

Evaluasi Keabsahan Dan Keandalan Instrumen Penelitian

Muhammad Fakhri Khairi¹, MHD. Nanang Hidayat², Nafia Sismita³, Nurhayati⁴, Herlinda⁵

^{1,2,3,4,5}Universitas Islam Negeri Sultan Syarif Kasim Riau

Correspondence: fakhrykhairi9@gmail.com, nng240102@gmail.com,

nafiasismita961@gmail.com, nurhayatii251995@gmail.com, herlinda@gmail.com

Abstrak.Evaluasi keabsahan dan keandalan instrumen penelitian merupakan langkah fundamental dalam menjamin kualitas data dan validitas hasil penelitian. Penelitian ini bertujuan untuk menjelaskan prosedur serta teknik pengujian validitas dan reliabilitas instrumen yang digunakan dalam penelitian kuantitatif. Evaluasi dilakukan melalui tiga bentuk validitas utama, yaitu validitas isi, validitas konstruk, dan validitas kriteria, serta pengujian reliabilitas yang mencakup konsistensi internal, reliabilitas uji ulang (test-retest), dan reliabilitas antar-penilai (inter-rater). Proses pengujian dilakukan dengan melibatkan analisis statistik seperti korelasi item-total, analisis faktor eksploratori, dan perhitungan koefisien Cronbach's Alpha menggunakan perangkat lunak statistik SPSS, R, atau Jamovi. Hasil kajian menunjukkan bahwa instrumen yang dinyatakan valid dan reliabel memiliki kontribusi signifikan terhadap ketepatan interpretasi hasil penelitian dan peningkatan kredibilitas temuan ilmiah. Dengan demikian, proses evaluasi instrumen tidak hanya berperan sebagai tahap teknis, tetapi juga sebagai bentuk pengendalian mutu metodologis dalam penelitian ilmiah.

Kata kunci : Validitas, Reliabilitas, Instrumen Penelitian, Evaluasi, Kualitas Data.

Abstract.Evaluating the validity and reliability of research instruments is a fundamental step in ensuring the quality of data and the validity of research results. This study aims to explain the procedures and techniques for testing the validity and reliability of instruments used in quantitative research. Evaluation is carried out through three main forms of validity, namely content validity, construct validity, and criterion validity, as well as reliability testing which includes internal consistency, test-retest reliability, and inter-rater reliability. The testing process is carried out by involving statistical analysis such as item-total correlation, exploratory factor analysis, and calculation of Cronbach's Alpha coefficient using SPSS, R, or Jamovi statistical software. The results of the study show that instruments that are declared valid and reliable have a significant contribution to the accuracy of the interpretation of research results and increase the credibility of scientific findings. Thus, the instrument evaluation process not only plays a role as a technical stage, but also as a form of methodological quality control in scientific research.

Keywords : Validity, Reliability, Research Instruments, Evaluation, Data Quality

PENDAHULUAN

Dalam penelitian, penggunaan instrumen yang valid dan reliabel sangat penting untuk memastikan data yang diperoleh akurat dan dapat dipercaya. Instrumen yang tidak valid atau tidak reliabel berpotensi menghasilkan data yang menyesatkan sehingga kesimpulan penelitian menjadi tidak valid. Oleh karena itu, evaluasi keabsahan dan keandalan instrumen menjadi tahap krusial sebelum instrumen digunakan secara luas. Validitas memastikan bahwa instrumen benar-benar mengukur konstruk yang dimaksud, sedangkan reliabilitas menjamin konsistensi hasil pengukuran dari waktu ke waktu maupun antar item. Dengan evaluasi yang tepat, kualitas data penelitian dapat ditingkatkan sehingga hasil dan rekomendasi yang diberikan lebih dapat dipertanggungjawabkan.

LANDASAN TEORI

Konsep Validitas Instrumen Penelitian

Validitas merupakan derajat ketepatan suatu instrumen dalam mengukur apa yang seharusnya diukur. Menurut Taherdoost (2018), validitas memastikan bahwa hasil pengukuran mencerminkan konstruk konseptual yang ditetapkan, bukan aspek lain yang tidak relevan. Tiga bentuk utama validitas yang sering digunakan adalah validitas isi (content validity), validitas konstruk (construct validity), dan validitas kriteria (criterion-related validity).

