa akan datang.

- 2) Validitas bandingan (*Concurrent Validity*) sering dikenal dengan empiris artinya ketepatan mengukur yang diperoleh dari pengalaman lapangan, ada juga yang menyebutnya validitas ada sekarang, sehingga terdapat dua pengertian: disebut validitas bandingan apabila tes tersebut dalam kurun waktu yang sama dengan secara tepat mampu menunjukkan adanya hubungan yang searah, antara tes pertama dengan tes berikutnya. Disebut validitas ada sekarang karena empiris selalu berkaitan dengan pengalaman, sedangkan pengalaman selalu mengenai hal yang telah lampau berarti pengalaman tersebut sudah ada sekarang. Contoh: Misalnya seorang guru ingin mengetahui apakah tes sumatif yang disusun sudah valid atau belum. Untuk ini diperlukan sebuah kriterium masa lalu yang sekarang datanya dimiliki. Misalnya nilai ulangan harian atau nilai ulangan sumatif yang lalu.

3. Faktor faktor yang mempengaruhi hasil tes evaluasi tidak valid

Secara garis besar faktor yang mempengaruhi tes evaluasi tidak valid ada empat (Sukardi, 2010):

- a) Faktor proses pembelajaran
 - 1) Cara guru menyampaikan materi (Kusaeri, 2014)
 - 2) Siswa kurang dibimbing dan diajar menyelesaikan contoh soal soal yang merupakan kisi-kisi yang akan diujikan

- b) Faktor yang berasal dari dalam tes
 - 1) Arahan tes yang disusun dengan makna tidak jelas sehingga dapat mengurangi validitas tes.
 - 2) Kata-kata yang digunakan dalam struktur instrumen soal terlalu sulit
 - 3) Tingkat kesulitan item soal tidak tepat dengan materi pembelajaran yang diterima siswa
 - 4) Jumlah item soal terlalu sedikit sehingga tidak mewakili materi pembelajaran
 - 5) Jawaban masing-masing soal bisa diprediksi siswa.
- c) Faktor yang berasal dari proses administrasi dan pensekoran
 - 1) Waktu pengerjaan tidak cukup sehingga siswa tergesa-gesa dalam menjawab
 - 2) Adanya kecurangan dalam pelaksanaan tes sehingga sulit untuk memberikan pensekoran antara siswa yang benar-benar belajar dengan siswa yang mencontek.
 - 3) Pemberian petunjuk pengawas tidak merata diberikan kepada siswa, sehingga beberapa siswa tidak dapat mengikuti arahan yang diberikan sebelum tes
 - 4) Teknik pensekoran yang tidak konsisten, misalnya pada soal essay.
- d) Faktor yang berasal dari jawaban siswa
Faktor ini berasal dari psikologis siswa saat menghadapi pengawas yang tegas, sehingga ada beberapa siswa yang takut serta tegang, hal ini berdampak pada jawaban siswa, sehingga nantinya juga akan berdampak pada hasil tes.

B. Pengertian Reliabilitas

1. Pengertian reliabilitas

Kata reliabilitas dalam bahasa Indonesia yang digunakan saat ini, sebenarnya diambil dari kata *reliability* dalam bahasa Inggris dan berasal dari kata *reliable* yang artinya dapat dipercaya, keajegan, konsisten, keandalan, kestabilan. Sedangkan dalam bahasa Arab disebut (*tsabit*= ثابت) yang artinya pasti, tetap, stabil, mantap, mapan, tidak berubah, tidak bergerak, tetap teguh (al maany, 2017). Apabila istilah tersebut dikaitkan dengan fungsi tes sebagai alat pengukur mengenai keberhasilan belajar peserta didik, maka sebuah tes hasil belajar dapat dinyatakan reliabel apabila hasil-hasil pengukuran yang dilakukan dengan menggunakan tes tersebut secara berulang kali terhadap subyek yang sama, senantiasa menunjukkan hasil yang tetap sama atau sifatnya ajeg dan stabil

