JURNAL ILMU PENDIDIKAN & SOSIAL
VOLUME 04, NOMOR 02, TAHUN 2026
e-ISSN: 3025-9231

Validitas Instrumen Asesmen sebagai Penjamin Mutu Evaluasi Pembelajaran Pendidikan Agama Islam: Kajian Literatur

Shoviyyatuz Zahro[1], Mila Ulfa[2], Dina Wulandari[3], Fahmi Ali[4], Annisa Amalia Rizqiyah[5], Mutmainnatuz Zahro[6], Muhammad Miftah[7]
[1,2,3,4,5,6,7]Pendidikan Agama Islam, Tarbiyah, Universitas Islam Negeri Sunan Kudus, Indonesia
Email: shoviyya@mhs.uinsuku.ac.id[1*], milaulfa@mhs.uinsuku.ac.id[2], dinawulanda@mhs.uinsuku.ac.id[3], fahmiali@mhs.uinsuku.ac.id[4], annisamalia@mhs.uinsuku.ac.id[5], mutmainna@mhs.uinsuku.ac.id[6], muhammadmiftah@iainkudus.ac.id[7]

Abstrak

Validitas instrumen asesmen merupakan aspek penting dalam menjamin mutu evaluasi pembelajaran Pendidikan Agama Islam (PAI). Namun, masih ditemukan instrumen evaluasi yang disusun tanpa memperhatikan kesesuaian indikator, konstruk, dan tujuan pengukuran sehingga berpotensi menghasilkan data yang kurang akurat. Penelitian ini bertujuan mengkaji konsep, jenis, langkah-langkah pengujian, dan cara menentukan validitas instrumen asesmen melalui studi literatur. Metode yang digunakan adalah penelitian kepustakaan (library research) dengan pendekatan kualitatif deskriptif. Data diperoleh dari berbagai buku, jurnal ilmiah, dan dokumen yang relevan, kemudian dianalisis secara deskriptif. Hasil kajian menunjukkan bahwa validitas instrumen mencakup validitas logis dan validitas empiris yang saling melengkapi dalam menjamin ketepatan pengukuran. Pengujian validitas dilakukan secara sistematis melalui perumusan konstruk, penyusunan kisi-kisi, validasi ahli, uji coba instrumen, serta analisis statistik. Instrumen yang valid mampu menghasilkan data yang akurat, kredibel, dan mendukung pengambilan keputusan pendidikan secara tepat.

Kata kunci: Asesmen PAI, evaluasi pembelajaran, validitas instrumen

Abstract

The validity of assessment instruments is a crucial aspect in ensuring the quality of learning evaluations in Islamic Religious Education (PAI). However, there are still evaluation instruments that are developed without considering the alignment of indicators, constructs, and measurement objectives, which may result in inaccurate data. This study aims to examine the concepts, types, testing procedures, and methods for determining the validity of assessment instruments through a literature review. The method used is library research with a descriptive qualitative approach. Data were obtained from various books, scientific journals, and relevant documents, and then analyzed descriptively. The results of the study indicate that the validity of an instrument encompasses logical validity and empirical validity, which complement each other in ensuring the accuracy of measurements. Validity testing was conducted systematically through construct definition, development of a test blueprint, expert validation, instrument pilot testing, and statistical analysis. A valid instrument is capable of producing data that is accurate, credible, and supports sound educational decision-making.

Keywords: *Islamic Education Assessment, learning evaluation, instrument validity.*

PENDAHULUAN

Validitas mengacu pada sejauh mana suatu instrumen pengukuran secara akurat menangkap konstruk atau variabel yang dimaksudkan untuk diukur,

Copyright © 2026 Shoviyyatuz Zahro, et al.
(cc) BY

sehingga memastikan bahwa data yang dihasilkan benar-benar mencerminkan fenomena yang diteliti (Buka dkk., 2025). Dalam penelitian, instrumen memainkan peran sentral dalam mengumpulkan informasi yang relevan dan andal yang diperlukan untuk menjawab pertanyaan penelitian dan mengevaluasi hipotesis yang diajukan (Sari Siagian & Meyniar Albina, 2025). Alat ukur dengan tingkat validitas yang tinggi dianggap mampu menghasilkan pengukuran yang bermakna dan dapat dipercaya, sedangkan alat ukur dengan validitas rendah mungkin gagal merepresentasikan konstruk yang dimaksud secara memadai (Maulana, 2022). Selain akurasi, validitas juga mencakup ketepatan alat ukur dalam mengukur variabel tertentu (Sanaky dkk., 2021a). Oleh karena itu, instrumen yang valid tidak hanya mencerminkan kondisi aktual, tetapi juga memberikan representasi yang tepat dari data yang diukur (Efendi, 2019). Dalam hal ini, pengujian validitas merupakan tahap penting dalam pengembangan instrumen, karena kualitas temuan penelitian dan keandalan kesimpulan yang diambil sangat bergantung pada validitas data yang dikumpulkan.

Dalam proses pembelajaran, evaluasi menjadi komponen penting untuk mengetahui tingkat ketercapaian tujuan pembelajaran dan perkembangan peserta didik (Idrus L, 2019). Berbagai lembaga pendidikan menggunakan beragam instrumen evaluasi, seperti tes, angket, observasi, dan penilaian kinerja. Namun, permasalahan yang sering muncul adalah Masih banyak guru yang memandang evaluasi sebatas kegiatan administratif untuk memenuhi tuntutan kurikulum, bukan sebagai bagian penting dari proses pembelajaran. Akibatnya, instrumen evaluasi sering disusun secara sederhana tanpa mempertimbangkan kebutuhan serta karakteristik peserta didik secara memadai. Permasalahan lain yang sering ditemukan adalah penggunaan instrumen evaluasi yang belum sepenuhnya disusun berdasarkan indikator pencapaian kompetensi. Instrumen yang digunakan kerap hanya berfokus pada aspek hafalan, sehingga kurang mampu mengukur kemampuan berpikir tingkat tinggi peserta didik. Akibatnya, hasil evaluasi belum dapat memberikan gambaran yang menyeluruh mengenai kemampuan peserta didik (Hasanah dkk., 2026). Oleh karena itu, pengujian instrumen, khususnya validitas, menjadi langkah penting dalam menjamin kualitas data penelitian maupun evaluasi pembelajaran. Validitas memiliki tujuan untuk mengukur tingkat keefktifan suatu alat ukur atau media ukur dalam memperoleh data dan juga digunakan untuk mengukur seberapa efektif suatu kuesioner untuk memperoleh data, lebih tepatnya pertanyaan-pertanyaan yang diajukan di kuesioner (Janna & Herianto, 2021b). Selain itu, uji validitas juga diperlukan untuk menganalisis kualitas setiap butir instrumen dan menentukan tingkat kelayakannya dalam proses pengumpulan data.

Beberapa penelitian terdahulu telah mengkaji mengenai uji validitas instrumen penelitian, terutama dalam pengembangan angket sebagai alat ukur variabel penelitian. Dalam penelitian yang dilakukan oleh Riko Al Hakim, Ika Mustika, dan Wiwin Yuliani yang berjudul (*Validitas Dan Reliabilitas Aangket Motivasi Berprestasi*) menekankan bahwa uji validitas untuk menguji ketepatan instrumen penelitian dalam mengukur variabel yang diteliti. Hasil uji validitas menunjukkan bahwa tidak semua item dapat digunakan, hanya item yang memenuhi kriteria (r hitung > r tabel) yang dinyatakan valid dan layak digunakan sebagai alat pengumpulan data (Al Hakim dkk., 2021). Sejalan dengan penelitian tersebut, Putu Gede Subhaktiyasa dalam jurnalnya yang berjudul *(Evaluasi Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif: Sebuah Studi Pustaka)* menegaskan bahwa validitas berperan penting dalam memastikan data yang diperoleh benar-benar mencerminkan variabel yang diteliti sehingga hasil penelitian menjadi lebih akurat dan dapat dipercaya.(Subhaktiyasa, 2024) Selain itu, Afifah Aulia Zayrin, Hayatun Nupus, Khalista Khansa Maizia, Siska Marsela, Rully Hidayatullah, dan Harmonedi dalam jurnal yang berjudul *Analisis Instrumen Penelitian Pendidikan (Uji Validitas Dan Relibilitas Instrumen Penelitian)* menegaskan bahwa validitas merupakan ukuran

yang menunjukkan tingkat ketepatan instrumen dalam mengukur variabel penelitian, sehingga instrumen yang valid mampu menghasilkan data yang akurat dan sesuai dengan kondisi yang sebenarnya (Afifah Aulia Zayrin dkk., 2025).

