

Evaluasi Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif: Sebuah Studi Pustaka

Putu Gede Subhaktiyasa
STIKES Wira Medika Bali, Indonesia

✉ Corresponding author
[pgs@stikeswiramedika.ac.id]

Abstrak

Validitas dan reliabilitas terbukti sebagai dua aspek penting yang menentukan kualitas instrumen dalam mengukur variabel penelitian, namun masih dibutuhkan konsistensi dalam penerapannya. Penelitian ini bertujuan untuk memberikan pemahaman terkait uji validitas dan reliabilitas yang perlu dilakukan untuk menghasilkan instrumen penelitian yang baik. Studi merupakan penelitian kualitatif melalui pendekatan studi pustaka. Temuan penelitian menunjukkan bahwa perubahan dalam konteks sosial dan budaya dapat mempengaruhi keandalan instrumen yang telah teruji di masa lalu. Studi menekankan pentingnya pengujian ulang instrumen melalui beberapa uji validitas dan reliabilitas untuk memastikan akurasi pengumpulan data. Hasil penelitian berkontribusi dalam pengembangan dan pengujian instrumen yang lebih efektif untuk meningkatkan kredibilitas hasil penelitian.

Kata Kunci: *validitas isi; validitas konstruk; validitas kriteria; reliabilitas; instrumen penelitian*

Abstract

Validity and reliability are two crucial aspects determining the quality of instruments in measuring research variables; however, consistency in their application is still required. This study aims to understand the validity and reliability tests that should be conducted to produce quality research instruments. The research is qualitative research through a literature study approach. The findings show that changes in social and cultural contexts can affect the reliability of instruments tested in the past. The study emphasizes the importance of retesting instruments through multiple validity and reliability tests to ensure data collection accuracy. The study's results contribute to developing and testing more effective instruments to increase the credibility of research results.

Keyword: *content validity; construct validity; criterion validity; reliability; research instrument*

PENDAHULUAN

Instrumen penelitian kuantitatif berperan penting dalam pengumpulan data dan berfungsi sebagai alat ukur untuk mengamati, mengidentifikasi, dan menganalisis variabel-variabel penelitian (Kittur, 2023). Kualitas instrumen ini menentukan keakuratan, relevansi, dan representativitas data yang dihasilkan, yang berdampak pada kredibilitas dan validitas hasil penelitian (Cisneros-Barahona et al., 2023). Oleh karena itu, uji validitas dan reliabilitas menjadi tahapan penting untuk memastikan instrumen mampu memberikan hasil konsisten dengan tujuan pengukuran.

Validitas dan reliabilitas merupakan konsep utama dalam menilai kualitas instrumen penelitian kuantitatif (Peeters & Harpe, 2020). Validitas merujuk pada sejauh mana instrumen mampu mengukur konsep yang seharusnya diukur, sehingga hasil yang diperoleh benar-benar mencerminkan realitas yang diteliti. Sedangkan reliabilitas merujuk pada konsistensi hasil yang diperoleh instrumen ketika digunakan pada pengukuran yang berbeda dalam kondisi yang serupa. Dengan demikian, uji validitas dan reliabilitas mampu meminimalkan risiko kesalahan pengumpulan data dan menjadi dasar yang kuat untuk pengambilan kesimpulan dan generalisasi temuan penelitian (Heale & Twycross, 2015).

Studi sebelumnya telah menekankan pentingnya pengujian instrumen dalam penelitian kuantitatif (Bushmakina & Cappelleri, 2022; Ediyanto et al., 2022; Peeters & Harpe, 2020; Ramadhan et al., 2024; Salmia, 2023). Namun, masih terdapat tantangan dalam penerapan uji validitas dan

reliabilitas secara nyata. Instrumen yang telah digunakan secara luas dan telah melalui uji validitas serta reliabilitas cenderung tidak diuji ulang. Sedangkan studi lainnya menunjukkan pentingnya pengujian ulang sebagai upaya untuk mempertahankan kredibilitas dan validitas hasil penelitian. Oleh karena itu diperlukan pemahaman komprehensif dan kontekstual dalam penerapan uji validitas dan reliabilitas instrumen penelitian.

