

Uji Validitas Dan Reliabilitas

¹Nurhalimah. s ²Imroatus Shalihah ³Raswan ⁴Ubaid Ridlo

¹⁻⁴Uin Syarif Hidayatullah Jakarta, Indonesia

Email: ¹nurhalimahsibar@gmail.com ²iiksholihah1998@gmail.com
³raswan@uinjkt.ac.id ⁴ubaiddridlo@uinjkt.ac.id

Abstrak

Evaluasi pembelajaran ialah bagian integral dalam sistem pendidikan yang berfungsi Untuk menilai sejauh mana kompetensi peserta didik tercapai serta seberapa efektif proses pembelajaran yang dilakukan. Salah satu aspek utama dalam evaluasi pembelajaran adalah validitas dan reliabilitas instrumen evaluasi. Validitas menilai sejauh mana alat pengukuran dapat mengukur hal yang sebenarnya ingin diukur, sementara reliabilitas mengindikasikan sejauh mana hasil pengukuran yang diperoleh dari alat tersebut tetap konsisten. Dalam makalah ini, dibahas mengenai konsep validitas, jenis-jenisnya (validitas isi, validitas konstruk, dan validitas kriteria), serta teknik-teknik pengujian validitas yang umum digunakan, seperti expert judgment dan analisis korelasi. Selain itu, dibahas pula tentang reliabilitas, Yang berhubungan dengan sejauh mana instrumen dapat konsisten mengukur variabel yang sama pada waktu yang berbeda. Teknik-teknik pengujian reliabilitas, seperti test-retest, parallel forms, dan koefisien Cronbach's Alpha, dijelaskan untuk memastikan konsistensi hasil tes. Makalah ini menggariskan pentingnya kedua konsep tersebut dalam menjamin kualitas data yang diperoleh dari proses evaluasi, yang pada akhirnya dapat meningkatkan kualitas pengambilan keputusan dalam pendidikan.

Kata Kunci: Evaluasi pembelajaran, validitas, reliabilitas, instrument pengukuran,

Astrack

Learning evaluation is an integral part of the education system, serving to measure students' competency achievement as well as the effectiveness of the learning process. One of the main aspects of learning evaluation is the validity and reliability of the evaluation instruments. Validity assesses the extent to which a measurement instrument measures what it is supposed to measure, while reliability indicates the consistency of the results obtained from the instrument. This paper discussed the concepts of validity, its types (content validity, construct validity, and criterion validity), and common techniques for testing validity, such as expert judgment and correlation analysis, furthermore, reliability, which relates to the consistency of an instrument in measuring the same variable at different times, is also examined. Techniques for testing reliability, including test-retest, parallel forms, and cronbach's alpha coefficient, are explained to ensure the consistency of the result. This paper emphasizes the importance of both concepts in ensuring the quality of data obtained from the evaluation process, which ultimately can improve decision-making quality in education.

Keywords: *Learning evaluation, validity, reliability, measurement instrument*

PENDAHULUAN

Dalam sektor pendidikan, evaluasi pembelajaran memiliki peranan yang sangat penting untuk menilai sejauh mana kemampuan peserta didik tercapai serta seberapa efektif proses pembelajarannya. Evaluasi bukan hanya berfungsi untuk mengukur sejauh mana siswa berhasil, akan tetapi juga menjadi dasar dalam memperbaiki metode pengajaran, membuat keputusan terkait kurikulum, dan meningkatkan kualitas pendidikan secara keseluruhan. Sehingga hasil dari evaluasi dapat digunakan sebagai dasar dalam menetapkan keputusan yang sesuai, diperlukan instrumen pengukuran yang valid, yang dapat dengan akurat mengukur hal yang menjadi sasaran pengukuran (Lia, 2021). Oleh karena itu, validitas instrumen pengukuran sangat penting Guna menjamin bahwa informasi hasil penilaian secara akurat merepresentasikan kapasitas dan penguasaan materi siswa dengan objektif serta dapat diverifikasi secara ilmiah.