Guna memastikan bahwa alat ukur yang digunakan dalam penelitian sudah tepat, maka diperlukan penelitian yang lebih mendalam tentang proses uji validitas instrumen. Meskipun beberapa penelitian telah meneliti pentingnya uji validitas instrumen, diskusi yang tersedia masih bersifat umum dan belum secara sistematis meninjau penerapan instrumen penelitian. Dengan demikian, tujuan penelitian tersebut agar memeriksa validitas instrumen penelitian secara menyeluruh. Hal ini sesuai konsep dan tujuan validitas, jenis validitas, dan langkah-langkah pelaksanaan uji validitas untuk menentukan kelayakan dan mutu instrumen. Hasil penelitian ini diharapkan dapat menjadi referensi bagi peneliti dan akademisi dalam mengembangkan instrumen yang lebih tepat dan bisa mengumpulkan data yang lebih akurat dan kredibel. Selain itu, penelitian ini dapat memberikan kontribusi terhadap pengembangan metodologi penelitian serta mendukung peningkatan kualitas hasil penelitian.

METODE

Penelitian ini menggunakan jenis penelititan kepustakaan (library research) dengan pendekatan kualitatif deskriptif. Pendekatan kualitatif deskriptif digunakan karena penelitan ini berfokus pada penggambaran dan pemeahaman secara mendalam terhadap konsep validitas instrumen berdasarkan kajian dari berbagai literatur yang relevan (Zed, 2008, hlm. 2). Tujuan penelitian ini adalah untuk mendeskripsikan dan menganalisis konsep validitas instrumen, khususnya validitas isi, validitas konstruk, dan validitas kriteria berdasarkan sumber-sumber kepustakaan yang kredibel, sehungga diperoleh pemahaman yang komprehensif mengenai penerapan dan penegembangannya dalam penelitian pendidikan. Penelitian kepustakaan dilakukan dengan memanfaatkan berbagai sumber tertulis sebagai data utama, seperti buku, jurnal, dan dokumen yang relevan dengan topik penelitian.

Subjek penelitian ini berupa sumber-sumber literatur yang membahas konsep validitas instrument, khususnya validitas isi, validitas konstruk, dan validitas kriteria. Teknik pengumpulan data dilakukan melalui studi dokumentasi, yaitu dengan cara mengumpulkan, membaca, dan mengkaji berbagai referensi yang relevan. Kemudian, Teknik keabsahan data dalam penelitian ini menggunakan triangulasi sumber, yaitu membandingkan dan mengkaji informasi dari berbagai literatur yang berbeda untuk memperoleh data yang konsisten dan dapat dipercaya. Sumber literatur yang digunakan dipilih berdasarkan kriteria relevansi, kredibilitas, dan kesesuaian dengan topik penelitian (Niam dkk., 2024, hlm. 151). Teknik analisis data yang digunakan adalah analisis deskriptif kualitatif, yaitu Teknik analisis yang dilakukan dengan cara menguraikan, menjelaskan, dan menafsirkan data secara sistematis sehingga menghasilkan pemahaman yang komprehensif terhadap objek yang diteliti. Analisis ini bertujuan untuk menggambarkan secara jelas konsep-konsep yang berkaitan dengan validitas instrumen berdasarkan kajian berbagai sumber literatur (Sugiyono, 2013).

HASIL DAN PEMBAHASAN

Uji Validitas

Instrumen asesmen merupakan alat yang digunakan untuk mengumpulkan data mengenai pencapaian belajar peserta didik, sehingga harus terukur serta memiliki tingkat keakuratan dan konsistensi yang baik. Keakuratan instrumen tercermin dalam validitas, yaitu kemampuan alat ukur dalam menilai kompetensi yang seharusnya diukur, Instrumen yang tidak valid atau tidak reliabel berpotensi menghasilkan data yang keliru dan dapat menyebabkan kesalahan dalam

Copyright © 2026 Shoviyyatuz Zahro, et al.
(cc) BY

pengambilan keputusan pembelajaran. Pengujian validitas dan reliabilitas menjadi langkah penting untuk memastikan instrumen asesmen layak digunakan (Azizah & Chalimatusadiah, 2025, hlm. 6638).

Istilah validitas berasal dari kata *validity*, yang merujuk pada standar keabsahan atau derajat kebenaran suatu hal (Sugiono dkk., 2020). Dalam konteks evaluasi, validitas menunjukkan tingkat ketepatan dan presisi instrumen dalam menjalankan fungsi ukurnya. Suatu instrumen dikatakan memiliki validitas tinggi apabila mampu menghasilkan data yang relevan serta merepresentasikan kondisi nyata sesuai dengan tujuan pengukuran (Ramadhan dkk., 2024). Karena bergantung pada konteks dan tujuan tertentu, validitas bersifat dinamis dan harus selalu dikaitkan dengan dasar pengambilan keputusan yang jelas. Penentuan validitas dilakukan melalui pembuktian empiris dengan berbagai data pendukung untuk memastikan bahwa instrumen layak digunakan (Nurhan Buka, 2025a).

Validitas juga dapat dipahami sebagai kesesuaian antara alat ukur dengan target yang dinilai. Oleh karena itu, suatu instrumen tidak dapat dianggap valid untuk semua situasi, melainkan harus disesuaikan dengan tujuan penggunaannya. Contohnya, tes seleksi sekolah dikatakan valid jika mampu mencerminkan potensi keberhasilan belajar siswa setelah mereka diterima (Ramadhan dkk., 2024). Validitas menjadi aspek penting dalam penelitian pendidikan, terutama untuk mengukur variabel yang bersifat abstrak seperti sikap dan motivasi. Instrumen yang tidak valid akan menghasilkan data yang tidak akurat dan berpotensi menimbulkan kesalahan dalam analisis serta penarikan kesimpulan. Maka dari itu, validitas sangat diperlukan sebagai bukti bahwa temuan penelitian memiliki dasar yang kuat, sahih, dan dapat dipercaya secara ilmiah.

Validitas instrumen sangat bergantung pada validitas datanya. Jika data yang dikumpulkan sudah terbukti valid, maka dapat dipastikan bahwa instrumen yang digunakan untuk memperoleh data tersebut juga merupakan alat ukur yang valid. Validitas menjamin bahwa data yang dikumpulkan benar-benar mencerminkan realita yang diteliti atau "mengukur apa yang seharusnya diukur". Kevalidan suatu kuesioner dapat dibuktikan melalui uji statistik dengan syarat nilai rhitung harus lebih besar dari rtabel. Indeks validitas inilah yang menjadi tolok ukur ketepatan sebuah instrumen evaluasi (Anggraini dkk., 2022).

Pengujian validitas dilakukan untuk menentukan apakah suatu instrumen penelitian secara akurat mengukur konsep yang ingin dievaluasi. Instrumen dengan validitas yang kuat menghasilkan data yang konsisten dengan tujuan pengukuran yang dimaksudkan, sedangkan instrumen dengan validitas yang lemah mungkin menghasilkan hasil yang tidak secara memadai mewakili konstruk yang diteliti (Janna & Herianto, 2021a). Pengujian validitas dilakukan untuk menentukan apakah suatu instrumen penelitian secara akurat mengukur konsep yang ingin dievaluasi. Instrumen dengan validitas yang kuat menghasilkan data yang konsisten dengan tujuan pengukuran yang dimaksudkan, sedangkan instrumen dengan validitas yang lemah mungkin menghasilkan hasil yang tidak secara memadai mewakili konstruk yang diteliti (Sanaky dkk., 2021b, hlm. 433). Akibatnya, alat ukur yang valid mampu menghasilkan data yang secara akurat mencerminkan kondisi sebenarnya, sehingga meningkatkan kredibilitas hasil pengukuran. Tujuan utama penilaian validitas adalah untuk memastikan bahwa suatu instrumen sesuai untuk digunakan dan mampu menghasilkan data yang akurat serta bermakna. Melalui pengujian validitas, peneliti dapat mengevaluasi apakah suatu instrumen memenuhi standar yang telah ditetapkan atau memerlukan penyempurnaan lebih lanjut sebelum diterapkan dalam suatu penelitian (Purwanti dkk., 2026, hlm. 187).

Urgensi validitas terletak pada perannya dalam menjaga ketepatan hasil penilaian agar tetap sesuai dengan tujuan pembelajaran. Instrumen yang valid memungkinkan pendidik memperoleh gambaran yang akurat mengenai kemampuan peserta didik dan mendukung perbaikan strategi pembelajaran (Nurhan Buka,

2025b). Sebaliknya, instrumen yang tidak memenuhi standar validitas berpotensi menghasilkan data yang menyimpang dari kondisi sebenarnya dan melemahkan kredibilitas hasil penelitian. Oleh karena itu, kedudukan validitas sangatlah krusial; ia bertindak sebagai pelindung integritas penelitian guna memastikan bahwa setiap temuan yang dihasilkan benar-benar sah dan dapat dipercaya (Reza dkk., 2026).

