TSAQOFAH
Jurnal Penelitian Guru Indonesia

p-ISSN: 2810-0395
e-ISSN: 2810-0042

Terindeks : Dimensions, Scilit, Lens, Semantic, Garuda, Google Scholar, Base, etc

https://doi.org/10.58578/tsaqofah.v6i2.8511

RELIABILITAS ALAT UKUR, JENIS-JENIS, CARA PENGUKURAN, DAN FAKTOR-FAKTOR YANG MEMPENGARUHINYA

Reliability of Measurement Instruments: Types, Methods of Assessment, and Influencing Factors

Achmad Rasyid, Rudy Hartanto, Catur Priyanto
Institut Islam Mamba'ul Ulum Surakarta
rudyzahrawain88@gmail.com, truth85@gmail.com

Article Info:

Submitted:	Revised:	Accepted:	Published:
Nov 15, 2025	Dec 7, 2025	Dec 19, 2025	Dec 24, 2025

Abstract

Measurement reliability is a crucial aspect of research because it concerns the consistency and trustworthiness of measurement results, whereby a reliable instrument is able to produce stable data when used repeatedly under the same conditions. This paper discusses the concept of measurement reliability, its types, testing methods, and the factors that influence the level of reliability of research instruments. The types of reliability reviewed include response consistency reliability, composite item consistency, external reliability, and internal reliability. The reliability testing methods discussed comprise test–retest, parallel forms, split-half, Cronbach's alpha, and inter-rater reliability. In addition, this study highlights various factors that affect reliability, such as item clarity, respondent conditions, measurement environment, number of items, time interval between measurements, and assessor quality. A sound understanding of reliability is expected to help researchers design and use instruments that are more

Tsaqofah journal is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License

accurate, consistent, and scientifically accountable, thereby improving the overall quality of research findings.

Keywords: Measurement Reliability; Research Instruments; Measurement Consistency; Data Validity; Reliability Testing Methods

Abstrak: Reliabilitas alat ukur merupakan aspek krusial dalam penelitian karena berkaitan dengan konsistensi dan keterpercayaan hasil pengukuran, di mana instrumen yang reliabel mampu menghasilkan data yang stabil ketika digunakan berulang kali pada kondisi yang sama. Kajian ini membahas konsep reliabilitas alat ukur, jenis-jenis reliabilitas, metode pengujiannya, serta faktor-faktor yang memengaruhi tingkat reliabilitas instrumen penelitian. Jenis reliabilitas yang diulas meliputi reliabilitas konsistensi tanggapan, konsistensi gabungan item, reliabilitas eksternal, dan reliabilitas internal. Metode pengukuran reliabilitas yang dibahas mencakup *test–retest*, bentuk sejajar, belah dua, *Cronbach's Alpha*, dan reliabilitas antarpenilai. Selain itu, kajian ini menyoroti berbagai faktor yang memengaruhi reliabilitas, seperti kejelasan item, kondisi responden, lingkungan pengukuran, jumlah item, interval waktu pengukuran, dan kualitas penilai. Pemahaman yang baik terhadap reliabilitas diharapkan dapat membantu peneliti merancang dan menggunakan instrumen yang lebih akurat, konsisten, dan dapat dipertanggungjawabkan secara ilmiah, sehingga meningkatkan kualitas temuan penelitian secara keseluruhan.

Kata Kunci: Reliabilitas Alat Ukur; Instrumen Penelitian; Konsistensi Pengukuran; Validitas Data; Metode Pengujian Reliabilitas

PENDAHULUAN

Dalam kegiatan penelitian, alat ukur memiliki peranan yang sangat penting untuk memperoleh data yang akurat dan dapat dipercaya. Alat ukur dalam konteks penelitian sosial, psikologi, maupun pendidikan tidak hanya mengacu pada alat fisik, tetapi juga mencakup instrumen seperti angket, kuesioner, atau tes. Keandalan dari alat ukur ini ditentukan oleh sejauh mana alat tersebut menghasilkan hasil yang konsisten ketika digunakan berulang kali pada kondisi yang sama. Konsep ini dikenal sebagai reliabilitas.