Penelitian ini menegaskan pentingnya validitas dan reliabilitas dalam penelitian kuantitatif serta memperjelas panduan metodologis terkait penggunaan instrumen yang telah diuji sebelumnya. Tujuan utama adalah memberikan pemahaman penerapan uji instrumen yang tepat dalam penelitian kuantitatif. Hasil penelitian diharapkan dapat menjadi referensi untuk peneliti dalam pengujian instrumen, serta memberikan kontribusi signifikan dalam meningkatkan kualitas penelitian kuantitatif.

METODE PENELITIAN

Penelitian ini merupakan penelitian kualitatif dengan desain studi pustaka. Pendekatan ini dipilih karena bertujuan untuk mengeksplorasi dan menganalisis konsep-konsep yang berkaitan dengan uji validitas dan reliabilitas instrumen dalam penelitian kuantitatif. Studi pustaka berfokus pada pengumpulan dan analisis kritis terhadap literatur yang relevan, termasuk artikel jurnal, buku, dan publikasi ilmiah lainnya (Turcios et al., 2012). Evaluasi data dilakukan secara sistematis guna memberikan pandangan yang lebih komprehensif tentang pengujian instrumen dalam penelitian kuantitatif seperti yang dilakukan Subhaktiyasa, (2024b). Tinjauan terhadap sumber-sumber akademis mencakup kajian tentang validitas isi, validitas konstruk, validitas kriteria, serta berbagai bentuk uji reliabilitas, termasuk konsistensi internal dan stabilitas. Data yang dikumpulkan dianalisis dengan analisis konten untuk mengidentifikasi penerapan uji validitas dan reliabilitas.

HASIL DAN PEMBAHASAN

Konsep Dasar Instrumen Penelitian Kuantitatif

Instrumen dalam penelitian kuantitatif merupakan alat atau perangkat yang digunakan untuk mengumpulkan data dalam penelitian yang bersifat numerik dan dapat dianalisis secara statistik (Little et al., 2024; Salmia, 2023). Instrumen berfungsi untuk mengukur variabel-variabel penelitian secara sistematis dan objektif untuk memperoleh data yang terukur, konsisten, dan dipercaya dalam menganalisis fenomena penelitian yang sesuai dengan tujuan penelitian dan hipotesis yang diajukan sehingga dapat digeneralisasikan pada populasi yang lebih luas. Variabel penelitian itu sendiri merupakan entitas yang dapat diidentifikasi, diukur, dan dianalisis untuk memahami hubungan, pengaruh, atau pola dalam sebuah studi. Variabel-variabel ini dapat berupa karakteristik fisik, perilaku, atau atribut lain yang dapat dikonversi ke dalam bentuk data numerik untuk dianalisis secara statistik (Fávero et al., 2023; Nermend, 2023). Oleh karena itu, peneliti perlu mengidentifikasi variabel dalam instrumen penelitian. Proses ini penting karena peneliti dapat merancang instrumen penelitian yang tepat dengan memberikan definisi operasional variabel yang jelas untuk memastikan bagaimana variabel tersebut dapat diukur. Definisi operasional variabel merupakan deskripsi konkret dan spesifik yang menjelaskan bagaimana suatu variabel diukur atau diamati dalam penelitian (Mthembu, 2018; Slife et al., 2016). Variabel yang abstrak dapat diterjemahkan ke dalam bentuk yang dapat diukur secara empiris, sehingga memungkinkan pengumpulan data yang akurat, analisis yang tepat, serta replikasi penelitian yang valid. Definisi operasional membantu menghubungkan konsep teoretis dengan realitas penelitian melalui indikator yang relevan dan terukur.