Jenis-jenis Validitas

Jika ditinjau dari jenisnya, validitas terbagi menjadi dua yakni validitas logis dan validitas empiris. Validitas logis mencakup validitas isi (content validity) dan validitas konstruk (construct validity), sedangkan validitas empiris meliputi validitas konkuren (validitas kriteria) dan validitas prediktif (Eliyah, 2019, hlm. 44).

1. Validitas logis

Validitas logis berkaitan dengan istilah "logis" yang berarti berdasarkan penalaran. Oleh karena itu, validitas logis merujuk pada instrumen evaluasi yang disusun atas dasar pertimbangan rasional. Suatu butir soal dinyatakan valid apabila penyusunannya mengacu pada teori dan ketentuan yang berlaku sehingga menjadi instrumen yang tepat. Validitas logis mencakup dua aspek utama, yaitu validitas isi dan validitas konstruk.

a. Validitas isi

Validitas isi adalah proses penilaian terhadap butir-butir soal dalam suatu instrumen untuk memastikan bahwa isi yang terkandung di dalamnya telah mencerminkan penguasaan materi secara tepat dan sesuai dengan tujuan yang akan dicapai. Dalam penelitian pendidikan, validitas isi berfokus pada keselarasan antara materi yang terdapat dalam instrumen dengan tujuan pembelajaran atau kompetensi yang hendak diukur. Validitas isi tidak dapat dinyatakan dalam bentuk angka, tetapi dinilai secara logis melalui pertimbangan atau penilaian para ahli (professional). Contoh pada mata pelajaran Bahasa Indonesia dengan materi cerpen, soal yang diberikan mencakup penentuan tema, tokoh, alur, dan amanat. Tes tersebut dikatakan memiliki validitas isi karena mampu mewakili seluruh materi yang seharusnya diujikan kepada siswa.

b. Validitas kontruk

Validitas konstruk merupakan validitas yang dilihat dari segi susunan, kerangka, atau rancangan suatu tes. Secara istilah, suatu tes dikatakan memiliki validitas konstruk apabila butir-butir soalnya mampu mengukur berbagai aspek kemampuan berpikir sebagaimana tercantum dalam standar kompetensi, kompetensi dasar, indikator, dan tujuan pembelajaran dalam kurikulum. Soal yang disusun mencakup aspek ingatan, pemahaman, penerapan, analisis, sintesis, hingga evaluasi. Contoh pada mata Pelajaran PAI, tujuan pembelajaran "siswa mampu menjelaskan hikmah beriman kepada hari akhir", maka instruksi atau perintah pada butir soal tes tersebut adalah: "Jelaskan hikmah beriman kepada haro akhir dalam kehidupan sehari-hari!" (Syaifudin, 2020, hlm. 110).

2. Validitas empiris

Validitas empiris, di mana istilah "empiris" merujuk pada pengalaman. Oleh karena itu, suatu instrumen dapat dinyatakan valid dari sisi empiris apabila telah teruji berdasarkan pengalaman atau melalui data hasil uji coba. Dalam validitas empiris ini terdapat beberapa jenis, yaitu validitas prediktif (*predictive validity*) dan validitas konkuren (*concurrent validity*).

a. Validitas prediktif

Validitas ini muncul ketika suatu instrumen mampu secara tepat memperkirakan hasil atau kejadian yang akan terjadi di masa depan. Artinya, instrumen tersebut memiliki kemampuan yang baik dalam meramalkan suatu peristiwa berdasarkan data yang diperoleh. Hal ini menunjukkan bahwa

Copyright © 2026 Shoviyyatuz Zahro, et al.
(cc) BY

instrumen dapat dipercaya untuk mengukur atau mengidentifikasi variabel yang berkaitan dengan prediksi hasil di waktu mendatang. Oleh karena itu, instrumen tersebut telah memenuhi kriteria validitas prediktif. Tingkat keakuratannya dapat dilihat dari adanya hubungan atau korelasi antara hasil pengukuran saat ini dengan hasil yang diperoleh pada masa yang akan datang.

Salah satu contohnya dapat dilihat pada pelaksanaan tes masuk perguruan tinggi, yang menghasilkan dua kemungkinan, yaitu lulus dan tidak lulus. Peserta yang dinyatakan lulus diperkirakan memiliki kemampuan untuk mengikuti perkuliahan dengan baik, sedangkan peserta yang tidak lulus diprediksi belum mampu menjalani proses perkuliahan secara optimal di masa yang akan datang.

b. Validitas konkuren

Validitas konkomitan mengacu pada sejauh mana nilai tes sesuai dengan ukuran kinerja lain yang diperoleh pada periode yang sama. Hubungan yang kuat antara ukuran-ukuran tersebut menunjukkan bahwa instrumen tersebut secara akurat mencerminkan konstruk yang sedang dinilai pada titik waktu tertentu tersebut. Dengan demikian, validitas konkomitan umumnya dianggap sebagai salah satu bentuk validitas empiris.

Misalnya, seorang pendidik yang ingin mengevaluasi validitas penilaian sumatif yang baru dikembangkan dapat membandingkan nilai siswa pada tes tersebut dengan indikator pencapaian yang sudah ada, seperti hasil kuis, penilaian di kelas, atau ujian sebelumnya. Jika nilai tes sumatif menunjukkan hubungan yang konsisten dengan ukuran-ukuran yang telah ditetapkan tersebut, tes tersebut dapat dianggap memberikan gambaran yang akurat mengenai hasil belajar dan kemajuan akademik siswa saat ini (Anshari dkk., 2024, hlm. 970).

Langkah-Langkah Uji Validitas

Pengujian validitas merupakan fase kritis dalam penelitian kuantitatif dan pengembangan instrumen, karena berfungsi untuk mengkonfirmasi bahwa alat ukur digunakan secara akurat dan tepat dalam menilai variabel yang dimaksud. Validitas tidak hanya menyangkut keakuratan isi instrumen tetapi juga mencakup keselarasan dengan konstruk teoretis dan hubungannya dengan kriteria eksternal. Akibatnya, proses pengujian validitas harus diprogram melalui beberapa langkah-langkah yang saling terkait.

1. Perumusan Struktur dan Definisi Operasional

Langkah awal dalam pengujian validitas melibatkan perumusan konseptual Struktur atau variabel penelitian, dengan mengacu pada teori-teori yang relevan.Struktur mewakili konsep abstrak yang membentuk dasar pengukuran dan oleh karena itu harus didefinisikan secara jelas dan operasional untuk memungkinkan pengukuran melalui instrumen. Perumusan Struktur yang tidak tepat akan menyebabkan ketidak akuratan pada indikator dan item instrumen yang dikembangkan. Selanjutnya, definisi operasional memiliki peran dalam menguraikan konsep menjadi indikator-indikator yang dapat diamati dan diukur secara empiris.

2. Pengembangan kisi-kisi Instrumen

Langkah berikutnya setelah konsep ditentukan adalah mengembangkan kisi kisi instrumen.kisi kisi ini berfungsi sebagai rancangan atau Struktur dasar untuk pembuatan item-item instrumen.kisi kisi tersebut mencakup dimensi dari variabel, indikator-indikatornya, jumlah pertanyaan, serta format pertanyaan atau pernyataan. Pengembangan kisi kisi ini bertujuan untuk memastikan bahwa semua aspek dari konsep terwakili secara seimbang di dalam instrumen. Hal ini sangat penting karena validitas instrumen tidak hanya bergantung pada kualitas item, tetapi juga pada keterwakilan semua indikator variabel yang diukur. Dengan adanya

kisi kisi, peneliti dapat mencegah terjadinya bias dalam pengukuran yang timbul akibat ketidakseimbangan jumlah item pada setiap indikator.