Reliabilitas merupakan salah satu kriteria utama dalam penilaian kualitas instrumen penelitian, selain validitas (Subhaktiayasa, 2024). Tanpa reliabilitas yang baik, data yang diperoleh akan sulit dipercaya karena dapat mengandung kesalahan pengukuran yang besar. Oleh karena itu, memahami konsep reliabilitas, jenis-jenisnya, cara pengukurannya, serta faktor-faktor yang mempengaruhi sangat penting bagi setiap peneliti agar dapat menghasilkan data yang berkualitas dan bermakna.

Selain memastikan konsistensi hasil pengukuran, reliabilitas juga berfungsi sebagai dasar dalam pengambilan kesimpulan penelitian. Instrumen yang reliabel memungkinkan peneliti meminimalkan bias dan kesalahan acak, sehingga hasil penelitian dapat merepresentasikan kondisi sebenarnya dari objek yang diteliti. Dengan demikian, reliabilitas berperan penting dalam menjaga objektivitas penelitian dan meningkatkan kepercayaan terhadap temuan yang dihasilkan (Kaddafi, 2025).

Dalam praktik penelitian, pengujian reliabilitas menjadi tahap yang tidak dapat diabaikan sebelum instrumen digunakan secara luas. Berbagai teknik dan pendekatan dapat digunakan untuk menguji reliabilitas, tergantung pada jenis instrumen dan karakteristik data yang dikumpulkan. Pemilihan metode uji reliabilitas yang tepat akan membantu peneliti memastikan bahwa instrumen yang digunakan benar-benar layak, sehingga proses penelitian dapat berjalan secara sistematis, terukur, dan bertanggung jawab secara ilmiah.

PEMBAHASAN

1. Pengertian Reliabilitas Alat Ukur

Reliabilitas berasal dari kata reliable yang berarti dapat dipercaya. Dalam konteks penelitian, reliabilitas diartikan sebagai tingkat konsistensi suatu alat ukur dalam menghasilkan hasil yang sama ketika digunakan berulang kali pada kondisi yang sama (Azwar, 2017).

Menurut Sugiyono (2018), reliabilitas menunjukkan sejauh mana hasil pengukuran dengan alat tertentu dapat dipercaya dan tidak mengalami perubahan berarti bila pengukuran diulang. Semakin kecil perbedaan antara hasil pengukuran pertama dan pengukuran berikutnya, semakin tinggi tingkat reliabilitas instrumen tersebut.

Secara khusus, konsep reliabilitas mengacu pada konsistensi hasil score pada item-item yang terdapat pada kuesioner Anda sehingga uji reliabilitas sesunggunya menguji ketepatan skala-skala pengukuran instrumen penelitian. Dengan demikian tujuan utama uji reliabilitas instrumen penelitian ialah untuk mengukur konsistensi alat ukur yang digunakan peneliti kuantitatif. Dalam konteks ini, peneliti hendak mengetahui apakah terdapat ketepatan hasil pengukuran pada sampel yang sama dalam waktu yang berbeda. Dengan kata lain, sebuah instrumen penelitian, misalnya kuesioner dinyatakan reliabel jika instrumen tersebut dapat menyediakan hasil skor yang konsisten pada setiap

pengukuran. Dengan demikian, alat pengukuran tersebut (butir-butir pernyataan/pertanyaan) tetap menyediakan hasilpengukuran yang konsisten dalam waktu yang berbeda (Ridha & Al Ghifari, 2025).

2. Jenis-Jenis Reliabilitas Alat Ukur

Reliabilitas merupakan salah satu unsur penting yang harus dipenuhi oleh instrumen penelitian agar data yang diperoleh dapat dipertanggungjawabkan secara ilmiah. Instrumen yang reliabel mampu menghasilkan data yang stabil dan konsisten. Djali dan Muljono (2008) menyatakan bahwa reliabilitas instrumen penelitian terdiri atas beberapa jenis yang digunakan untuk menilai tingkat keajegan hasil pengukuran.