Validitas isi menilai sejauh mana butir instrumen mencakup keseluruhan aspek konstruk teoretis. Proses ini biasanya dilakukan melalui penilaian ahli atau expert judgment untuk memastikan

keterwakilan aspek yang diukur (Bolarinwa, 2018). Sementara itu, validitas konstruk berfokus pada sejauh mana item dalam instrumen berkorelasi secara konsisten dengan konstruk teoretis yang mendasari (Hair et al., 2019). Validitas kriteria menguji sejauh mana hasil instrumen berkorelasi dengan ukuran eksternal yang telah terstandar atau terbukti valid sebelumnya (Heale & Twycross, 2018).

Konsep Reliabilitas Instrumen Penelitian

Reliabilitas menunjukkan konsistensi hasil pengukuran ketika instrumen digunakan secara berulang dalam kondisi yang sama. Menurut Taber (2018), instrumen yang reliabel menghasilkan data yang stabil dari waktu ke waktu dan antar item. Reliabilitas dapat dibedakan menjadi reliabilitas internal (internal consistency), reliabilitas uji ulang (test-retest reliability), dan reliabilitas antar-penilai (inter-rater reliability).

Reliabilitas internal sering diukur menggunakan koefisien Cronbach's Alpha, yang menunjukkan derajat konsistensi antarbutir dalam satu konstruk (Gliem & Gliem, 2021). Nilai Cronbach's Alpha di atas 0,70 dianggap memadai untuk penelitian sosial, sementara nilai di atas 0,80 menunjukkan reliabilitas tinggi. Test-retest reliability digunakan untuk mengukur stabilitas temporal hasil pengukuran, sedangkan inter-rater reliability digunakan untuk menilai konsistensi antar pengamat atau penilai dalam data kualitatif (Koo & Li, 2016).

Hubungan Validitas dan Reliabilitas terhadap Kualitas Data

Validitas dan reliabilitas saling berkaitan erat dalam menentukan kualitas instrumen penelitian. Instrumen yang reliabel belum tentu valid, tetapi instrumen yang valid haruslah reliabel. Menurut Taherdoost (2018), kedua konsep ini merupakan syarat utama agar data yang dikumpulkan akurat, kredibel, dan dapat dipertanggungjawabkan secara ilmiah. Dalam penelitian sosial dan pendidikan, validitas dan reliabilitas yang tinggi meningkatkan kredibilitas hasil serta memperkuat dasar pengambilan keputusan kebijakan berbasis data (Heale & Twycross, 2018; Ardiansyah & Jailani, 2023).

Selain itu, perkembangan perangkat lunak statistik seperti SPSS, R, dan Jamovi telah memudahkan proses pengujian validitas dan reliabilitas. Analisis seperti Exploratory Factor Analysis (EFA) dan korelasi item-total digunakan untuk mengevaluasi keterpaduan butir terhadap konstruk (Maulana, 2022). Dengan demikian, evaluasi instrumen penelitian bukan sekadar prosedur teknis, melainkan langkah metodologis penting untuk menjamin mutu dan integritas hasil penelitian ilmiah.

METODE PENELITIAN

Proses evaluasi instrumen penelitian dilakukan melalui beberapa tahap utama. Pertama, validitas isi diukur dengan meminta penilaian dari ahli untuk memastikan bahwa butir-butir instrumen mencakup seluruh aspek yang relevan dengan konstruk yang diteliti. Selanjutnya, validitas konstruk diuji menggunakan analisis faktor eksploratori dan korelasi item-total untuk melihat keterkaitan antar butir dengan keseluruhan skor. Validitas kriteria diuji dengan membandingkan hasil instrumen dengan instrumen standar melalui korelasi Pearson.

Untuk mengukur reliabilitas, digunakan uji konsistensi internal dengan menghitung nilai Cronbach's Alpha dan analisis split-half. Selain itu, reliabilitas uji ulang dilakukan dengan mengadministrasikan instrumen kepada responden yang sama pada dua waktu berbeda, kemudian menghitung koefisien korelasi hasil pengukuran kedua. Data hasil uji coba instrumen dianalisis menggunakan perangkat lunak statistik guna memastikan instrumen memenuhi standar validitas dan reliabilitas sebelum digunakan dalam penelitian utama.