Secara umum, terdapat beberapa jenis variabel yang digunakan dalam penelitian kuantitatif yaitu variabel independen, variabel dependen, variabel kontrol, variabel moderator dan variabel mediasi (Andrade, 2021; Flannelly et al., 2014; Rossi & Merenda, 2010). Variabel independen merupakan variabel yang dianggap sebagai faktor penyebab yang dimanipulasi dalam penelitian untuk mengamati pengaruhnya terhadap variabel lain. Variabel dependen adalah variabel yang dipengaruhi oleh variabel independen dan merupakan hasil dari manipulasi terhadap variabel independent. Variabel kontrol adalah variabel lain yang berpotensi memengaruhi variabel dependen, namun dikendalikan agar tidak memengaruhi hubungan antara variabel independen dan dependen, guna memastikan keakuratan pengamatan hubungan tersebut. Variabel moderator adalah variabel yang memengaruhi kekuatan atau arah hubungan antara variabel independen dan dependen, dengan kemampuannya memperkuat, memperlemah, atau mengubah hubungan tersebut. Variabel mediasi

adalah variabel yang menjelaskan pengaruh variabel independen terhadap variabel dependen dengan berperan sebagai perantara dalam hubungan tersebut. Selain itu, terdapat variabel eksogen dan variabel endogen yang digunakan dalam penelitian berbasis model struktural atau analisis jalur dan model persamaan struktural (Jin et al., 2021; Subhaktiyasa, 2024b). Meskipun variabel tersebut memiliki konsep yang mirip dengan variabel independen dan variabel dependen, namun terdapat beberapa perbedaan dalam penggunaannya, terutama dalam konteks analisis hubungan kausal. Variabel Independen dan Dependen umumnya digunakan dalam konteks penelitian sederhana dengan satu hubungan kausal langsung, sedangkan Variabel Eksogen dan Endogen digunakan dalam konteks model penelitian dengan struktur hubungan kausal yang lebih kompleks, lihat (Candrawati et al., 2024; Subhaktiyasa et al., 2023; Subhaktiyasa, Agung, et al., 2024; Subhaktiyasa, Sintari, et al., 2024; Subhaktiyasa, Sutrisna, et al., 2024; Subhaktiyasa & Sintari, 2024). Variabel eksogen adalah variabel yang berasal dari luar model penelitian dan tidak dipengaruhi oleh variabel lain dalam model penelitian, berfungsi sebagai faktor penyebab yang mempengaruhi variabel lain. Sedangkan variabel endogen adalah variabel yang dipengaruhi oleh variabel lain dalam model penelitian dan dapat pula mempengaruhi variabel lainnya, berfungsi sebagai variabel hasil dalam hubungan kausal.

Instrumen penelitian tidak hanya digunakan untuk mengukur variabel, tetapi juga melibatkan penggunaan skala pengukuran yang sesuai untuk menilai setiap variabel penelitian. Skala pengukuran dapat berupa skala nominal, ordinal, interval dan skala rasio yang masing-masingnya menghasilkan jenis data nominal, ordinal, interval dan skala rasio (Fávero et al., 2023; Miksa et al., 2023). Skala-skala tersebut bersifat matematis dan digunakan untuk mengukur sifat-sifat objektif yang dapat diukur secara langsung. Skala nominal termasuk dalam pengukuran kualitatif dan merupakan jenis skala pengukuran paling sederhana yang digunakan untuk mengklasifikasikan data ke dalam kategori yang berbeda tanpa urutan atau hirarki. Skala ordinal adalah jenis skala pengukuran yang digunakan untuk mengurutkan data berdasarkan peringkat atau tingkatan, namun jarak antara setiap kategori tidak selalu sama atau nilai intervalnya tidak diketahui. Skala interval adalah jenis skala pengukuran yang memiliki urutan yang jelas dan jarak yang sama antara setiap kategori dan memiliki makna yang konsisten serta bersifat arbitrer. Skala rasio adalah jenis skala pengukuran yang paling lengkap dan informatif di antara keempat skala pengukuran yang umum digunakan dalam statistik. Skala ini tidak hanya memiliki urutan dan jarak yang sama antara kategori, tetapi juga memiliki titik nol mutlak sehingga memungkinkan untuk melakukan semua jenis perhitungan dan menarik kesimpulan pasti.