Jenis reliabilitas yang pertama adalah reliabilitas konsistensi tanggapan, yaitu reliabilitas yang menekankan pada keseragaman respons responden terhadap instrumen yang digunakan. Instrumen dikatakan memiliki konsistensi yang baik apabila hasil pengukuran yang diperoleh relatif sama ketika pengukuran dilakukan secara berulang pada kondisi yang sama. Apabila hasil pengukuran menunjukkan perbedaan yang signifikan, maka instrumen tersebut dianggap kurang mampu mencerminkan kondisi objek penelitian secara akurat. Untuk mengetahui tingkat konsistensi ini, pengujian dilakukan pada objek yang sama dengan menggunakan beberapa teknik, antara lain teknik *test-retest*, teknik belah dua dengan membagi item instrumen ke dalam dua bagian yang setara, serta teknik bentuk ekuivalen dengan menggunakan dua instrumen yang sebanding pada waktu yang sama (Ridha & Al Ghifari, 2025).

Jenis berikutnya adalah reliabilitas konsistensi gabungan item, yaitu reliabilitas yang berfokus pada keterpaduan antaritem dalam satu instrument. Ketidakkonsistenan hasil pengukuran antaritem terhadap objek yang sama menunjukkan adanya kelemahan pada instrumen atau alat ukur yang digunakan. Instrumen dengan kondisi tersebut dinilai memiliki reliabilitas yang rendah. Pengujian reliabilitas konsistensi gabungan item dilakukan dengan mempertimbangkan jenis instrumen yang digunakan, seperti tes objektif, tes uraian, dan tes afektif.

Secara umum, pengujian reliabilitas instrumen dapat dilakukan melalui dua pendekatan, yaitu reliabilitas eksternal dan reliabilitas internal. Reliabilitas eksternal merupakan pengujian reliabilitas yang menggunakan kriteria di luar instrumen. Metode yang termasuk dalam reliabilitas eksternal antara lain metode tes ulang, metode paralel, dan metode belah dua. Metode tes ulang dilakukan dengan memberikan tes yang sama kepada responden

pada waktu yang berbeda, kemudian skor hasil pengukuran dikorelasikan menggunakan teknik korelasi *product moment*. Metode paralel menggunakan dua instrumen yang setara dan diberikan kepada kelompok responden yang sama, lalu hasil keduanya dikorelasikan untuk menentukan tingkat reliabilitas instrumen. Metode ini digunakan untuk mengurangi kelemahan yang muncul pada metode tes ulang, khususnya pengaruh pengulangan tes (Ridha & Al Ghifari, 2025).

Sementara itu, reliabilitas internal merupakan pengujian reliabilitas yang dilakukan satu kali menggunakan satu instrumen. Data hasil pengujian dianalisis dengan rumus reliabilitas yang disesuaikan dengan bentuk instrumen. Pada instrumen dengan skor dikotomi, seperti benar-salah dengan skor 0 dan 1, reliabilitas dapat dihitung menggunakan rumus Kuder Richardson 20 (KR-20) atau Kuder Richardson 21 (KR-21). Adapun instrumen yang berbentuk kontinum atau skala, seperti skala penilaian atau skala Likert, pengujian reliabilitasnya dilakukan dengan menggunakan rumus Alpha Cronbach.

3. Cara Pengukuran Reliabilitas

Dalam penelitian, reliabilitas merupakan indikator penting untuk menilai konsistensi hasil pengukuran. Terdapat beberapa metode yang umum digunakan dalam mengukur reliabilitas alat ukur, antara lain: Metode Ulang-Uji (Test-Retest), Metode Bentuk Sejajar (Equivalent Forms), Metode Belah Dua (Split-Half), Metode Konsistensi Internal (Cronbach's Alpha), dan Metode Antar Penilai (Inter-Rater Reliability). Berikut penjelasan lengkap beserta contoh penerapannya (Murti, 2011).

a. Metode Ulang-Uji (Test-Retest Method)

Metode Test-Retest digunakan untuk menilai sejauh mana suatu alat ukur memberikan hasil yang stabil dari waktu ke waktu. Jika instrumen yang sama digunakan pada kelompok responden yang sama di dua waktu berbeda, maka hasilnya harus konsisten.