HASIL DAN PEMBAHASAN

Validitas menunjukkan sejauh mana instrumen mengukur apa yang seharusnya diukur. Ada beberapa jenis validitas:

Validitas Isi (*Content Validity*)

Menilai apakah butir-butir instrumen mencakup semua aspek konstruk yang diukur merupakan bagian penting dari validitas isi (*content validity*). Langkah ini bertujuan untuk memastikan

bahwa seluruh dimensi atau aspek dari konstruk yang ingin diukur benar-benar terwakili dalam butir-butir instrumen. Penilaian ini biasanya dilakukan dengan mengacu pada definisi teoretis konstruk, telaah literatur, dan masukan dari para ahli di bidang terkait.

Instrumen dianggap memiliki cakupan yang baik jika setiap aspek utama konstruk terwakili oleh satu atau lebih butir pertanyaan, sehingga tidak ada aspek penting yang terlewatkan dan tidak ada butir yang mengukur hal di luar konstruk.

Misalnya, sebuah instrumen dikembangkan untuk mengukur *kepuasan kerja* karyawan. Berdasarkan literatur, konstruk kepuasan kerja terdiri dari beberapa aspek seperti kepuasan terhadap gaji, hubungan dengan atasan, lingkungan kerja, kesempatan pengembangan karier, dan keseimbangan kerja-hidup. Dalam proses penilaian isi, tim pengembang instrumen harus memastikan bahwa ada butir-butir yang secara eksplisit mengukur masing-masing aspek tersebut.

Jika ternyata dalam instrumen tersebut hanya terdapat butir mengenai gaji dan hubungan dengan atasan, tetapi tidak ada butir yang mengukur aspek lingkungan kerja atau keseimbangan kerja-hidup, maka dapat disimpulkan bahwa instrumen tersebut belum mencakup seluruh aspek konstruk kepuasan kerja secara komprehensif. Dengan demikian, revisi instrumen perlu dilakukan agar semua aspek konstruk terwakili dan hasil pengukuran menjadi lebih valid.

Validitas Konstruk (*Construct Validity*)

Mengukur apakah instrumen benar-benar mencerminkan konstruk teoretis yang ingin diteliti merupakan bagian dari evaluasi validitas konstruk (*construct validity*). Validitas konstruk menilai sejauh mana suatu instrumen mampu mengukur konstruk yang dimaksud, sesuai dengan definisi dan kerangka teoretis yang mendasarinya. Penilaian ini melibatkan proses yang lebih mendalam dibandingkan validitas isi, karena tidak hanya memeriksa cakupan butir, tetapi juga hubungan antar butir, keterkaitan dengan konstruk lain, dan kesesuaian dengan teori yang ada. Proses ini dapat melibatkan analisis statistik seperti *factor analysis*, serta uji hipotesis teoretis seperti konvergen dan diskriminan validitas.

Contoh: Jika sebuah peneliti mengembangkan instrumen untuk mengukur *self-efficacy* (keyakinan diri seseorang dalam menyelesaikan tugas), maka ia harus memastikan bahwa semua butir dalam instrumen tersebut benar-benar mencerminkan konstruk *self-efficacy* sebagaimana didefinisikan dalam teori Bandura. Misalnya, butir seperti “Saya yakin dapat menyelesaikan tugas sulit jika saya berusaha keras” mencerminkan aspek inti dari *self-efficacy*.

Namun, jika terdapat butir seperti “Saya senang bekerja dalam kelompok,” maka butir tersebut lebih mencerminkan preferensi sosial daripada *self-efficacy*, dan tidak sesuai dengan konstruk teoretis yang ingin diukur. Oleh karena itu, instrumen perlu dikaji secara kritis agar hanya mencakup indikator yang benar-benar sesuai dengan konstruk, memastikan bahwa pengukuran yang dilakukan valid secara teoretis.