Tanpa adanya validitas yang cukup, hasil dari evaluasi dapat memberikan gambaran yang salah dan menciptakan bias dalam proses penilaian, Dampaknya akan merambat ke seluruh aspek mutu pendidikan, mencakup ketidakakuratan dalam mengenali kebutuhan pembelajaran siswa sampai kekeliruan dalam merumuskan kebijakan di bidang pendidikan. Maka dari itu, penguasaan komprehensif tentang pengertian validitas ragam bentuknya dan factor penentunya berkontribusi terhadap validitas menjadi krusial bagi seluruh pemangku kepentingan yang berkecimpung dalam penyusunan dan implementasi evaluasi pembelajaran—termasuk pendidik, dosen, akademisi bidang pendidikan, dan perancang kurikulum. Validitas tidak semata-mata persoalan teknis dalam pembuatan instrumen asesmen, melainkan juga terkait langsung dengan kredibilitas dan reliabilitas sistem pendidikan secara menyeluruh (Ida & Musyarofah, 2021).

Dalam konteks evaluasi pembelajaran, selain validitas, reliabilitas adalah aspek yang sangat penting untuk diperhatikan. Reliabilitas mengacu

pada seberapa jauh hasil pengukuran yang dilakukan dengan instrumen tertentu mampu menghasilkan data yang stabil konsisten dan terpercaya ketika instrumen tersebut diterapkan secara berulang dalam situasi yang sama. Tanpa reliabilitas yang baik, meskipun instrumen pengukuran valid, hasil evaluasi tetap tidak dapat diandalkan karena ketidakstabilan atau ketidakseragaman hasil yang diperoleh. Maka, pengujian reliabilitas menjadi langkah penting guna untuk memastikan bahwa instrumen yang digunakan dalam evaluasi pembelajaran benar-benar mampu menyajikan informasi yang konsisten dan dapat dipercaya. Karena Reliabilitas bukan hanya memastikan bahwa instrumen tersebut mengukur dengan konsisten, tetapi juga memastikan bahwa instrumen dapat diandalkan dalam berbagai kondisi. Dalam evaluasi pembelajaran, instrumen yang reliabel memastikan bahwa skor yang diperoleh peserta didik dapat mencerminkan kemampuan mereka secara stabil dan tidak terpengaruh oleh faktor luar yang tidak berkaitan. Berbagai teknik digunakan untuk menguji reliabilitas, di antaranya adalah Test-Retest, Bentuk Paralel (Parallel Forms), Konsistensi Internal (Internal Consistency), dan Reliabilitas Antar Penilai (Inter-Rater Reliability). Setiap metode memiliki cara yang berlainan dalam menilai keberhasilan dan kemantapan hasil pengukuran.

Dengan menguji validitas dan reliabilitas melalui berbagai teknik, diharapkan instrumen evaluasi yang digunakan dalam proses pembelajaran dapat memberikan data yang bukan hanya valid, tetapi juga konsisten dan dapat diandalkan. Dengan demikian, evaluasi yang dilakukan dapat lebih akurat, objektif, dan mampu memberikan gambaran yang jelas tentang pencapaian kompetensi peserta didik.

METODE PENELITIAN

Penelitian ini menerapkan pendekatan kualitatif dengan rancangan studi literatur untuk mengeksplorasi konsep validitas dan reliabilitas dalam konteks evaluasi pembelajaran Bahasa Arab. Penelitian ini bertujuan untuk menganalisis dan menggali pemahaman dari berbagai sumber literatur yang

relevan, termasuk artikel jurnal, buku, dan penelitian terdahulu, terkait dengan uji validitas dan reliabilitas dalam instrumen evaluasi pendidikan. Sumber data utama Pada penelitian ini ialah artikel-artikel ilmiah yang mengulas mengenai validitas dan reliabilitas dalam bidang pendidikan. Teknik Pengumpulan Data dalam penelitian ini dilakukan melalui studi literatur dengan cara mengumpulkan dan menelaah berbagai artikel ilmiah, jurnal, dan buku yang membahas uji validitas dan reliabilitas dalam evaluasi pembelajaran. Analisis data dilakukan melalui pendekatan kualitatif deskriptif.