Prosedur pengujian reliabilitas dilakukan dengan cara memberikan instrumen penelitian kepada sekelompok responden pada tahap awal. Setelah melewati jangka waktu tertentu, instrumen yang sama kembali diberikan kepada responden yang sama. Hasil skor dari pengukuran pertama dan kedua kemudian dianalisis dengan menghitung nilai korelasi menggunakan rumus korelasi Pearson. Nilai korelasi yang mendekati angka 1 menunjukkan bahwa instrumen memiliki tingkat reliabilitas yang tinggi, karena mampu menghasilkan skor yang konsisten dari waktu ke waktu.

Contoh: Seorang peneliti memberikan angket motivasi belajar kepada 50 siswa, lalu mengulanginya dua minggu kemudian. Hasil korelasi r = 0,89 menunjukkan reliabilitas tinggi.

Kelebihan: Mengukur stabilitas terhadap waktu. Kekurangan: Butuh dua kali pengukuran.

b. Metode Bentuk Sejajar (Equivalent / Parallel Forms Method)

Metode ini menggunakan dua versi alat ukur yang setara untuk menilai kesetaraan hasil pengukuran. Kedua versi diberikan kepada kelompok yang sama, lalu hasilnya dikorelasikan.

Contoh: Peneliti membuat dua versi tes pemahaman bahasa Inggris (Tes A dan Tes B) dengan tingkat kesulitan sama. Korelasi skor antara Tes A dan Tes B sebesar r = 0,82 menunjukkan reliabilitas tinggi.

Kelebihan: Tidak terpengaruh faktor waktu. Kekurangan: Sulit membuat dua versi tes yang setara.

c. Metode Belah Dua (Split-Half Method)

Metode ini menilai konsistensi internal dengan membagi item tes menjadi dua bagian (misalnya soal ganjil dan genap), kemudian mengkorelasikan hasil kedua bagian. Reliabilitas total dihitung dengan rumus Spearman-Brown: $r_{sb} = 2r_{xy} / (1 + r_{xy})$

Contoh: Tes matematika 40 soal dibagi menjadi dua bagian, dan korelasi antar bagian = 0,76. Reliabilitas total = 2(0,76)/(1+0,76) = 0,86, menunjukkan reliabilitas sangat tinggi.

Kelebihan: Efisien karena hanya satu kali pengukuran. Kekurangan: Pembagian item harus seimbang.

d. Metode Konsistensi Internal (Cronbach's Alpha)

Metode ini paling sering digunakan pada penelitian sosial atau pendidikan yang menggunakan skala Likert. Cronbach's Alpha mengukur sejauh mana item-item dalam instrumen mengukur hal yang sama. Rumus: $r_{\alpha} = (k / (k - 1)) * (1 - (\Sigma s_i^2 / s_t^2))$

Keterangan: k = jumlah item, s_i^2 = varians tiap item, s_t^2 = varians total. Nilai α antara 0–1. Umumnya, nilai di atas 0,70 dianggap reliabel.

Contoh: Kuesioner motivasi belajar dengan 15 item menghasilkan Cronbach's Alpha = 0,88. Artinya, instrumen tersebut sangat reliabel.

Kelebihan: Praktis, hanya butuh satu kali pengambilan data. Kekurangan: Alpha bisa

meningkat hanya karena banyaknya item.

e. Metode Antar Penilai (Inter-Rater Reliability)

Metode ini digunakan ketika pengukuran melibatkan lebih dari satu penilai (rater). Tujuannya untuk menilai konsistensi atau kesepakatan antar penilai terhadap objek yang sama.

Jika data kategorikal, digunakan Cohen's Kappa (κ); jika data numerik, digunakan Intraclass Correlation Coefficient (ICC).