Selain skala-skala tersebut, terdapat skala dalam instrumen penelitian yang digunakan untuk mengukur konsep-konsep abstrak yang tidak dapat diukur secara langsung dan berfokus pada konstruksi psikologis seperti sikap, persepsi, atau pengalaman (Dalati, 2018; Fenn et al., 2020). Skala tersebut diantaranya adalah skala Likert, Guttman, Rating Scale, Semantic Differential, dan skala Thurstone (Agung, 2018; Hidayat, 2021; Sugiono, 2021). Skala Likert paling umum digunakan dalam penelitian untuk mengukur sikap, pendapat dan persepsi seseorang terhadap suatu pernyataan tentang fenomena sosial yang telah ditetapkan secara spesifik sebagai variabel penelitian. Tingkat persetujuan atau ketidaksetujuan pernyataan sebaiknya disajikan dalam bentuk ganjil dengan skala 5 atau 7 poin untuk meningkatkan validitas dan reliabilitas data. Menggunakan skala 4 poin dengan menghilangkan pilihan jawaban alternatif tengah terbukti dapat menurunkan validitas respon dan dapat meningkatkan varians kesalahan acak dalam respon (Wang & Krosnick, 2020). Skala Guttman berfungsi untuk mengukur sikap atau perilaku dengan cara mengurutkan item-item yang berkaitan dengan suatu konstruk dari yang paling mudah hingga yang paling sulit. Syarat penggunaan skala ini adalah bahwa item-item yang disusun harus saling berhubungan secara hierarkis, sehingga menciptakan sebuah skala yang mencerminkan peningkatan kompleksitas. Penggunaan skala Guttman akan menghasilkan jawaban yang ketat dan tegas seperti iya-tidak, setuju-tidak setuju, pernah-tidak pernah, positif-negatif, atau lainnya yang bersifat dikotomi (dua alternatif). Semantic Differential adalah skala yang digunakan untuk mengukur makna yang diberikan responden terhadap suatu konsep dengan menggunakan dua kata yang berlawanan dan tersusun dalam garis kontinum, seperti " baik " dan " buruk. " sikap yang dimiliki responden terhadap konsep tertentu. Rating Scale adalah jenis skala yang memungkinkan responden untuk memberikan penilaian terhadap objek atau pernyataan dengan menetapkan nilai tertentu dalam bentuk angka. Penggunaan skala ini memerlukan kriteria yang jelas mengenai aspek yang dinilai, sehingga dapat memastikan konsistensi dalam pengukuran. Skala Thurstone adalah skala yang dirancang untuk mengukur intensitas sikap seseorang

terhadap suatu objek atau isu tertentu menggunakan metode penilaian interval. Skala ini berfungsi untuk memberikan bobot obyektif pada berbagai pernyataan yang mewakili sikap, memungkinkan peneliti untuk tidak hanya mengukur arah sikap (positif atau negatif), tetapi juga intensitasnya. Skala ini sering digunakan dalam penelitian sosial dan psikologi untuk memperoleh data yang lebih terukur dan terstruktur mengenai sikap individu atau kelompok. Walaupun skala Thurstone menawarkan tingkat akurasi yang tinggi dalam pengukuran sikap, proses pengembangannya dianggap kompleks dan memakan waktu karena melibatkan penilaian dari panel ahli untuk menentukan bobot setiap pernyataan.