HASIL DAN PEMBAHASAN

Konsep Validitas

Istilah validitas berasal dari kata *validity*, merujuk pada seberapa jauh akurasi dan ketepatan suatu instrumen pengukuran dalam melaksanakan fungsi pengukurannya. Dengan kata lain, validitas merupakan konsep yang menggambarkan seberapa jauh sebuah tes mampu mengukur hal yang seharusnya diukur (Dr. Sudaryono, 2018). Validitas ialah konsep bersifat mendasar dalam proses pengukuran dan evaluasi pendidikan. Secara etimologis, validitas berasal dari kata "validity" yang berarti kesahihan atau keabsahan. Berdasarkan kamus besar Bahasa Indonesia (KBBI), validitas diartikan sebagai karakteristik kebenaran berdasarkan bukti yang tersedia, penalaran logis, atau kekuatan hukum, sifat shahih, keabsahan (Ega Najwa Ismail:2025).

Menurut Arikunto (1999), validitas merupakan sebuah indicator yang menggambarkan seberapa jauh kesahihan suatu tes. Sebuah tes dikatakan valid jika tes tersebut mampu mengukur apa yang seharusnya diukur. Tes dianggap mempunyai validitas tinggi jika hasil yang diperoleh sesuai dengan kriteria yang ditetapkan, yakni ada kesesuaian antara tes tersebut dengan kriteria yang digunakan. Konsep yang dikemukakan oleh Maxwell (1996, h.87) bahwa 'validity is the correctness or credibility of a description,

conclusion, explanation, interpretation". Konsep ini menekankan validitas sebagai suatu akurasi atau kredibilitas dari sebuah deskripsi, konklusi, penjelasan serta interpretasi temuan penelitian.

Menurut Mareceki (2009), validitas dijelaskan sebagai 'penilaian sejauh mana bukti penelitian mendukung atau membenarkan interpretasi dan kesimpulan yang dihasilkan'. Dalam perspektif ini, validitas dipahami sebagai proses asesmen guna menetapkan interpretasi dan konklusi yang diambil dari penelitian didasarkan pada bukti atau data yang tersedia. Oleh karena itu, dapat disimpulkan bahwa validitas dalam penelitian kualitatif berkaitan dengan keakuratan dalam menjalankan prosedur penelitian, sehingga hasil dari konklusi yang didapat dapat diandalkan sebagai suatu kebenaran yang abash.

Validitas suatu tes umumnya dibagi dalam dua jenis, ialah validitas logis dan validitas empiris. Validitas logis merujuk pada analisis kualitatif terhadap suatu soal untuk menentukan apakah soal tersebut berfungsi dengan baik, berdasarkan kriteria yang telah ditetapkan, seperti kriteria materi, struktur, dan bahasa. Suatu tes atau instrument pengukuran dinyatakan memiliki validitas tinggi apabila instrument tersebut dapat menjalankan fungsinya dengan baik, yaitu menghasilkan hasil pengukuran yang selaras sesuai dengan tujuan yang hendak dicapai. Dengan demikian, hasil pengukuran yang didapat harus menggambarkan secara akurat fakta atau kondisi sesunguhnya dari apa yang diukur.

Sakah satu metode guna mengenali instrumen yang valid adalah dengan memeriksa terlebih dahulu apakah data yang dipakai valid atau tidak. Dengan demikian, dapat dipastikan bahwa jika data tersebut memenuhi kriteria validitas, maka hasil yang diperoleh akan menghasilkan instrumen yang juga valid.