Contoh: Dua guru menilai hasil presentasi 30 siswa. Nilai Cohen's Kappa = 0,81 menunjukkan tingkat kesepakatan sangat tinggi.

Kelebihan: Cocok untuk penilaian observasi. Kekurangan: Dapat dipengaruhi perbedaan persepsi antar penilai.

Kesimpulan umum :

Metode	Fokus Reliabilitas	Data yang Umum Digunakan	Kelebihan Utama
Test-Retest	Stabilitas waktu	Tes/kuesioner kuantitatif	Mengukur konsistensi dari waktu ke waktu
Bentuk Sejajar	Kesetaraan antar Bentuk	Tes kemampuan	Menilai dua versi alat ukur
Split-Half	Konsistensi antar Bagian	Tes objektif	Efisien, satu kali Pengukuran
Cronbach's Alpha	Konsistensi internal	Skala Likert	Praktis untuk penelitian sosial
Inter-Rater	Kesepakatan antar penilai	Observasi/performa	Mengukur konsistensi antar Pengamat

4. Faktor-Faktor yang Mempengaruhi Reliabilitas

a. Kejelasan Item Instrumen

Kejelasan item atau butir pertanyaan dalam instrumen sangat menentukan reliabilitas (Utami, 2023). Jika setiap item disusun dengan bahasa yang jelas, tidak ambigu, dan mudah dipahami oleh responden, maka kemungkinan kesalahan interpretasi menjadi kecil. Sebaliknya, item yang membingungkan atau memiliki makna ganda dapat menyebabkan jawaban responden tidak konsisten, sehingga reliabilitas menurun.

Contoh: pertanyaan seperti 'Apakah Anda sering membaca buku?' tanpa penjelasan tentang jenis buku atau frekuensi 'sering' dapat ditafsirkan berbeda oleh setiap responden.

b. Kondisi Responden

Reliabilitas juga dipengaruhi oleh keadaan fisik dan psikologis responden saat menjawab instrumen (Febriana et al, 2023). Jika responden sedang lelah, tidak fokus, tergesa-gesa, atau mengalami tekanan emosional, maka jawabannya cenderung tidak stabil. Contoh: siswa yang sedang cemas atau lapar ketika mengerjakan tes mungkin tidak menunjukkan kemampuan sebenarnya, sehingga hasilnya tidak konsisten jika tes diulang pada waktu lain.

c. Kondisi Lingkungan Pengukuran

Lingkungan tempat pengukuran dilakukan dapat memengaruhi konsistensi hasil. Faktor-faktor seperti kebisingan, suhu ruangan, penerangan, gangguan dari luar, atau kenyamanan tempat duduk dapat membuat responden tidak fokus atau cepat lelah. Contoh: tes yang dilakukan di ruangan bising atau panas mungkin menghasilkan skor berbeda dibanding tes di ruangan yang tenang dan nyaman.

d. Panjang Tes atau Jumlah Item

Semakin banyak jumlah item dalam instrumen, biasanya semakin tinggi reliabilitasnya, karena variasi data menjadi lebih luas dan kesalahan acak dapat saling meniadakan. Namun, terlalu banyak item juga bisa menurunkan kualitas jika membuat responden bosan atau kelelahan. Contoh: tes dengan 10 soal mungkin kurang stabil dibanding tes dengan 50 soal yang mengukur aspek yang sama, asalkan setiap item tetap relevan dan jelas.

e. Interval Waktu Pengukuran

Reliabilitas test-retest (ulang-uji) sangat bergantung pada jarak waktu antara dua kali pengukuran (Forester et al, 2024). Jika interval waktu terlalu pendek, responden mungkin masih mengingat jawaban sebelumnya (menyebabkan reliabilitas tampak tinggi secara semu). Sebaliknya, jika terlalu lama, kondisi responden atau faktor eksternal bisa berubah (menurunkan konsistensi). Contoh: mengulang tes motivasi belajar setelah 2 hari mungkin terlalu cepat, tapi setelah 6 bulan mungkin terlalu lama karena motivasi bisa berubah secara alami.

f. Kualitas Penilai atau Pengamat

Dalam instrumen yang melibatkan penilaian subjektif, seperti observasi, wawancara, atau penilaian kinerja, reliabilitas sangat tergantung pada konsistensi antar penilai. Penilai yang kurang terlatih atau tidak menggunakan pedoman penilaian yang sama dapat memberikan skor berbeda terhadap responden yang sama. Contoh: dua guru yang menilai presentasi siswa tanpa rubrik yang jelas bisa memberikan nilai yang berbeda, sehingga reliabilitas antar-penilai menjadi rendah.