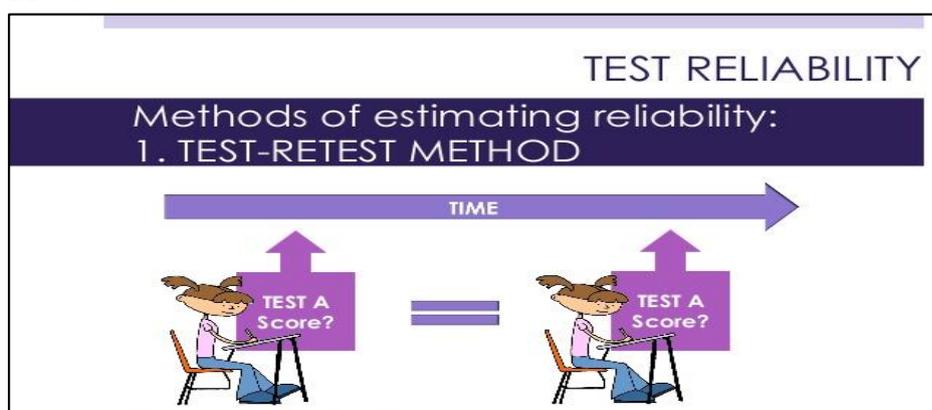
(Sudijono, 1996). Jika dikaitkan dengan ujian sekolah atau nasional maka suatu ujian dapat dikatakan reliabilitas apabila skor-skor atau nilai-nilai yang diperoleh para peserta ujian untuk pekerjaan ujiannya adalah stabil, kapan saja, dimana saja dan oleh siapa saja ujian itu dilaksanakan, diperiksa dan dinilai.

Menurut Sugiono (2005) Reliabilitas adalah serangkaian pengukuran atau serangkaian alat ukur yang memiliki konsistensi bila pengukuran yang dilakukan dengan alat ukur itu dilakukan secara berulang. Sedangkan Sukadji (2000) mengatakan bahwa reliabilitas suatu tes adalah seberapa besar derajat tes mengukur secara konsisten sasaran yang diukur.

Melalui beberapa pendapat para tokoh dapat dipahami bahwa prinsip reliabilitas dapat berupa pertanyaan “seberapa jauhkah pengukuran yang dilakukan secara berulang-ulang terhadap subyek atau kelompok subyek yang sama, memberikan hasil-hasil yang relatif tidak mengalami perubahan?” jika hasil-hasil yang diperoleh selalu sama atau paling tidak mendekati sama maka dapat dikatakan bahwa alat pengkuran berupa tes tersebut telah memiliki reabilitas yang tinggi.

2. Jenis-jenis reliabilitas

1. Metode *test – retest*



Gambar (1) model tes-retest (Surapranata, 2009)

Test-retest adalah derajat yang menunjukkan konsistensi hasil sebuah tes dari waktu ke waktu (Sukardi, 2010). Pengetesan dua kali dengan menggunakan satu test (satu seri tes) yang sama pada waktu yang berbeda. Test seperti ini juga dikenal dengan istilah lain yaitu *single-test-double-trial method* (satu jenis tes-dicobakan dua kali). Kemudian hasil dari kedua kali tes tersebut dihitung korelasinya, maka nilai tersebut adalah reabilitas dari

tes. Reliabilitas tes-retes dapat dilakukan dengan cara seperti berikut:

- a) Selenggarakan tes pada suatu kelompok yang tepat sesuai dengan rencana,
- b) Setelah selang waktu tertentu, misalnya satu minggu atau dua minggu, lakukan kembali tes yang sama dengan kelompok yang sama tersebut,
- c) Korelasikan hasil kedua tes tersebut.

Jika hasil koefisien korelasi menunjukkan tinggi, berarti reliabilitas tes adalah bagus, jika korelasi rendah, berarti tes tersebut mempunyai konsistensi rendah. Tes ini memiliki kekurangan yaitu faktor waktu jeda atau tenggang antara tes pertama dengan tes kedua, mengenai interval waktu yang baik, Gay (1983:118) memberikan referensi bahwa satu hari terlalu pendek, sebaliknya satu bulan terlalu panjang, oleh karena itu, interval waktunya adalah satu atau dua minggu.

2. Metode tes sejajar equivalen

Tes paralel atau tes equivalen adalah dua buah tes yang mempunyai kesamaan tujuan, tingkat kesukaran, dan susunan, tetapi butir-butir soalnya berbeda. Dalam menggunakan metode tes ini, pembuat tes harus menyiapkan dua buah tes dan masing-masing dicobakan pada kelompok siswa yang sama. Skor dari kedua kelompok tes tersebut dikorelasikan untuk mendapatkan reliabilitas tes. Misalnya: dua buah tes paralel, tes matematika seri A dan tes seri B, kemudian diteskan kepada kelompok siswa yang sama, kemudian hasilnya dikorelasikan, jika koefisiennya tinggi maka tes tersebut sudah reliabel. Kelemahan metode ini adalah akan menyulitkan guru karena harus membuat dua seri tes.