Validitas dibagi menjadi tiga jenis utama, yaitu:

(a) validitas isi (*content validity*)

Validitas isi merupakan validitas yang mempersoalkan seberapa jauh suatu tes atau alat ukur menilai tingkat penguasaan terhadap konten materi tertentu yang semestinya dikuasai sesuai dengan sasaran pembelajaran. Validitas isi mengambarkan sejauh mana instrument mempresentasikan konten yang diinginkan. Validitas isi tidak dapat diekspresikan dalam bentuk angka, melainkan dikaji secara rasional atau melalui penilaian professional. Pengujian validitas isi dapat dijalankan dengan membandingkan antara konten instrument dengan materi pembelajaran atau dengan meminta pendapat para ahli (expert judgment).

(b) validitas konstruk (*construct validity*) (Muhammad Isa Anshari: 2024)

Validitas konstruk merupakan validitas yang berhubungan dengan kemampuan suatu instrumen mengukur konsep-konsep yang terdapat dalam materi yang diukurnya. Validitas konstruk mempersoalkan apakah item-item pertanyaan dalam instrument telah selaras dengan konsep teori yang dipakai. Pengujian validitas konstruk dapat dijalankan dengan cara mengorelasikan antara skor butir dengan skor total atau menggunakan analisis faktor. Validitas konstruk sangat penting dalam penelitian yang mengukur variabel psikologis atau konsep abstrak lainnya.

(c) validitas empiris atau validitas kriteria

Validitas ramalan bermakna dikaitkan dengan kriteria tertentu. Dalam validitas ini yang diutamakan bukanlah isi tes melainkan kriterianya, apakah alat ukur tersebut dapat digunakan untuk memprediksi suatu ciri atau perilaku tertentu atau kriteria tertentu yang diinginkan. Misalnya alat ukur motivasi belajar, apakah dapat digunakan untuk meramal prestasi belajar yang dicapai.

Validitas tidak melekat pada tes itu sendiri, melainkan pada hasil atau skor yang dihasilkan dari proses pengukuran. Suatu tes hasil belajar dapat dinyatakan valid apabila benar-benar mampu mengukur hasil belajar yang dimaksud, bukan sekedar mengukur kemampuan mengingat atau keterampilan berhsa semata.

Dalam uji validitas kuesioner, mempunyai dua jenis yang perlu dipahami, yaitu validitas faktor dan validitas item. Validitas faktor diukur saat item-item yang disusun mencakup lebih dari satu faktor, di mana antar faktor tersebut memiliki kesamaan. Pengujian validitas faktor dijalankan dengan mengorelasikan skor faktor (hasil penjumlahan item dalam satu faktor) dengan skor total faktor (jumlah dari seluruh faktor). Sedangkan validitas item mengacu pada korelasi atau dukungan terhadap total skor item. Pengukuran validitas butir dilakukan dengan mengorelasikan skor setiap item dengan skor total. Apabila lebih dari satu faktor digunakan, pengujian validitas item dilakukan dengan mengorelasikan skor setiap item dengan skor faktor, selanjutnya diteruskan dengan mengorelasikan skor item dengan skor total faktor (jumlah dari beberapa faktor) (Musrifah Mardiani Sanaky, 2021). Uji validitas ini dijalankan dengan memakai program SPSS. Salah satu teknik yang umum dipakai oleh peneliti guna menguji validitas yaitu dengan menerapkan korelasi Bivariate Pearson (Produk Momen Pearson).

Faktor-faktor Mempengaruhi Validitas:

Berbagai faktor yang memengaruhi hasil evaluasi, sehingga hasilnya menjadi tidak valid dan tidak mencerminkan kondisi sebenarnya sesuai dengan tujuan penggunaannya. Beberapa faktor tersebut berasal dari alat ukur evaluasi itu sendiri. Dalam konteks proses pembelajaran, faktor-faktor tersebut dapat mengurangi efektivitas fungsi utama dari evaluasi yang diharapkan, sehingga menurunkan validitas alat evaluasi tersebut. Faktor-faktor yang dimaksud antara lain adalah: (Ega Najwa Ismail: 2025)