3. Validasi Isi (Content Validity) melalui Penilaian Ahli

Tahap selanjutnya adalah melaksanakan validasi isi dengan melibatkan partisipasi Penilaian ahli (expert judgment). Validitas isi berfokus pada tingkat relevansi dan penggambaran elemen-elemen dalam instrumen terhadap struktur yang sedang diukur (Ihsan, 2015, hlm. HAL:173-174). Secara praktis, ahli diminta untuk mengevaluasi berbagai aspek, meliputi:

a. Kerelevanan item dengan indikator
b. Ketepatan penggunaan bahasa
c. Struktur instrumen
d. Kesesuaian skala pengukuran

Proses validasi ini dapat dilakukan melalui pendekatan kualitatif maupun kuantitatif. Dalam konteks kuantitatif, evaluasi ahli dapat dikuantifikasi menggunakan indeks validitas seperti *Aiken's V* atau *Content Validity Ratio (CVR)*. Studi terdahulu menunjukkan bahwa proses validasi terkadang melibatkan sejumlah pakar dari disiplin ilmu pembelajaran, evaluasi, dan pengembangan instrumen guna meraih perspektif yang menyeluruh (Erviana, 2023, hlm. HAL:200-201). Validitas isi merupakan fondasi sangat penting sebelum instrumen diuji lebih lanjut secara nyata. Instrumen yang belum memenuhi kriteria validitas isi tidak direkomendasikan untuk melanjutkan ke tahapan selanjutnya.

4. Penyesuaian Instrumen Berdasarkan masukan ahli

Setelah validasi oleh ahli, langkah berikutnya Adalah dilakukan penyesuaian pada instrumen berdasarkan rekomendasi dan umpan balik yang diterima. Penyesuaian ini dirancang untuk meningkatkan kualitas instrumen, baik dalam aspek materi maupun aspek teknis. Suatu instrumen dianggap valid secara isi jika telah memenuhi standar kelayakan sesuai penilaian para ahli dan tidak memerlukan modifikasi tambahan.Apabila masih ditemukan ketidaksesuaian, maka proses penyesuaian perlu diulang sampai instrumen tersebut benar-benar siap untuk digunakan.

5. Pengujian Instrumen (Empirical Testing)

Instrumen yang telah melalui validasi teoretis kemudian diujicobakan pada sampel responden dalam skala yang lebih kecil. Pengujian ini dimaksudkan untuk mengumpulkan data empiris yang selanjutnya akan digunakan untuk analisis validitas yang lebih mendalam.Jumlah partisipan dalam tahap pengujian ini harus mencukupi agar hasil analisis dapat diandalkan. Secara umum, pengujian dapat dilaksanakan dengan partisipan minimal berjumlah 30 orang, namun untuk analisis yang lebih rumit seperti analisis faktor, disarankan partisipan melebihi 100 orang. Melalui proses pengujian ini, peneliti dapat mengidentifikasi unsur-unsur yang tidak berkinerja optimal, misalnya unsur yang tidak jelas maknanya atau tidak sanggup membedakan antara responden.

6. Evaluasi Validitas Struktur

Validitas Struktur adalah tahapan lanjutan yang berupaya menguji sejauh mana suatu instrumen secara akurat mengukur konstruk yang telah ditentukan sesuai dengan struktur teori. Evaluasi validitas ini dilaksanakan melalui metode analisis statistik, contohnya analisis faktor eksploratori (EFA) atau analisis faktor konfirmatori (CFA). Dalam proses analisis ini, setiap unsur akan dievaluasi korelasi nya dengan konstruk melalui skor *loading factor*. Unsur dianggap valid jika menunjukkan skor *loading factor* yang melebihi ambang batas minimum yang ditetapkan (misalnya, lebih besar dari 0,3). Lebih lanjut, analisis faktor juga dimanfaatkan untuk menelaah kesesuaian struktur instrumen dengan dimensi teoritis yang telah dirumuskan sebelumnya.

Apabila hasil analisis struktur menunjukkan kualitas yang baik, hal tersebut menandakan bahwa instrumen tersebut memiliki bentuk yang logis dan selaras

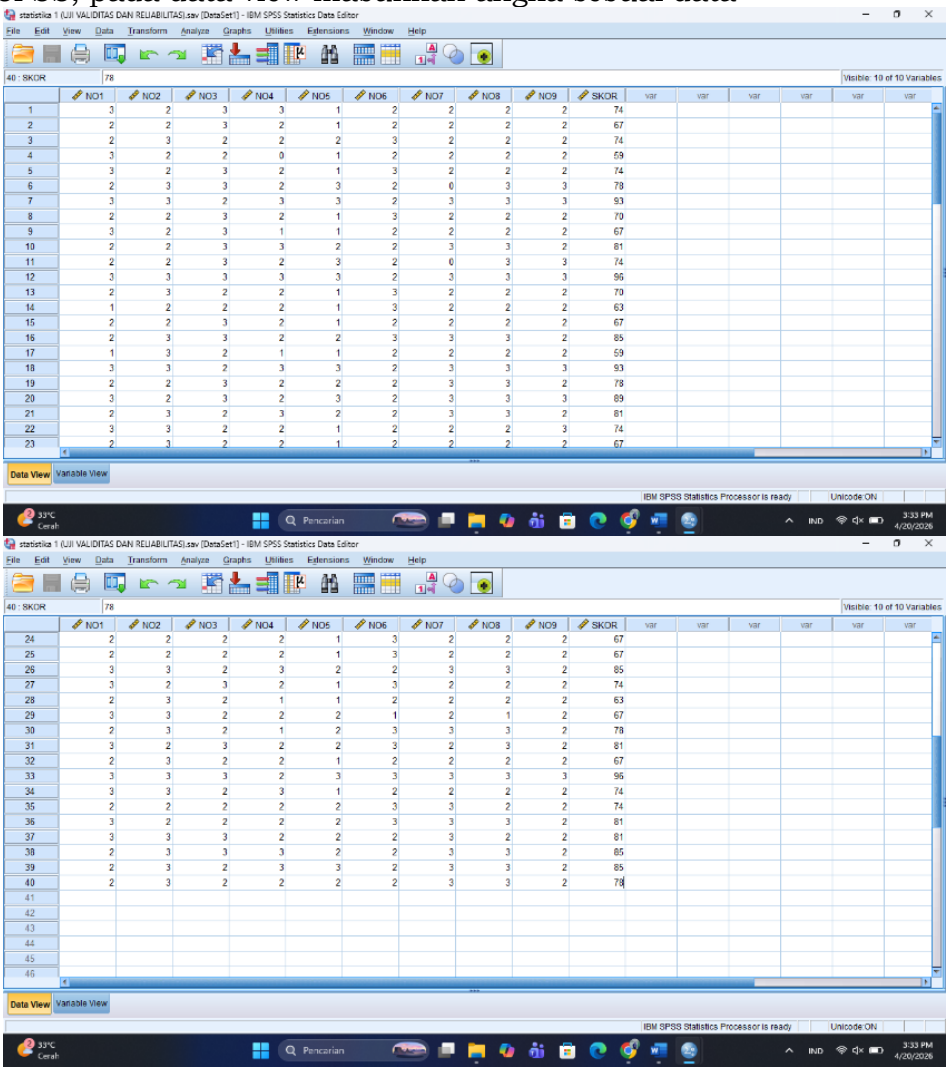
Copyright © 2026 Shoviyyatuz Zahro, et al.
(cc) BY

dengan konsep teoritis, sehingga efektif dalam pengukuran variabel secara presisi. Untuk memperoleh bukti empiris mengenai validitas struktur ini, diperlukan analisis statistik untuk mengidentifikasi pola hubungan antar item dalam instrumen. Metode yang umum digunakan adalah Analisis Faktor Eksploratori (EFA), yang bertujuan untuk mengelompokkan item berdasarkan karakteristik yang sama dan mengungkap dimensi laten yang dibentuk oleh data. Analisis ini memungkinkan peneliti untuk memastikan apakah struktur faktor yang dihasilkan selaras dengan konstruk teoritis yang telah dirumuskan sebelumnya.