3. Metode belah dua atau metode tes tunggal (*Single Test – Single Trial*)

Metode ini diciptakan oleh para ahli untuk menjawab kekurangan metode sebelumnya yaitu tes ulang dan tes paralel, metode ini juga dinamakan tes tunggal dilihat dari kepraktisannya lebih praktis dari pada dua metode sebelumnya yaitu: (1) metode ini hanya menyajikan satu tes untuk sekali tes kepada sekelompok subjek; (2) membagi tes tersebut menjadi dua bagian yang sama (sama banyak soal); (3) mengkorelasikan skor kedua belahan ini untuk mengestimasi reliabilitas tes. Koefisien reliabilitas dapat diperoleh dengan cara membelah instrument menjadi dua, tiga, empat, atau bahkan sebanyak

butir yang dimiliki oleh instrumen tersebut. Teknik perhitungannya tergantung pada banyaknya belahan, bentuk, serta sifat alat ukurnya. Terdapat berbagai macam cara membagi dua suatu tes, salah satu contohnya yaitu: 40 soal bahasa Arab dibelah menjadi dua belahan, belahan pertama terdiri atas soal nomor 1-20. Belahan kedua terdiri atas soal nomor 21-40. Antara soal belahan pertama dan kedua mempunyai tingkat kesulitan yang berbeda, nomor-nomor awal biasanya mudah dibanding nomor-nomor akhir, sehingga peserta tes mengalami kelelahan ketika mengerjakan soal belahan kedua.

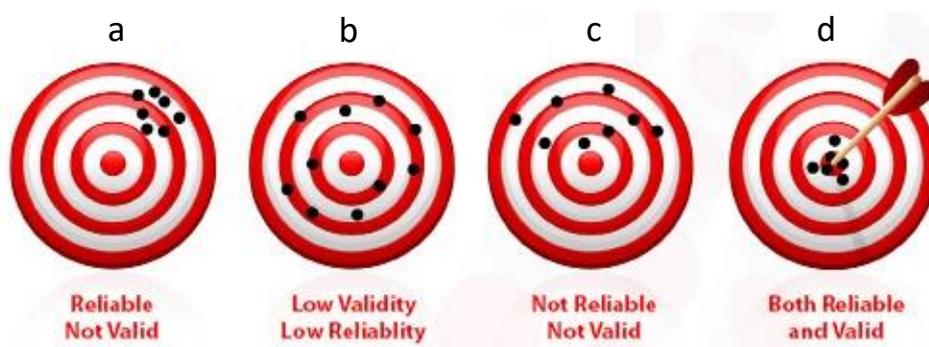
3. Faktor-faktor yang mempengaruhi reliabilitas

- a) Karakteristik umum permanen peserta didik
 - 1) Kemampuan yang dimiliki peserta didik dalam menghadapi tes
 - 2) Kemampuan umum dan teknik yang digunakan ketika mengambil tes
 - 3) Kemampuan umum untuk memahami petunjuk tes
- b) Karakteristik khusus permanen peserta didik
 - 1) Khusus yang berkaitan dengan tes secara keseluruhan
 - a. kemampuan peserta didik yang berkaitan dengan atribut yang diukur dalam sebuah tes
 - b. pengetahuan dan kemampuan khusus yang berkaitan dengan soal
 - c. keajegan respon peserta didik terhadap pilihan jawaban (misalnya mereka cenderung memberi jawaban A dari empat alternatif yang disediakan atau cenderung memilih B dari soal benar-salah yang disajikan).
 - 2) Khusus yang berkaitan dengan soal
 - a. Pengertian khusus yang berkaitan dengan fakta atau konsep khusus
 - b. Pengetahuan dan kemampuan khusus yang berkaitan dengan soal
- c) Karakteristik umum yang temporer seperti:
 - 1) Kelelahan
 - 2) Kesehatan
 - 3) Motivasi
 - 4) Gangguan emosi
 - 5) Kemampuan umum dan teknik yang digunakan ketika mengambil tes
 - 6) Pemahaman mekanisme tes
 - 7) Faktor panas, cahaya, ventilasi, dan lain sebagainya