1. Faktor Isi Tes

Konten tes sangat berpengaruh pada sejauh mana butir-butir soal yang disusun dapat merepresentasikan seluruh kompetensi atau materi yang akan diukur. Jika isi tes tidak sesuai dengan tujuan, maka hasil tes tersebut tidak dapat dikatakan valid

2. Faktor Konstruksi Instrumen

Konstruksi instrumen terkait dengan bagaimana tes tersebut disusun dan dirancang, mencakup aspek format, struktur, serta penyusunan soal. Validitas alat pengukuran akan menurun jika instrumen disusun dengan format yang kurang terorganisir, penggunaan bahasa yang membingungkan, atau petunjuk penggerjaan yang tidak jelas. Agar validitas tetap terjaga, konstruksi instrumen harus memperhatikan beberapa hal penting, seperti kejelasan instruksi, keterbacaan soal, pemilihan tipe soal yang sesuai dengan tujuan pengukuran, serta urutan soal yang logis.

3. Faktor Jawaban peserta didik

Validitas instrumen pengukuran dapat terpengaruh jika jawaban yang diberikan peserta didik tidak sepenuhnya mencerminkan kemampuan atau kompetensi yang sesungguhnya. Faktor-faktor seperti kondisi fisik dan mental peserta didik selama tes, motivasi dalam mengerjakan tes, atau bahkan kesalahan dalam memahami soal, dapat menjadi penyebab terjadinya hal tersebut.

4. Pengaturan soal yang kurang tepat

Rumus Koefisien Korelasi Pearson (r_{xy}):

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{[N \sum X^2 - (\sum X)^2][N \sum Y^2 - (\sum Y)^2]}}$$

r_{xy} adalah koefisien korelasi Pearson antara variabel X dan Y.

N adalah jumlah pasangan data.

X dan Y adalah nilai-nilai dari dua variabel yang sedang dianalisis.

$\sum X$ adalah jumlah dari seluruh nilai X.

$\sum Y$ adalah jumlah dari seluruh nilai Y.

$\sum XY$ adalah jumlah dari hasil perkalian pasangan data X dan Y.

$\sum X^2$ adalah jumlah kuadrat dari nilai X.

$\sum Y^2$ adalah jumlah kuadrat dari nilai Y.

Pengertian Reliabilitas

Uji reliabilitas sering digunakan untuk mengukur konsistensi suatu alat pengukur, seperti kuesioner yang memuat indikator-indikator berkaitan dengan variabel atau konstruk tertentu. Tujuan utama dari uji reliabilitas adalah untuk memastikan bahwa instrumen pengukur tersebut menghasilkan hasil yang konsisten dan berkesinambungan ketika digunakan dalam percobaan atau pengukuran yang dilakukan berulang kali, serta dapat dipercaya. Dalam uji reliabilitas, sebuah instrumen pengukuran dianggap reliabel jika menghasilkan hasil yang konsisten saat diukur pada kesempatan yang berbeda. Dengan kata lain, alat ukur tersebut dapat diandalkan karena mampu menghasilkan temuan yang relative sama meskipun digunakan dalam kondisi pengukuran yang berbeda. Hal ini menegaskan bahwa instrument tersebut memiliki stabilitas dan konsistensi yang memadai untuk evaluasi yang akurat dan dapat diandalkan (Ghozali, 2018).

Menurut Nursalam, Reliabilitas berkaitan dengan tingkat keseragaman hasil pengukuran atau pengamatan apabila dilakukan pada waktu yang berbeda. Oleh karena itu, alat ukur dan metode yang digunakan perlu diperhatikan secara serius

karena keduanya saling berkaitan dalam memastikan keseragaman hasil pada waktu yang bersamaan. Dalam konteks kuesioner, reliabilitas menunjukkan sejauh mana jawaban seseorang terhadap suatu pertanyaan dalam tes tetap konsisten. Jika respons yang diberikan relatif sama atau konsisten saat tes dilakukan berulang kali, maka kuesioner tersebut dianggap reliabel dan dapat dipercaya. Uji reliabilitas digunakan untuk mengetahui sejauh mana tes mampu menghasilkan data yang konsisten dan akurat. Dengan kata lain, uji reliabilitas memastikan bahwa tes yang dilakukan menghasilkan hasil yang stabil dan hampir sama setiap kali tes tersebut digunakan (Nursalam, 2003).