Validitas Kriteria (*Criterion Validity*)

Mengukur sejauh mana skor pada instrumen berkorelasi dengan standar atau hasil eksternal yang sudah diketahui merupakan bagian dari **validitas kriteria** (*criterion-related validity*). Validitas ini menilai seberapa baik skor dari suatu instrumen dapat memprediksi atau berkaitan dengan suatu kriteria eksternal yang relevan dan telah diakui. Validitas kriteria terbagi menjadi dua jenis utama: **validitas prediktif** (jika skor instrumen digunakan untuk memprediksi hasil di masa depan) dan **validitas konkuren** (jika skor dibandingkan dengan hasil yang diperoleh pada waktu yang sama).

Contoh: Misalnya, sebuah lembaga mengembangkan instrumen untuk mengukur *kemampuan akademik calon mahasiswa baru*. Untuk menguji validitas kriteria dari instrumen ini, peneliti dapat membandingkan skor instrumen dengan IPK tahun pertama mahasiswa (validitas prediktif) atau dengan hasil tes standar nasional yang telah diakui (validitas konkuren). Jika ditemukan korelasi yang tinggi dan signifikan antara skor instrumen dengan

IPK atau hasil tes tersebut, maka dapat disimpulkan bahwa instrumen memiliki validitas kriteria yang baik. Sebaliknya, jika tidak ada korelasi atau korelasi sangat rendah, maka efektivitas instrumen dalam mencerminkan kriteria eksternal yang diharapkan perlu dipertanyakan dan kemungkinan harus direvisi.

Reliabilitas (Keandalan) Instrumen

Reliabilitas menunjukkan konsistensi hasil pengukuran jika instrumen digunakan berulang kali dalam kondisi yang sama.

Reliabilitas Internal

Menunjukkan konsistensi antar item dalam satu instrumen merupakan bagian dari reliabilitas internal atau konsistensi internal (*internal consistency*). Reliabilitas ini menunjukkan sejauh mana butir-butir (item) dalam satu instrumen saling berkaitan dan secara konsisten mengukur konstruk yang sama. Dengan kata lain, jika semua item dalam suatu instrumen memang dimaksudkan untuk mengukur aspek yang sama, maka respons individu terhadap item-item tersebut seharusnya menunjukkan pola yang konsisten.

Untuk mengukur konsistensi internal, metode yang paling umum digunakan adalah koefisien Cronbach's Alpha (α). Nilai alpha berkisar antara 0 hingga 1, dengan nilai yang lebih tinggi menunjukkan konsistensi internal yang lebih baik. Secara umum, nilai $\alpha \geq 0,70$ dianggap memadai untuk penelitian eksploratif, dan $\geq 0,80$ untuk penelitian yang lebih aplikatif atau pengambilan keputusan penting.

Contoh: sebuah instrumen terdiri dari 10 butir yang dirancang untuk mengukur *stres akademik* mahasiswa. Jika seseorang memberikan jawaban yang tinggi pada satu butir seperti "Saya merasa cemas ketika menghadapi ujian," maka seharusnya ia juga cenderung memberikan jawaban tinggi pada butir lain yang serupa seperti "Saya sulit tidur karena memikirkan tugas kuliah." Jika korelasi antar butir tinggi dan koefisien Cronbach's Alpha mendekati atau melebihi 0,80, maka dapat disimpulkan bahwa instrumen tersebut memiliki konsistensi internal yang baik. Jika tidak, maka perlu dilakukan evaluasi ulang terhadap item-item yang lemah atau tidak relevan.

Reliabilitas Uji Ulang (*Test-Retest Reliability*)

Mengukur konsistensi hasil dari waktu ke waktu merupakan bagian dari reliabilitas test-retest. Reliabilitas ini menilai sejauh mana suatu instrumen menghasilkan skor yang stabil dan konsisten ketika digunakan pada individu yang sama dalam dua waktu yang berbeda, dengan asumsi bahwa konstruk yang diukur tidak berubah secara signifikan di antara dua pengukuran tersebut.