KESIMPULAN

Reliabilitas merupakan indikator penting dalam menentukan sejauh mana suatu alat ukur dapat menghasilkan hasil yang konsisten dan dapat dipercaya. Jenis-jenis reliabilitas meliputi test-retest, parallel forms, internal consistency, split-half, dan inter-rater reliability. Pengukuran reliabilitas dapat dilakukan dengan berbagai metode statistik seperti korelasi Pearson, Spearman-Brown, Cronbach's Alpha, dan Cohen's Kappa. Faktor-faktor seperti kejelasan item, kondisi responden, lingkungan, panjang tes, interval waktu, dan kualitas penilai sangat menentukan tinggi rendahnya reliabilitas.

DAFTAR PUSTAKA

Apikunto, S. (2019). *Prosedur Penelitian: Suatu Pendekatan Praktik*. Rineka Cipta.

Azwar, S. (2017). *Reliabilitas dan Validitas*. Pustaka Pelajar.

Creswell, J. W. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE Publications.

Febriana, A., Pefbrianti, D., Ifansyah, M. N., & Lestari, D. H. (2023). Validitas dan Reliabilitas Instrumen Kualitas Hidup Lansia dengan Hipertensi. *Media Publikasi Promosi Kesehatan Indonesia (MPPKI)*, *6*(7), 1401–1406.

Forester, B. J., Khater, A. I. A., Afgani, M. W., & Isnaini, M. (2024). Penelitian Kuantitatif: Uji Reliabilitas. *Edu Society: Jurnal Pendidikan, Ilmu Sosial, dan Pengabdian Kepada Masyarakat*, *4*(3), 1812–1820. https://doi.org/10.56832/edu.v4i3.577

Ghozali, I. (2021). *Aplikasi Analisis Multivariate dengan Program IBM SPSS 26*. Badan Penerbit Universitas Diponegoro.

Khaddafi, M., Nasution, A. F., Angkat, N. A., & Pitriyani, S. (2025). Peran Metodologi Penelitian dalam Menentukan Kualitas Hasil Penelitian Ilmiah. *Jurnal Intelek Insan Cendikia*, *2*(7), 13250–13255.

Murti, B. (2011, May). *Validitas dan Reliabilitas Pengukuran* [Materi matrikulasi program studi doktoral, Fakultas Kedokteran]. Universitas Sebelas Maret.

https://rossisanusi.wordpress.com/wp-content/uploads/2013/09/validitas-relibilitas-pengukuran.pdf

Ridha, A. R., & Al Ghifari, F. (2025). Reliabilitas Alat Ukur, Jenis-Jenis, Cara Pengukuran dan Faktor-Faktor. *Inovasi: Jurnal Ilmiah Pengembangan Pendidikan*, *4*(1), 121–131. https://ejournal.lpipb.com/index.php/inovasi/article/view/82

Subhaktiyasa, P. G. (2024). Evaluasi Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif: Sebuah Studi Pustaka. *Journal of Education Research*, *5*(4), 5599–5609. https://doi.org/10.37985/jer.v5i4.1747

Sugiyono. (2018). *Metode Penelitian Pendidikan: Pendekatan Kuantitatif, Kualitatif, dan R&D*. Alfabeta.

Utami, Y. (2023). Uji Validitas dan Uji Reliabilitas Instrument Penilaian Kinerja Dosen. *Jurnal Sains dan Teknologi*, *4*(2), 21–24. https://doi.org/10.55338/saintek.v4i2.730