Secara umum, reliabilitas (keandalan) merupakan ukuran sejauh mana instrumen pengukuran menghasilkan hasil yang konsisten atau stabil jika digunakan secara berulang dalam situasi yang serupa. Dalam konteks tes pendidikan, reliabilitas menyangkut sejauh mana skor yang diperoleh peserta dari tes tersebut dapat dipercaya sebagai refleksi dari kondisi yang sama jika tes tersebut diulang atau sedikit diubah variabelnya (Singh, Tejinder). Sebagaimana dijelaskan dalam artikel “A Primer on the Validity of Assessment Instruments”, “Reliability refers to whether an assessment instrument gives the same results each time it is used in the same setting with the same type of subjects (Faiz Zahfa dkk: 2025). Dengan demikian, dapat dirumuskan bahwa reliabilitas merupakan indicator konsistensi dan kestabilan skor yang diperoleh oleh suatu instrumen pengukuran (misalnya tes), dalam kondisi yang serupa.

Teknik-Teknik Pengujian Reliabilitas Tes

1) Test-Retest (Tes Ulang)

Pendekatan ini mencangkup pemberian tes yang sama kepada kelompok yang sama dalam dua waktu berbeda, kemudian dihitung korelasi antara skor pertama dan kedua. Semakin tinggi korelasi, semakin tinggi reliabilitasnya.

2) Bentuk Paralel (Parallel Forms)

Metode ini menggunakan dua versi tes yang berbeda (versi A dan versi B) yang setara dan diberikan pada kelompok yang sama. Korelasi antara kedua versi mencerminkan tingkat reliabilitas.

3) Konsistensi Internal (Internal Consistency)

Pendekatan ini mengukur sejauh mana butir-butir dalam tes tersebut mengukur konstruk yang sama. Teknik yang sering digunakan ialah Split-Half Reliability dan Cronbach's Alpha.

4) Reliabilitas Antar Penilai (Inter-Rater Reliability)

Digunakan ketika skor didasarkan pada penilaian dari beberapa penilai (rater). Semakin tinggi kesepakatan antar penilai, semakin tinggi reliabilitasnya (Supriadi, G).

KESIMPULAN

Validitas dan reliabilitas merupakan dua komponen penting yang tidak bisa dipisahkan dalam proses pengembangan dan evaluasi alat pengukuran. Validitas merujuk pada sejauh mana tes atau instrumen secara tepat mampu mengukur aspek yang memang seharusnya diukur, yang mencakup validitas isi, validitas konstruk, dan validitas kriteria. Selain itu, reliabilitas berkaitan dengan konsistensi dan kestabilan hasil pengukuran saat tes dilakukan berulang dalam situasi yang sama. Kedua konsep ini saling melengkapi; meskipun suatu instrumen bisa reliabel tetapi tidak valid, instrumen yang valid tentu harus reliabel. Karena itu, pemahaman yang mendalam tentang kedua konsep ini sangat krusial bagi para peneliti, pendidik, dan praktisi yang menggunakan instrumen pengukuran di berbagai bidang, seperti pendidikan, psikologi, dan penelitian sosial. Penerapan prinsip validitas dan reliabilitas dalam praktik memerlukan pemilihan teknik pengujian yang tepat sesuai dengan tujuan dan karakteristik instrumen yang dikembangkan. Untuk menguji validitas, dapat digunakan expert judgment, analisis korelasi, atau analisis faktor, sedangkan reliabilitas dapat diuji melalui metode test-retest, parallel forms, split-half, atau koefisien Alpha Cronbach. Hasil pengujian kedua

aspek ini menjadi dasar untuk menentukan apakah suatu instrumen layak digunakan atau perlu diperbaiki. Dengan memastikan apabila instrument memiliki tingkat validitas dan reliabilitas yang baik, maka data yang dihasilkan akan tepat, dapat diandalkan, dan memberikan informasi yang bermakna untuk pengambilan keputusan atau pengembangan ilmu pengetahuan.