Prosedur test-retest melibatkan pemberian instrumen kepada responden pada dua waktu yang berbeda, biasanya dipisahkan oleh jeda waktu yang cukup (misalnya satu atau dua minggu), dan kemudian menghitung korelasi antara dua set skor. Korelasi yang tinggi menunjukkan bahwa instrumen memiliki reliabilitas temporal yang baik, artinya hasil pengukuran bersifat stabil dan tidak dipengaruhi oleh faktor situasional sesaat.

Contoh: seorang peneliti mengembangkan instrumen untuk mengukur self-esteem (harga diri). Ia memberikan kuesioner kepada 50 mahasiswa pada minggu pertama, lalu mengulangnya pada minggu ketiga. Jika skor yang diperoleh oleh masing-masing mahasiswa pada dua waktu tersebut memiliki korelasi yang tinggi (misalnya $r \geq 0,80$), maka instrumen tersebut dianggap memiliki reliabilitas test-retest yang baik. Namun, jika korelasi rendah, hal ini mengindikasikan bahwa hasil pengukuran tidak konsisten dari waktu ke waktu, sehingga perlu ditinjau kembali kemungkinan adanya masalah pada instrumen atau faktor luar yang memengaruhi jawaban.

Reliabilitas Antar Penilai (*Inter-Rater Reliability*)

Digunakan dalam data kualitatif atau pengamatan mengacu pada reliabilitas antar-rater (*inter-rater reliability*). Reliabilitas ini penting ketika data dikumpulkan melalui observasi, wawancara, atau penilaian subjektif, di mana lebih dari satu penilai (rater) terlibat. Tujuannya adalah untuk mengukur sejauh mana dua atau lebih pengamat memberikan penilaian yang konsisten terhadap objek atau peristiwa yang sama.

Reliabilitas antar-rater memastikan bahwa hasil pengamatan atau interpretasi tidak tergantung pada siapa yang melakukan penilaian, tetapi mencerminkan kenyataan yang objektif atau konsisten. Untuk mengukurnya, biasanya digunakan koefisien seperti Cohen's Kappa (untuk data kategorikal), ICC (Intraclass Correlation Coefficient) untuk data interval atau rasio, atau persentase kesepakatan antar-rater.

Contoh: dua guru mengamati perilaku siswa dalam kelas untuk menilai tingkat partisipasi aktif. Mereka masing-masing mencatat apakah siswa menunjukkan perilaku aktif atau pasif dalam diskusi. Jika kedua guru memberikan penilaian yang serupa untuk sebagian besar siswa, maka reliabilitas antar-rater tinggi. Namun, jika penilaian mereka sering berbeda untuk siswa yang sama, berarti reliabilitas antar-rater rendah, dan mungkin diperlukan pelatihan tambahan atau pedoman penilaian yang lebih jelas untuk meningkatkan konsistensi antar penilai.

Langkah-langkah Evaluasi Praktis

1).Melakukan uji coba instrumen (pilot test) kepada sampel kecil yang mewakili populasi penelitian adalah langkah awal yang penting dalam proses pengembangan instrumen. Tujuan pilot test adalah untuk menguji kelayakan, kejelasan, dan fungsi butir-butir instrumen sebelum digunakan secara luas. Dengan pilot test, peneliti dapat mengidentifikasi masalah teknis, seperti butir yang ambigu, sulit dipahami, atau tidak relevan, serta mengevaluasi reliabilitas dan validitas awal instrumen.

Pilot test biasanya dilakukan pada sampel kecil yang karakteristiknya mirip dengan populasi sasaran agar hasilnya dapat menggambarkan bagaimana instrumen akan bekerja pada populasi sebenarnya. Setelah pilot test, peneliti dapat melakukan revisi berdasarkan masukan dan analisis data, seperti memperbaiki bahasa butir, menyingkirkan item yang kurang valid, atau menambah butir yang diperlukan.

Contoh peneliti ingin mengembangkan kuesioner untuk mengukur motivasi belajar siswa SMA. Sebelum menggunakan kuesioner pada seluruh sampel penelitian, ia melakukan pilot test pada 30 siswa dari sekolah yang sama. Dari hasil pilot test tersebut, ditemukan bahwa beberapa butir kurang jelas dan siswa memberikan jawaban yang tidak konsisten pada bagian tertentu. Peneliti kemudian merevisi butir tersebut agar lebih mudah dipahami dan memastikan instrumen lebih valid serta reliabel untuk pengukuran motivasi belajar.