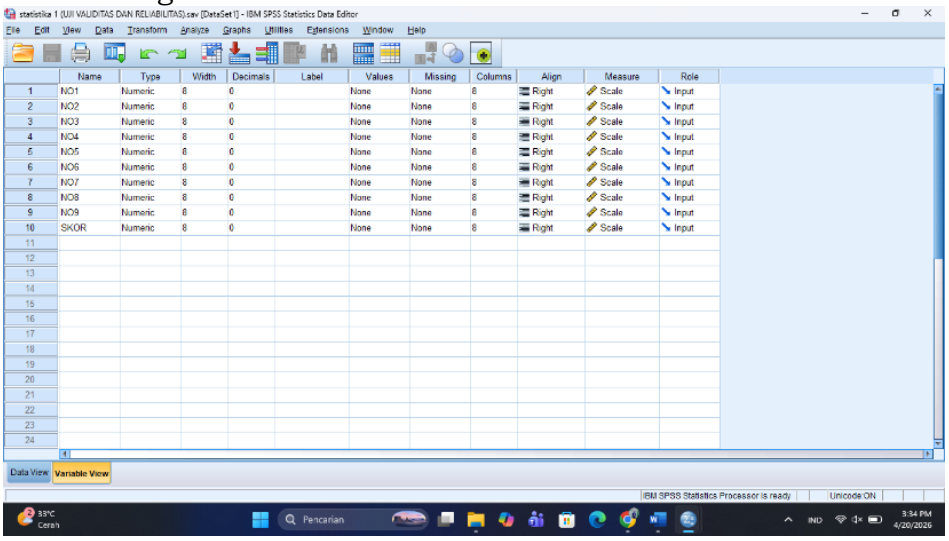
Implementasi analisis faktor dalam penelitian ini difasilitasi oleh perangkat lunak statistik untuk meningkatkan akurasi perhitungan dan menyederhanakan interpretasi hasil. Pemanfaatan perangkat lunak statistik memungkinkan peneliti untuk secara sistematis dan objektif menguji kesesuaian data, mengekstrak faktor, dan menentukan nilai muatan faktor. Oleh karena itu, temuan dari analisis ini akan menyajikan fondasi yang kokoh untuk menetapkan signifikansi konstruksi instrumen penelitian. Selanjutnya, guna memverifikasi signifikansi konstruksi secara empiris, analisis dilaksanakan dengan memanfaatkan perangkat lunak IBM SPSS Statistics melalui langkah-langkah sebagai berikut:

Data Uji Validitas menggunakan SPSS

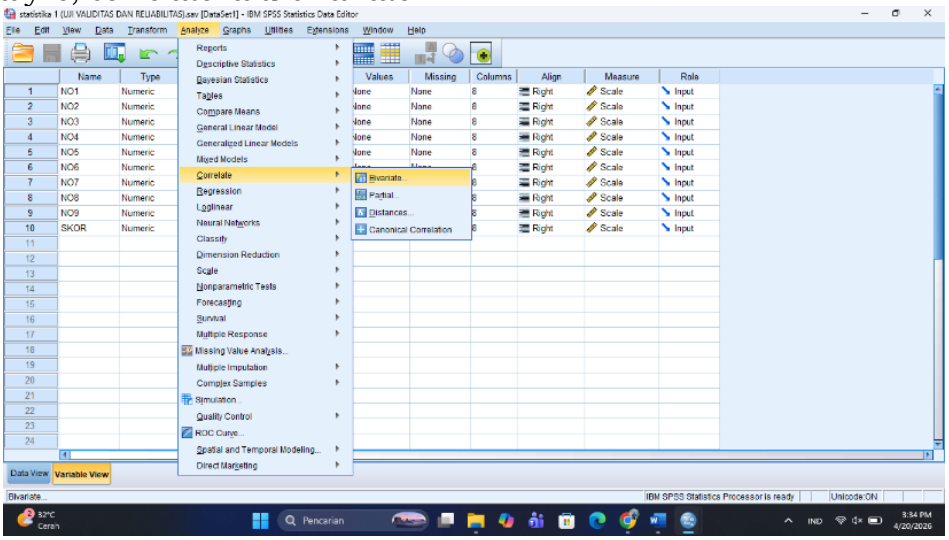
a. Buka SPSS, pada data view masukkan angka sesuai data



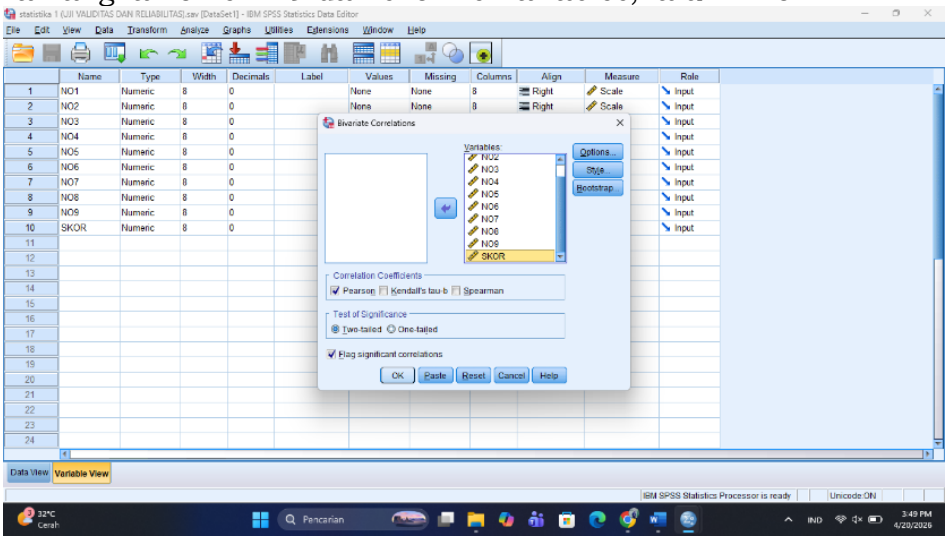
b. Pada variabel view bagian nama diganti sesuai pada kolom data, decimal diganti 0, dan measure diganti scale



c. Klik analyze, correlate lalu bivariate



d. Pindahkan angka nomer 1-9 dan skor ke Variables, lalu klik ok



Copyright © 2026 Shoviyyatuz Zahro, et al.
CC BY

e. Muncul data tabel corelations

Correlations

		NO1	NO2	NO3	NO4	NO5	NO6	NO7	NO8	NO9	SKOR
NO1	Pearson Correlation	1	-.022	.196	.163	.212	-.156	.187	.050	.324*	.439**
	Sig. (2-tailed)		.894	.227	.315	.190	.336	.247	.760	.041	.005
	N	40	40	40	40	40	40	40	40	40	40
NO2	Pearson Correlation	-.022	1	-.394*	.212	.295	-.305	.199	.153	.201	.310
	Sig. (2-tailed)	.894		.012	.189	.065	.056	.219	.346	.214	.051
	N	40	40	40	40	40	40	40	40	40	40
NO3	Pearson Correlation	.196	-.394*	1	.091	.164	.015	-.129	.218	.176	.266
	Sig. (2-tailed)	.227	.012		.577	.312	.929	.428	.177	.278	.097
	N	40	40	40	40	40	40	40	40	40	40
NO4	Pearson Correlation	.163	.212	.091	1	.393*	-.167	.298	.368*	.207	.622**
	Sig. (2-tailed)	.315	.189	.577		.012	.303	.062	.019	.199	.000
	N	40	40	40	40	40	40	40	40	40	40
NO5	Pearson Correlation	.212	.295	.164	.393*	1	-.173	.238	.736**	.652**	.816**
	Sig. (2-tailed)	.190	.065	.312	.012		.286	.140	.000	.000	.000
	N	40	40	40	40	40	40	40	40	40	40
NO6	Pearson Correlation	-.156	-.305	.015	-.167	-.173	1	.052	.042	-.193	.002
	Sig. (2-tailed)	.336	.056	.929	.303	.286		.751	.796	.234	.993
	N	40	40	40	40	40	40	40	40	40	40
NO7	Pearson Correlation	.187	.199	-.129	.298	.238	.052	1	.350*	-.139	.536**
	Sig. (2-tailed)	.247	.219	.428	.062	.140	.751		.027	.393	.000
	N	40	40	40	40	40	40	40	40	40	40
NO8	Pearson Correlation	.050	.153	.218	.368*	.736**	.042	.350*	1	.415**	.767**
	Sig. (2-tailed)	.760	.346	.177	.019	.000	.796	.027		.008	.000
	N	40	40	40	40	40	40	40	40	40	40
NO9	Pearson Correlation	.324*	.201	.176	.207	.652**	-.193	-.139	.415**	1	.559**
	Sig. (2-tailed)	.041	.214	.278	.199	.000	.234	.393	.008		.000
	N	40	40	40	40	40	40	40	40	40	40
SKOR	Pearson Correlation	.439**	.310	.266	.622**	.816**	.002	.536**	.767**	.559**	1
	Sig. (2-tailed)	.005	.051	.097	.000	.000	.993	.000	.000	.000	
	N	40	40	40	40	40	40	40	40	40	40

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

f. Menghitung r tabel

Df : (N-2)