Reliabilitas merupakan aspek penting dalam penyusunan dan penggunaan instrumen penelitian maupun tes hasil belajar di bidang pendidikan. Reliabilitas menunjukkan sejauh mana suatu instrument pengukuran memberikan hasil yang seragam dan stabil saat digunakan secara berulang dalam kondisi yang relative sama. Instrumen yang memiliki reliabilitas tinggi menandakan bahwa hasil pengukuran dapat dipercaya dan mencerminkan keadaan yang sebenarnya dari objek yang diukur. Berbagai teknik pengujian reliabilitas, seperti *test-retest*, bentuk paralel (*parallel forms*), konsistensi internal (*internal consistency*), dan reliabilitas antar-penilai (*inter-rater reliability*), dapat digunakan sesuai dengan jenis data dan tujuan penelitian. Pemilihan teknik yang tepat akan meningkatkan keakuratan hasil evaluasi dan memperkuat validitas kesimpulan penelitian.

DAFTAR PUSTAKA

- Dr. Budiaستuti Dyah, (2018) *Validitas dan reliabilitas penelitian*
- Ega Najwa Ismail, (2025) VALIDITAS ALAT UKUR DALAM EVALUASI PEMBELAJARAN: STUDI FAKTOR YANG MEMPENGARUHI VALIDITAS ALAT UKUR, *Jurnal Kajian Ilmiah Interdisiplinier*, 9
- Faiz, Z., dkk. (2025). *Validity and Reliability of Educational Evaluation Tests*. *Jurnal Riset Ilmu Pendidikan*, 5(1), 1–5.
- Ghozali, I. (2018). *Aplikasi Analisis Multivariate dengan Program IBM SPSS 25* (Edisi 9). Semarang: Badan Penerbit Universitas Diponegoro.
- Khumaedi, M. (2012). *Reliabilitas Instrumen Penelitian Pendidikan*. *Jurnal Pendidikan Teknik Mesin*, 12(1). <https://doi.org/10.15294/jptm.v12i1.5273>

Muhammad Isa Anshari, (2024) Analisi Validitas dan Reliabilitas butir soal sumatif akhir semester ganjil mata pelajaran PAI, Edukatif: *Jurnal Ilmu Pendidikan*.

Musrifah Mardiani Sanaky dkk, (2021) Ananlisis factor-faktor penyebab keterlambatan pada proyek pembangunan gedung asrama MAN 1 Tulehu Maluku tengah, *Jurnal Simetrik*.

Nursalam. (2003). *Konsep dan Penerapan Metodologi Penelitian Ilmu Keperawatan: Pedoman Skripsi, Tesis, dan Instrumen Penelitian Keperawatan*. Jakarta: Salemba Medika.

Singh, T. (2024). *Reliability in Student Assessment: Implications for Competency-Based Curriculum*. *Indian Journal of Health Sciences and Biomedical Research (KLEU)*, 17(2), 95–99. https://doi.org/10.4103/kleuhsj.kleuhsj_15_2

Supriadi, G. (2016). *Reliabilitas Tes Hasil Belajar dan Aplikasinya dalam Evaluasi Pembelajaran*. *Himmah*, 7, 51–62.

Sudaryono, Dr. (2017) *Penelitian kuantitatif, kualitatif, dan mix method*
Rasyid, Arbanur, and Sawaluddin Siregar. "Fenomena Menarik Perkawinan

Dibawah Umur." *AL-MANHAJ: Jurnal Hukum Dan Pranata Sosial Islam* 4, no. 1 (2022): 61–68. <https://doi.org/10.37680/almanhaj.v4i1.1571>.