2). Setelah melakukan uji coba instrumen (pilot test), langkah selanjutnya adalah analisis data hasil uji coba menggunakan software statistik seperti SPSS, R, atau Jamovi. Analisis ini bertujuan untuk menilai kualitas butir, reliabilitas, dan validitas instrumen secara kuantitatif sebelum digunakan pada penelitian utama.

Beberapa analisis umum yang dilakukan meliputi:

- a. Analisis butir (item analysis): Melihat statistik deskriptif tiap butir, seperti rata-rata, standar deviasi, serta item-total correlation untuk mengetahui seberapa baik butir tersebut berkorelasi dengan total skor instrumen.
- b. Uji reliabilitas (misalnya Cronbach's Alpha): Mengukur konsistensi internal instrumen secara keseluruhan.
- c. Analisis faktor eksploratori (Exploratory Factor Analysis/EFA): Mengidentifikasi struktur faktor dalam instrumen dan memastikan butir-butir mengelompok sesuai konstruk yang diukur.
- d. Uji validitas konstruk awal: Misalnya dengan melihat korelasi antar butir dan faktor.

Dengan menggunakan software seperti SPSS, R, atau Jamovi, proses ini bisa dilakukan secara cepat dan hasilnya disajikan dalam bentuk tabel dan grafik yang memudahkan interpretasi.

Contoh: Seorang peneliti memasukkan data hasil pilot test kuesioner motivasi belajar ke SPSS. Setelah menjalankan analisis reliabilitas, didapatkan nilai Cronbach's Alpha sebesar 0,82, yang menunjukkan konsistensi internal yang baik. Selanjutnya, peneliti melakukan analisis faktor eksploratori untuk memastikan butir-butir terkonsentrasi pada beberapa faktor yang sesuai teori motivasi belajar. Hasil analisis membantu peneliti memutuskan butir mana yang perlu direvisi atau dihapus sebelum instrumen digunakan lebih luas.

1. Kaji validitas isi dengan meminta masukan dari pakar

Mengkaji validitas isi dengan meminta masukan dari pakar adalah langkah krusial dalam pengembangan instrumen untuk memastikan bahwa butir-butir yang disusun benar-benar mewakili semua aspek konstruk yang ingin diukur secara komprehensif dan sesuai dengan teori. Proses ini biasanya dilakukan setelah draft instrumen selesai disusun.

Caranya adalah dengan menyerahkan instrumen tersebut kepada beberapa ahli atau pakar di bidang yang relevan misalnya akademisi, praktisi, atau peneliti yang berpengalaman untuk memberikan penilaian dan masukan. Para pakar akan mengevaluasi setiap butir dari segi relevansi, kejelasan, kelayakan, dan cakupan aspek konstruk. Masukan dari pakar ini sangat berguna untuk memperbaiki atau menyesuaikan butir agar instrumen menjadi lebih valid secara isi.

Contoh: sebuah instrumen untuk mengukur *kepuasan pelanggan* telah disusun. Peneliti kemudian mengirimkan draft instrumen tersebut ke tiga pakar pemasaran dan psikometri untuk ditinjau. Para pakar memberikan komentar seperti "Butir nomor 4 kurang spesifik dan perlu diperjelas," atau "Aspek pelayanan purna jual kurang terwakili, sebaiknya ditambahkan beberapa butir." Berdasarkan masukan ini, peneliti merevisi instrumen sehingga butir-butirnya menjadi lebih relevan dan lengkap dalam mengukur kepuasan pelanggan sesuai dengan teori dan kebutuhan penelitian.

2. Lakukan analisis reliabilitas, terutama Cronbach's Alpha

Melakukan analisis reliabilitas, terutama menggunakan Cronbach's Alpha, adalah langkah penting untuk mengukur konsistensi internal suatu instrumen. Cronbach's Alpha menunjukkan sejauh mana butir-butir dalam satu instrumen saling berkorelasi dan bersama-sama mengukur konstruk yang sama secara konsisten.