- d) Karakteristik khusus yang temporer seperti:
- 1) Khusus yang berkaitan dengan tes secara keseluruhan
 - a. Pemahaman terhadap petunjuk khusus
 - b. Trik atau teknik-teknik mengatasi tes
 - c. Pengalaman/latihan menghadapi tes terlebih lagi dalam tes psikomoto
 - d. Kebiasaan menghadapi sebuah tes
 - 2) Khusus yang berkaitan dengan soal
 - a. Fluktuasi ingatan yang dimiliki peserta didik
 - b. Hal-hal yang berkaitan dengan perhatian dan keakuratan
- e) Faktor penyelenggaraan
- 1) Waktu, bebas dari gangguan, dan petunjuk yang jelas
 - 2) Pengawasan
 - 3) Pensekoran
- f) Faktor yang tidak pernah diperhitungkan
- 1) Keberuntungan karena faktor menebak
 - 2) Mengingat soal yang dilihatnya

C. Korelasi Antara Validitas dan Reliabilitas

validitas dan reliabilitas itu seperti orang menembak, atau membidik satu titik sasaran, semakin tepat bidikan ke sasaran tersebut, maka ia semakin valid dan juga reliabel, begitu juga dengan penilaian. suatu penilaian dikatakan valid dan reliabel apabila dapat mengukur apa yang diukur dengan tepat dan ajeg. Gronlund & linn (1990) memberikan ilustrasi kaitan antara validitas dan reliabilitas dengan mengibaratkan seorang penembak atau pemanah yang sedang membidik papan sasaran. Ada tiga kemungkinan hasil tembakannya, (di sini saya menambahkan satu dari referensi lain) sehingga seluruhnya berjumlah empat (Kusaeri, 2014).



ditetapkan

Gambar (b) hasil tembakan yang memiliki validitas dan reliabilitas rendah, karena tembakan menyebar pada bidang sasaran, tidak fokus pada titik yang telah ditetapkan

Gambar (c) hasil tembakan yang tidak valid dan tidak reliabel, karena tembakan labil dan merata ke semua luasan target yang telah ditetapkan

Gambar (d) hasil tembakan yang reliabel dan valid, karena tembakan mendekati atau mengumpul pada titik sasaran atau tepat sasaran dan masih dalam luasan konsisten.

PENUTUP

Berdasarkan pemaparan secara teoritis yang telah disampaikan sebelumnya, maka dapat disimpulkan bahwa validitas dan reliabilitas instrument tes baik pada mata pelajaran bahasa Arab maupun pelajaran lain, tidak terlepas dari tujuan atau standar kompetensi yang diharapkan, disamping itu pula banyak faktor lain yang dapat mempengaruhi valid atau tidaknya dan reliabel atau tidaknya suatu tes yaitu dengan cara membimbing siswa mengerjakan contoh soal, memperhatikan teknik mengajar, pembuatan soal yang mudah dipahami siswa, mengelola emosional siswa agar tidak tegang dalam mengerjakan soal serta mengarahkan siswa untuk memahami petunjuk soal dengan baik. Dengan demikian, Secara garis besar dapat disebutkan bahwa selain dari soal tes itu sendiri faktor internal dan eksternal dari siswa juga menjadi penyebab valid dan reliabelnya suatu tes atau soal.

DAFTAR PUSTAKA

- al maany. (2017). *Al-maany*. Www.Almaany.Com.
- Arikunto, S. (2012). *Dasar-Dasar Evaluasi Pendidikan*. Bumi Aksara.
- Kusaeri. (2014). *Acuan dan Teknik Penilaian Proses dan Hasil Belajar dalam Kurikulum 2013*. Ar-ruzz media.
- Sudijono, A. (1996). *Pengantar Evaluasi Pendidikan*. PT. Raja Grafindo Persada.
- Sudijono, A. (2011). *Pengantar Evaluasi Pendidikan*. PT. Raja Grafindo Persada.
- Sukardi. (2010). *Evaluasi Pendidikan, Prinsip dan Operasionalnya*. Bumi Aksara.
- Supardi. (2016). *Penilaian Autentik Pembelajaran Afektif, Kognitif, dan Psikomotor (Konsep dan Aplikasi)*. Rajawali Press.
- Surapranata, S. (2009). *Analisis, validitas, Reliabilitas dan Interpretasi Hasil Tes (Implementasi Hasil Tes)*. Remaja Rosda Karya.
- Uno.B.hamzah. Koni, S. (2014). *Assesment Pembelajaran*. Bumi Aksara.
- Zed, M. (2004). *Metode Penelitian Kepustakaan*. Yayasan Obor Indonesia.