N: Jumlah responden

Df = (N-2) = 40- 2

= 38

df = (N-2)	Tingkat signifikansi untuk uji satu arah				
	0.05	0.025	0.01	0.005	0.0005
	Tingkat signifikansi untuk uji dua arah				
	0.1	0.05	0.02	0.01	0.001
1	0.9877	0.9968	0.9995	0.9999	1.0000
2	0.9000	0.9500	0.9800	0.9900	0.9990
3	0.8054	0.8783	0.9343	0.9587	0.9911
4	0.7293	0.8114	0.8822	0.9172	0.9741
5	0.6694	0.7545	0.8329	0.8745	0.9509
6	0.6215	0.7067	0.7887	0.8343	0.9249
7	0.5822	0.6664	0.7498	0.7977	0.8983
8	0.5494	0.6319	0.7155	0.7646	0.8721
9	0.5214	0.6021	0.6851	0.7348	0.8470
10	0.4973	0.5760	0.6581	0.7079	0.8233
11	0.4762	0.5529	0.6339	0.6835	0.8010
12	0.4575	0.5324	0.6120	0.6614	0.7800
13	0.4409	0.5140	0.5923	0.6411	0.7604
14	0.4259	0.4973	0.5742	0.6226	0.7419
15	0.4124	0.4821	0.5577	0.6055	0.7247
16	0.4000	0.4683	0.5425	0.5897	0.7084
17	0.3887	0.4555	0.5285	0.5751	0.6932
18	0.3783	0.4438	0.5155	0.5614	0.6788
19	0.3687	0.4329	0.5034	0.5487	0.6652
20	0.3598	0.4227	0.4921	0.5368	0.6524
21	0.3515	0.4132	0.4815	0.5256	0.6402
22	0.3438	0.4044	0.4716	0.5151	0.6287
23	0.3365	0.3961	0.4622	0.5052	0.6178
24	0.3297	0.3882	0.4534	0.4958	0.6074
25	0.3233	0.3809	0.4451	0.4869	0.5974
26	0.3172	0.3739	0.4372	0.4785	0.5880
27	0.3115	0.3673	0.4297	0.4705	0.5790
28	0.3061	0.3610	0.4226	0.4629	0.5703
29	0.3009	0.3550	0.4158	0.4556	0.5620
30	0.2960	0.3494	0.4093	0.4487	0.5541
31	0.2913	0.3440	0.4032	0.4421	0.5465
32	0.2869	0.3388	0.3972	0.4357	0.5392
33	0.2826	0.3338	0.3916	0.4296	0.5322
34	0.2785	0.3291	0.3862	0.4238	0.5254
35	0.2746	0.3246	0.3810	0.4182	0.5189
36	0.2709	0.3202	0.3760	0.4128	0.5126
37	0.2673	0.3160	0.3712	0.4076	0.5066
38	0.2638	0.3120	0.3665	0.4026	0.5007
39	0.2605	0.3081	0.3621	0.3978	0.4950
40	0.2573	0.3044	0.3578	0.3932	0.4896
41	0.2542	0.3008	0.3536	0.3887	0.4843
42	0.2512	0.2973	0.3496	0.3843	0.4791
43	0.2483	0.2940	0.3457	0.3801	0.4742
44	0.2455	0.2907	0.3420	0.3761	0.4694
45	0.2429	0.2876	0.3384	0.3721	0.4647
46	0.2403	0.2845	0.3348	0.3683	0.4601

Diketahui Df 38, sehingga ditarik kesimpulan r tabel dengan taraf signifikansi 0,05% adalah 0,3120

Copyright © 2026 Shoviyyatuz Zahro, et al.
(cc) BY

NO	Person Corelation	R tabel	Keterangan
1.	0, 439		Valid
2.	0,310		Tidak valid
3.	0,266		Tidak valid
4.	0,622		Valid
5.	0,816	0,3120	Valid
6.	0,002		Tidak valid
7.	0,536		Valid
8.	0,767		Valid
9.	0,559		Valid

Jumlah item yang teridentifikasi adalah 9 butir soal. Berdasarkan evaluasi validitas yang dilakukan dengan korelasi Product Moment terhadap 40 responden, nilai r tabel yang diperoleh adalah 0,312.

Hasil analisis mengindikasikan bahwa:

1) Sebanyak 6 item dikategorikan valid, yaitu NO1, NO4, NO5, NO7, NO8, dan NO9, dikarenakan nilai r hitung mereka melebihi nilai r tabel.
2) Sementara itu, 3 item dinilai tidak valid, yaitu NO2, NO3, dan NO6, disebabkan nilai r hitung mereka lebih rendah dari nilai r tabel.

Dengan demikian, dapat ditarik kesimpulan bahwa mayoritas item pada instrumen telah memenuhi kriteria validitas dan sesuai untuk digunakan, meskipun terdapat beberapa item yang memerlukan revisi atau penghapusan.

7. Analisis Validitas Kriteria

Penilaian kriteria dilaksanakan dengan membandingkan luaran pengukuran instrumen terhadap tolak ukur eksternal yang bersesuaian. Tolak ukur ini umumnya mencakup instrumen lain yang telah terverifikasi validitasnya atau standar spesifik. Uji coba dilakukan melalui aplikasi analisis korelasi, misalnya korelasi Product Moment. Besaran koefisien korelasi mengindikasikan tingkat keterkaitan antara instrumen dan tolak ukurnya. Semakin substantial nilai keterkaitan, semakin kuat validitas instrumen tersebut. Validitas kriteria dapat diklasifikasikan menjadi:

a. Validitas prediktif (predictive validity)

Validitas prediktif mewakili sebuah tipe validitas kriteria yang mengukur kemampuan sebuah alat ukur dalam memperkirakan hasil di periode mendatang. Dengan kata lain, skor yang diperoleh dari pengukuran saat ini berfungsi untuk mengestimasi kondisi atau performa subjek di waktu yang akan datang. Proses pengujiannya melibatkan perbandingan data dari alat ukur dengan kriteria yang dinilai pada periode waktu yang berbeda (masa depan). Apabila ditemukan keterkaitan yang nyata antara skor alat ukur dan luaran di masa depan, maka alat ukur tersebut dianggap memiliki validitas prediktif yang kuat.

Contoh:

1) Ujian penerimaan universitas diaplikasikan untuk mengestimasi pencapaian akademis mahasiswa pada semester selanjutnya.
2) Uji seleksi pekerjaan dimanfaatkan untuk mengantisipasi performa pegawai setelah menjalani masa kerja.

Karakteristik utama:
1) Terdapat rentang temporal antara waktu pengukuran alat ukur dan hasil kriteria
2) Dipergunakan untuk estimasi (prediction)
3) Diverifikasi melalui uji korelasi antara skor awal dan luaran di masa depan

b. Validitas bersamaan (concurrent validity)

Validitas bersamaan merupakan salah satu tipe validitas kriteria yang mengindikasikan sejauh mana skor dari suatu instrumen berhubungan dengan skor dari kriteria eksternal yang diukur pada periode waktu yang identik. Dalam konteks ini, instrumen yang sedang divalidasi akan diadministrasikan dan dibandingkan dengan instrumen lain yang telah terbukti valid sebelumnya. Kedua instrumen tersebut diberikan kepada partisipan secara simultan. Apabila terdapat korelasibe yang signifikan antara hasil dari kedua instrumen, maka instrumen yang baru dikembangkan tersebut dianggap memiliki validitas konkuren.

Contoh:
1) Sebuah kuesioner mengenai motivasi belajar dibandingkan dengan skala motivasi yang telah distandarisasi, diadministrasikan pada waktu yang sama.
2) Sebuah tes baru untuk mengukur kesehatan mental dibandingkan dengan hasil diagnosis yang diberikan oleh seorang psikolog pada periode waktu yang sama.

Karakteristik utama:
1) Proses pengukuran dilakukan pada interval waktu yang sama
2) Bertujuan untuk mengevaluasi tingkat kesesuaian (agreement)
3) Memerlukan perbandingan dengan instrumen pengukuran yang sudah tervalidasi (Yusup, 2018a, hlm:18-19).

8. Penentuan Keputusan Validitas Instrumen

Fase final melibatkan penetapan keputusan untuk setiap item instrumen berdasarkan hasil analisis. Item yang memenuhi parameter validitas akan dipertahankan, sementara item yang tidak memenuhi parameter dapat:
a. Disempurnakan (direvisi)
b. Dihapus(dibuang)

Keputusan ini memiliki signifikansi yang besar karena akan menentukan mutu akhir dari instrumen yang akan dipakai dalam sebuah riset. Instrumen yang sahih akan mengantarkan pada perolehan data yang presisi dan dapat diyakini, oleh sebab itu berkontribusi pada kesahihan temuan riset secara komprehensif (Setiawan dkk., 2024, hlm. HAL:50-53).

Cara Menentukan Kevalidan Instrumen

Validitas instrument dapat ditinjau melalui beberapa bentuk pembuktian. Secara umum, validitas tersebut meliputi validitas isi atau validitas konten yang berkaitan dengan kesesuaian konten atau isi dari suatu instrumen, validitas konstruk yang menunjukkan kesesuaian instrument dengan konnsep atau teori yang mendasarinnya, serta validitas kriteria yang dilihat dari hubungan instrumen dengan ukuran lain yang relevan (Zayrin dkk., 2025b). Masing-masing jenis validitas memiliki cara penentuan yang berbeda sesuai dengan karakteristiknya.