Berikut langkah-langkah umum melakukan analisis Cronbach's Alpha:

- a. Input data hasil uji coba instrumen ke software statistik seperti SPSS, R, atau Jamovi.
- b. Pilih menu analisis reliabilitas dan masukkan semua butir instrumen yang akan diuji.
- c. Jalankan analisis dan lihat nilai Cronbach's Alpha yang dihasilkan.

Interpretasi nilai alpha:

Nilai $\alpha \geq 0,70$ dianggap cukup baik untuk penelitian eksploratif.

Nilai $\alpha \geq 0,80$ menunjukkan reliabilitas yang baik.

Nilai $\alpha < 0,60$ menandakan konsistensi internal rendah dan perlu evaluasi ulang butir.

Selain nilai keseluruhan, perhatikan juga item-total correlation dan dampak penghapusan item terhadap nilai alpha, sehingga butir yang bermasalah bisa diidentifikasi dan diperbaiki atau dihapus.

Contoh dalam pengujian kuesioner motivasi belajar dengan 15 butir, peneliti menggunakan SPSS untuk menghitung Cronbach's Alpha dan mendapatkan nilai 0,85. Ini menunjukkan instrumen memiliki konsistensi internal yang baik. Namun, analisis item menunjukkan butir nomor 7 memiliki korelasi item-total yang rendah dan jika dihapus, nilai alpha naik menjadi 0,87. Peneliti memutuskan untuk merevisi atau menghapus butir tersebut agar instrumen semakin andal.

Revisi Instrumen Berdasarkan Hasil Analisis Untuk Memperbaiki Butir Yang Lemah

Revisi instrumen berdasarkan hasil analisis merupakan langkah penting untuk meningkatkan kualitas dan keandalan instrumen. Setelah melakukan uji coba dan analisis seperti reliabilitas dan validitas, peneliti harus mengevaluasi butir-butir yang menunjukkan performa kurang baik misalnya butir dengan *item-total correlation* rendah, butir yang mengurangi nilai Cronbach's Alpha jika dimasukkan, atau butir yang tidak relevan berdasarkan masukan pakar.

Dalam proses revisi, butir-butir yang lemah bisa diperbaiki dengan cara memperjelas bahasa, menyederhanakan kalimat, menghilangkan ambiguitas, atau mengganti kata-kata yang sulit dipahami. Jika butir tersebut memang tidak sesuai atau tidak penting, maka bisa diputuskan untuk dihapus agar instrumen lebih fokus dan reliabel. Revisi ini juga bisa mencakup penambahan butir baru untuk menutupi aspek konstruk yang kurang terwakili.

Contoh: dari hasil analisis reliabilitas kuesioner motivasi belajar, ditemukan bahwa butir nomor 5 dan 12 memiliki korelasi rendah dengan total skor dan mengurangi nilai Cronbach's Alpha. Peneliti kemudian mengkaji ulang butir tersebut dan menyadari bahwa bahasa yang digunakan terlalu teknis dan sulit dipahami responden. Setelah revisi dengan bahasa yang lebih sederhana dan jelas, kedua butir tersebut diuji kembali pada pilot test berikutnya. Bila masih bermasalah, butir tersebut diputuskan untuk dihapus agar instrumen menjadi lebih valid dan reliabel.

KESIMPULAN

Evaluasi keabsahan dan keandalan instrumen sangat penting untuk memastikan bahwa alat ukur yang digunakan benar-benar sesuai dengan tujuan penelitian dan menghasilkan data yang konsisten. Dengan melakukan pengujian validitas (isi, konstruk, dan kriteria) serta reliabilitas (internal dan uji ulang), instrumen dapat dioptimalkan sehingga mendukung validitas hasil penelitian secara keseluruhan. Oleh karena itu, proses evaluasi ini menjadi langkah krusial dalam pengembangan instrumen penelitian yang berkualitas dan dapat dipercaya.

DAFTAR PUSTAKA

- Anshari, Muhammad Isa, and others. (2024). Analisis Validitas dan Reliabilitas Butir Soal Sumatif Akhir Semester Ganjil Mata Pelajaran PAI', *Edukatif: Jurnal Ilmu Pendidikan*, 6.1, pp. 964–75, doi:10.31004/edukatif.v6i1.5931
- Ardiansyah, Risnita, and M. Syahrani Jailani. (2023). Teknik Pengumpulan Data Dan Instrumen Penelitian Ilmiah Pendidikan Pada Pendekatan Kualitatif dan Kuantitatif', *Jurnal IHSAN : Jurnal Pendidikan Islam*, 1.2, pp. 1–9, doi:10.61104/ihsan.v1i2.57
- Bolarinwa, O. A. (2018). *Principles and methods of validity and reliability testing of questionnaires used in social and health science researches*. Nigerian Postgraduate Medical Journal, 25(2), 67–75. https://doi.org/10.4103/npmj.npmj_13_18
- Fauziyah, Anisa, Zahro As Sakinah, and Dase Erwin Juansah. (2023). Instrumen Tes Dan Non Tes Pada Penelitian', 08.
- Gliem, J. A., & Gliem, R. R. (2021). *Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales*. Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education, 82–88. <https://scholarworks.iupui.edu/handle/1805/344>
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2019). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed.). SAGE Publications.
- Heale, R., & Twycross, A. (2018). *Validity and reliability in quantitative studies*. Evidence-Based Nursing, 18(3), 66–67. <https://doi.org/10.1136/eb-2015-102129>
- Mahmud, Dira Kaimel, and Yama Pradhana Sumbodo, 'Komunikasi Keluarga Dalam Kemampuan Remaja Mengambil Keputusan Memilih Sekolah Lanjutan Tingkat Atas',

1.1 (2025)

- Maulana, Andi. (2022). Analisis Validitas, Reliabilitas, dan Kelayakan Instrumen Penilaian Rasa Percaya Diri Siswa. *Jurnal Kualita Pendidikan*, 3.3, pp. 133–39, doi:10.51651/jkp.v3i3.331
- Mekarisce, Arnild Augina. (2020). Teknik Pemeriksaan Keabsahan Data pada Penelitian Kualitatif di Bidang Kesehatan Masyarakat', *Jurnal Ilmiah Kesehatan Masyarakat : Media Komunikasi Komunitas Kesehatan Masyarakat*, 12, pp. 145–51, doi:10.52022/jikm.v12i3.102
- Ono, Sugi. (2020). Uji Validitas dan Reliabilitas Alat Ukur SG Posture Evaluation', *Jurnal Keterampilan Fisik*, 5.1, pp. 55–61, doi:10.37341/jkf.v5i1.167
- Saadah, Muftahatus, Yoga Catur Prasetyo, and Gismina Tri Rahmayati. (2022). Strategi Dalam Menjaga Keabsahan Data Pada Penelitian Kualitatif', *Al-'Adad : Jurnal Tadris Matematika*, 1.2, pp. 54–64, doi:10.24260/add.v1i2.1113
- Sanaky, Musrifah Mardiani, La Moh Saleh, and Henriette D Titaley. (2021). Analisis Faktor-Faktor Penyebab Keterlambatan Pada Proyek Pembangunan Gedung Asrama Man 1 Tulehu Maluku Tengah', 11.1
- Sholikhah, Ananda Amilus, and others. (2024). Analisis Tingkat Kepuasan Mahasiswa Terhadap Penggunaan Quizizz pada Mata Kuliah Data Mining', *Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)*, 7.4, pp. 481–89, doi:10.32672/jnkti.v7i4.7675
- Koo, T. K., & Li, M. Y. (2016). *A guideline of selecting and reporting intraclass correlation coefficients for reliability research*. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Maulana, A. (2022). Analisis validitas, reliabilitas, dan kelayakan instrumen penilaian rasa percaya diri siswa. *Jurnal Kualita Pendidikan*, 3(3), 133–139. <https://doi.org/10.51651/jkp.v3i3.331>
- Taber, K. S. (2018). *The use of Cronbach's alpha when developing and reporting research instruments in science education*. *Research in Science Education*, 48(6), 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Taherdoost, H. (2018). *Validity and reliability of the research instrument: How to test the validation of a questionnaire/survey in a research*. *International Journal of Academic Research in Management*, 7(3), 28–36. <https://hal.science/hal-02546799>