Validitas isi ditentukan melalui penilaian ahli (*expert judgement*). Instrumen dinyatakan valid apabila seluruh komponen dalam instrument, seperti kesesuaian dengan indikator, kejelasan Bahasa, format, skala, serta tata letak, telah dinilai sesuai oleh ahli. Apabila setelah dilakukan penilaian tidak terdapat lagi suatu revisi atau perbaikan dari sang ahli, maka instrument tersebut dapat dinyatakan valid secara isi. Sebaliknya, jika masih terdapat saran perbaikan dari ahli, maka

Copyright © 2026 Shoviyyatuz Zahro, et al.
(cc) BY

instrumen perlu diperbaiki hingga dinyatakan layak oleh ahli (Yusup, 2018b, hlm. 19).

Selanjutnya, validitas konstruk ditentukan dengan melihat kesesuain antara instrument dengan konsep atau teori yang mendasari variabel yang diukur. Instrument dinyatakan valid apabila setiap butir pernyataan atau pertanyaan telah mencerminkan konstruk atau defines variabel yang berasal dari teori yang tepat. Dengan demikian, kejelasan definisi variabel dan kesesuaian item dengan aspek-aspek teori menjadi dasar utama dalam menentukan validitas konstruk (Yusup, 2018b, hlm. 19).

Sedangkan validitas kriteria ditentukan dengan cara membandingkan instrument yang dikembangkan dengan ukuran lain yang dijadikan sebagai kriteria. Kriteria tersebut dapat berupa instrument lain yang telah terbukti valid atau ukuran eksternal yang relevan. Intrumen dinyatakan valid apabila memiliki hubungan atau korelasi yang tinggi dengan kriteria tersebut. Dalam praktiknya, penentuan validitas kriteria dilakukan dengan memebandingkan nilai koefisien korelasi antara skor instrumen dengan kriteria. Instrumen dinyatakan valid apabila nilai koefisien korelasi yang diperoleh memenuhi kriteria sebagai berikut:

$$r_{\{hitung\}} \geq r_{\{tabel\}} \Rightarrow Valid$$

Semakin tinggi nilai koefisien korelasi atau mendekati +1, maka semakin kuat hubungan antara instrumen dengan kriteria, sehingga tingkat validitas instrument semakin tinggi (Yusup, 2018b, hlm. 19).

Koefisien korelasi tersebut dapat dihitung dengan menggunakan rumus korelasi product Moment (pearson) sebagai berikut:(Saputri dkk., 2023)

$$r_{xy} = \frac{n\Sigma x_i y_i - (\Sigma x_i)(\Sigma y_i)}{\sqrt{[(n\Sigma {x_i}^2 - (\Sigma x_i)^2)(n\Sigma {y_i}^2 - (\Sigma y_i)^2)]}}$$

Berdasarkan uraian tersebut, validitas instrument merupakan aspek penting yang harus diperhatikan untuk meamstikan bahwa instrument mampu mengukur variabel secara tepat. Penentuan validitas instrument dapat dilakukan melalui beberapa jenis validitas, yaitu validitas isi, validitas konstruk, dan validitas kriteria. Instrument dinyatakan valid apabila telah sesuai dengan indikator yang diukur, mencerminkan konsep atau teori yang mendasarinya, serta memiliki hubungan yang kuat dengan kriteria tertentu berdasarkan hasil pengujian. Dengan demikian, instrument yang valid akan menghasilkan data yang akurat dan dapat dipercaya dalam penelitian.

PENUTUP

Kesimpulan

Uji validitas menjadi fondasi utama dalam menjamin kualitas instrumen asesmen. Instrumen yang valid mampu mengukur aspek yang memang menjadi tujuan pengukuran serta menghasilkan data yang akurat dan dapat dipercaya sebagai dasar pengambilan keputusan dalam pembelajaran maupun penelitian. Ketika instrumen tidak melalui uji validitas yang memadai, data yang dihasilkan berpotensi menyimpang dari kondisi sebenarnya dan dapat melemahkan keabsahan hasil penelitian. Uji validitas hadir sebagai bagian penting dalam menjaga integritas dan ketepatan proses ilmiah.

Dari sisi klasifikasi, validitas terbagi menjadi validitas logis dan validitas empiris yang memiliki peran saling melengkapi. Validitas logis menekankan kesesuaian instrumen berdasarkan kajian teori, konsep, dan penalaran yang kuat, sedangkan validitas empiris berangkat dari hasil uji coba instrumen di lapangan.

Keduanya menunjukkan bahwa keabsahan instrumen tidak cukup dilihat dari satu pendekatan saja, melainkan perlu dikaji secara konseptual sekaligus dibuktikan melalui data. Instrumen yang berkualitas lahir dari perpaduan antara kekuatan teori dan ketepatan hasil pengukuran di lapangan.

Pada aspek yang lebih spesifik, validitas logis mencakup validitas isi dan validitas konstruk yang menekankan keselarasan antara materi, indikator, serta tujuan pembelajaran. Sementara itu, validitas empiris seperti validitas prediktif dan konkuren menunjukkan kemampuan instrumen dalam mencerminkan kondisi nyata bahkan memberikan gambaran terhadap kemungkinan hasil di masa mendatang. Instrumen asesmen tidak hanya berfungsi sebagai alat ukur, tetapi juga sebagai sarana untuk memahami perkembangan peserta didik secara lebih utuh dan mendalam.

Proses uji validitas memperlihatkan bahwa pengembangan instrumen harus dilakukan secara sistematis dan terencana. Tahapan dimulai dari perumusan konsep, penyusunan kisi-kisi, hingga proses validasi oleh para ahli sebelum diujicobakan. Setiap langkah memiliki kontribusi penting dalam memastikan instrumen benar-benar merepresentasikan variabel yang diukur. Kualitas instrumen terbentuk melalui proses ilmiah yang matang dan berkelanjutan, sehingga hasil yang diperoleh dapat dipertanggungjawabkan secara akademik maupun praktis.

Saran

Berdasarkan hasil kajian literatur, validitas instrumen asesmen merupakan aspek yang sangat penting dalam menjamin mutu evaluasi pembelajaran Pendidikan Agama Islam. Oleh karena itu, guru dan pengembang instrumen perlu memastikan bahwa setiap instrumen yang digunakan telah melalui proses validasi yang sesuai agar hasil evaluasi yang diperoleh benar-benar mencerminkan kompetensi peserta didik. Selain itu, lembaga pendidikan perlu memberikan dukungan melalui pelatihan dan pendampingan terkait penyusunan serta pengujian validitas instrumen asesmen agar kompetensi guru dalam melaksanakan evaluasi pembelajaran semakin meningkat. Penelitian selanjutnya disarankan untuk mengkaji validitas instrumen asesmen PAI melalui penelitian lapangan pada berbagai jenjang pendidikan dan konteks pembelajaran yang berbeda sehingga diperoleh gambaran yang lebih komprehensif mengenai implementasi dan efektivitasnya. Di samping itu, penelitian mendatang dapat mengembangkan model atau panduan validasi instrumen asesmen yang lebih praktis dan sesuai dengan kebutuhan pembelajaran PAI di era digital. Dengan demikian, instrumen asesmen yang digunakan tidak hanya memenuhi kriteria validitas, tetapi juga mampu mendukung peningkatan kualitas pembelajaran dan pencapaian tujuan pendidikan secara optimal.

DAFTAR PUSTAKA

Afifah Aulia Zayrin, Hayatun Nupus, Khalista Khansa Maizia, Siska Marsela, Rully Hidayatullah, & Harmonedi, H. (2025). Analisis Instrumen Penelitian Pendidikan (Uji Validitas Dan Relibilitas Instrumen Penelitian). *Jurnal QOSIM Jurnal Pendidikan Sosial & Humaniora*, *3*(2), 780–789. https://doi.org/10.61104/jq.v3i2.1070

Al Hakim, R., Mustika, I., & Yuliani, W. (2021). Validitas Dan Reliabilitas Angket Motivasi Berprestasi. *FOKUS (Kajian Bimbingan & Konseling dalam Pendidikan)*, *4*(4), 263–268. https://doi.org/10.22460/fokus.v4i4.7249

Anggraini, F. D. P., Aprianti, A., Setyawati, V. A. V., & Hartanto, A. A. (2022). Pembelajaran Statistika Menggunakan Software SPSS untuk Uji Validitas dan Reliabilitas. *Jurnal Basicedu*, *6*(4), 6492. https://doi.org/10.31004/basicedu.v6i4.3206

Copyright © 2026 Shoviyyatuz Zahro, et al.
(cc) BY

Anshari, M. I., Nasution, R., Irsyad, M., Alifa, A. Z., & Zuhriyah, I. A. (2024). Analisis Validitas dan Reliabilitas Butir Soal Sumatif Akhir Semester Ganjil Mata Pelajaran PAI. *EDUKATIF : JURNAL ILMU PENDIDIKAN*, *6*(1), 964–975. https://doi.org/10.31004/edukatif.v6i1.5931

Azizah, N., & Chalimatusadiah, C. (2025). Uji Validitas dan Uji Reliabilitas Instrumen Penelitian Pemahaman Konsep Dasar Aljabar. *Jurnal Pendidikan Tambusai*, *9*(1), 6637–6643. https://doi.org/10.31004/jptam.v9i1.25533

Buka, N., Jayamin, M. A., Kholis, N., Nursalam, & Nur Akbar Rasyid. (2025). Peran Validitas dan Rehabilitas Dalam Evaluasi Pembelajaran. *Socius: Jurnal Penelitian Ilmu-Ilmu Sosial*, *2*(11), 225. https://doi.org/10.5281/ZENODO.15557188

Efendi, Y. (2019). Uji Validitas dan Reliabilitas Instrumen Tes Shooting Sepak Bola pada Pemain Tim Persiwu FC Jatiyoso. *Jurnal Kesehatan Olahraga*, *07*(02), 368–369.

Eliyah. (2019). Validitas Tes Sebagai Media Ukur Ketepatan Evaluasi Mengukur Pencapaian Tujuan Pembelajaran Di SD/MI. *Jurnal Edukatif*, *V*(1).

Erviana, V. Y. (2023). The Validity and Reliability of IoJEPD Model Instruments to Improve Elementary School Teacher Competence: *Pegem Journal of Education and Instruction*, *13*(2), 200–206. https://doi.org/10.47750/pegegog.13.02.24

Hasanah, U., Ilham Ilham, Syarifuddin Syarifuddin, & Luthfiyah Luthfiyah. (2026). Pengembangan Instrumen Evaluasi Pembelajaran: Menggunakan pendekatan penelitian pengembangan (Research and Development/R&D). *SOKO GURU: Jurnal Ilmu Pendidikan*, *5*(3), 506. https://doi.org/10.55606/sokoguru.v5i3.6712

Idrus L. (2019). Evaluasi dalam Proses Pembelajaran. *Adara: Jurnal Manajemen Pendidikan Islam*, *9*(2), 920. https://doi.org/https://doi.org/10.35673/ajmpi.v9i2.427

Ihsan, H. (2015). Validitas Isi Alat Uukur Penelitian: Konsep Dan Panduan Penilaiannya. *PEDAGOGIA Jurnal Ilmu Pendidikan*, *13*, 173. https://doi.org/10.17509/pedagogia.v13i3.6004

Janna, N. M., & Herianto. (2021a). *Konsep Uji Validitas Dan Reliabilitas Dengan Menggunakan SPSS*. https://doi.org/10.31219/osf.io/v9j52

Janna, N. M., & Herianto, H. (2021b). *Konsep Uji Validitas Dan Reliabilitas Dengan Menggunakan SPSS*. Open Science Framework. https://doi.org/10.31219/osf.io/v9j52

Maulana, A. (2022). Analisis Validtas, Reliabilitas, dan Kelayakan Instrumen Penilaian Rasa Percaya Diri Siswa. *Jurnal Kualita Pendidikan*, *3*(3), 134. https://doi.org/10.51651/jkp.v3i3.331

Niam, M. F., Rumahlewang, E., Umiyati, H., Dewi, N. P. S., Atiningsih, S., Haryati, T., Magfiroh, I. S., Anggraini, R. I., Mamengko, R. P., Fathin, S., Mola, M. S. R., Syaifudin, A. A., & Wajdi, F. (2024). *Metode Penelititan Kulaitatif*. Widina Media Utama.

Nurhan Buka, M. A. J. (2025a). Peran Validitas dan Rehabilitas Dalam Evaluasi Pembelajaran. *Socius: Jurnal Penelitian Ilmu-Ilmu Sosial*, *02*(11), 225. https://doi.org/10.5281/ZENODO.15557188

Nurhan Buka, M. A. J. (2025b). Peran Validitas dan Rehabilitas Dalam Evaluasi Pembelajaran. *Socius: Jurnal Penelitian Ilmu-Ilmu Sosial*, *02*(11), 226. https://doi.org/10.5281/ZENODO.15557188

Purwanti, R., Ummah, M., Pulungan, A. M. K., Zulpan, Syafi, A., Asmoro, E. I., & Wulandari, M. N. (2026). Analisis Konseptual Validitas: Instumen , Jenis-Jenis dan Contohnya Dalam Penelitian Pendidikan. *Teaching and Learning Research Journal*, *2*(1). https://jurnal.permataedukasi.com/TLRJ/article/view/99/93

Ramadhan, M. F., Siroj, R. A., & Afgani, M. W. (2024). Validitas and Reliabilitas. *Journal on Education*, *06*(02), 10969.

Reza, I. A., Alfi, M. A., & Anam, S. (2026). validitas dan reliabilitas dalam penelitian kuantitatif. *Psikosospen : Jurnal Psikososial dan Pendidikan*, *2*(1), 830.

Sanaky, M. M., Saleh, L. M., & Titaley, H. D. (2021a). Analisis Faktor-Faktor Penyebab Keterlambatan pada Proyek Pembangunan Gedung Asrama MAN 1 Tulehu Maluku Tengah. *Jurnal Simetrik*, *11*(1), 433. https://doi.org/https://doi.org/10.31959/js.v11i1.615

Sanaky, M. M., Saleh, L. M., & Titaley, H. D. (2021b). Analisis Faktor-Faktor Penyebab Keterlambatan Pada Proyek Pembangunan Gedung Asrama MAN 1 Tulehu Maluku Tengah. *Jurnal Simetrik*, *11*(1).

Saputri, H. A. S., Zulhijrah, Larasati, N. J., & Shaleh. (2023). Analisis Instrumen Assesmen: Validitas, Reliabilitas, Tingkat Kesukaran Dan Daya Beda Butir Soal. *Didaktik : Jurnal Ilmiah PGSD STKIP Subang*, *9*(5), 2986–2995. https://doi.org/10.36989/didaktik.v9i5.2268

Sari Siagian, B. & Meyniar Albina. (2025). Konsep, Jenis, dan Penyusunan Instrumen Penelitian dalam Pendidikan. *QAZI : Journal of Islamic Studies*, *2*(1), 252. https://doi.org/10.61104/qz.v2i1.285

Setiawan, R., Wagiran, W., & Alsamiri, Y. (2024). Construction of an instrument for evaluating the teaching process in higher education: Content and construct validity. *REID (Research and Evaluation in Education)*, *10*(1), 50–63. https://doi.org/10.21831/reid.v10i1.63483

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

Sugiono, Noerdjanah, & Wahyu, A. (2020). Uji Validitas dan Reliabilitas Alat Ukur SG Posture Evaluation. *Jurnal Keterapian Fisik*, *5*(1), 55. https://doi.org/10.37341/jkf.v5i1.167

Sugiyono. (2013). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D* (Cetakan ke-19). Alfabeta.

Syaifudin. (2020). Validitas dan Reliabilitas Instrumen Penilaian Pada Mata Pelajaran Bahasa Arab. *Jurnal Kajian Perbatasan Antarnegara*, *3*(2).

Yusup, F. (2018a). Uji Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif. *Tarbiyah : Jurnal Ilmiah Kependidikan*, *7*(1). https://doi.org/10.18592/tarbiyah.v7i1.2100

Yusup, F. (2018b). Uji Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif. *Tarbiyah : Jurnal Ilmiah Kependidikan*, *7*(1). https://doi.org/10.18592/tarbiyah.v7i1.2100

Copyright © 2026 Shoviyyatuz Zahro, et al.
(cc) BY

Zayrin, A. A., Nupus, H., Maizia, K. K., Marsela, S., Hidayatullah, R., & Harmonedi, H. (2025). Analisis Instrumen Penelitian Pendidikan (Uji Validitas Dan Relibilitas Instrumen Penelitian). *QOSIM : Jurnal Pendidikan Sosial & Humaniora*, *3*(2), 780–789. https://doi.org/10.61104/jq.v3i2.1070

Zed, M. (2008). *Metode Penelitian Kepustakaan*. Yayasan Pustaka Obor Indonesia.