Penyusunan dan Pengembangan Instrumen Penelitian

Instrumen penelitian berfungsi untuk mengukur variabel-variabel penelitian secara sistematis dan objektif, sehingga peneliti dapat memperoleh data yang akurat, konsisten, dan dapat diandalkan dalam menganalisis fenomena penelitian. Kualitas instrumen sangat menentukan apakah data yang diperoleh mencerminkan realitas atau fenomena yang diteliti (Turcios & Watkins, 2012). Instrumen baku yang telah teruji validitas dan reliabilitasnya dapat digunakan kembali untuk penelitian selanjutnya dengan mempertimbangkan kesesuaian instrumen dengan landasan teori serta konstruk variabel yang akan diukur dalam penelitian tersebut (Liu et al., 2023; Selvi & Özge, 2023; Winkler & Berenbon, 2021). Apabila instrumen baku tersebut tidak tersedia, maka perlu melakukan pengembangan instrumen baru melalui serangkaian langkah sistematis yang berbasis pada sintesis teori (Amir Behghadami et al., 2019; Ediyanto et al., 2022; Pittman & Bakas, 2010; Thomas, 2022). Langkah awal adalah dengan perumusan konstruk variabel yang dirumuskan berdasarkan landasan teori yang kuat sehingga menjadi dasar bagi pengembangan dimensi dan indikator variabel. Selanjutnya, peneliti menyusun kisi-kisi instrumen yang mencakup dimensi, indikator, dan jumlah butir pertanyaan atau pernyataan yang berfungsi sebagai peta atau panduan dalam penyusunan butir instrumen untuk mengukur setiap aspek dari konstruk variabel. Kemudian peneliti menetapkan rentang skala pengukuran. Skala yang umum digunakan adalah skala Likert, yang mengukur persepsi responden dalam rentang tertentu, misalnya dari sangat tidak setuju hingga sangat setuju. Butir instrumen disusun dalam bentuk pernyataan positif dan negatif untuk menghindari bias jawaban akibat pola respon yang sama. Langkah berikutnya adalah validasi teoritik, yang dilakukan melalui diskusi dengan pakar atau panel ahli (*expert judgment*) yang bertujuan untuk memastikan kesesuaian antara konstruk, dimensi, indikator, dan butir yang dikembangkan. Saran dari pakar kemudian digunakan untuk merevisi instrumen.

Setelah melalui proses validasi teoritik, instrumen perlu diuji coba secara empiris pada sejumlah sampel untuk mengumpulkan data (Eisend & Kuss, 2019). Menentukan jumlah sampel minimum yang diperlukan untuk pengujian empiris bervariasi secara signifikan di berbagai bidang dan metodologi. Ukuran sampel harus cukup untuk memastikan akurasi dan keandalan dalam hasil (Subhaktiyasa, 2024a). Sugiono, (2021) menyebutkan pengujian dapat menggunakan 30 sampel. Namun pengujian yang bertujuan utama untuk menguji validitas konstruk melalui analisis faktor disarankan menggunakan sampel lebih dari 100 responden (Anthoine et al., 2014; Clark & Watson, 2019; Holton III et al., 1997; Thompson & Daniel, 1996). Sampel untuk uji coba instrumen tersebut dapat berasal dari populasi yang berbeda dengan populasi utama dengan memperhatikan karakteristik sampel yang sebanding dengan populasi utama (Cha et al., 2016; Stuart et al., 2017). Ini penting untuk memastikan bahwa hasil uji instrumen dapat digeneralisasi ke populasi utama. Namun apabila sampel diambil dari populasi utama, maka sampel uji coba instrumen penelitian harus berbeda dengan sampel yang digunakan untuk penelitian utama agar menghindari bias "efek pengujian" (*testing effect*), di mana subjek yang pernah terpapar instrumen mungkin memberikan jawaban yang berbeda pada pengumpulan data utama (Pelletier & Sharma, 2015). Data yang terkumpul dari sampel yang menjadi percobaan terhadap instrumen penelitian dianalisis untuk menguji validitas dan reliabilitas instrumen. Hasil dari uji validitas dan reliabilitas ini akan menentukan apakah instrumen dapat digunakan dalam penelitian utama. Meskipun langkah-langkah tersebut memberikan pendekatan terstruktur, tantangan seperti karakteristik sampel dan keterbatasan metodologis dapat berdampak pada proses pengembangan instrumen.

Teknik Pengumpulan Data

Kualitas instrumen penelitian sangat dipengaruhi validitas dan reliabilitasnya, sedangkan kualitas pengumpulan data dapat ditentukan dari ketepatan kondisi alamiah, sumber data baik data primer atau sekunder, prosedur yang jelas, etika penelitian, dan teknik pengumpulan data seperti kuesioner, wawancara, observasi atau gabungan ketiganya yang disesuaikan dengan tujuan penelitian (Agung, 2018; Sugiono, 2021). Kuesioner merupakan alat pengumpulan data yang berbentuk serangkaian pertanyaan tertulis yang harus diisi oleh responden. Instrumen ini digunakan untuk mendapatkan informasi yang terstruktur dan terukur terkait fenomena yang sedang diteliti. Kuesioner sering digunakan karena efisien untuk mengumpulkan data dari sampel yang besar, serta mudah dianalisis dengan pendekatan statistik. Desain kuesioner yang baik harus memperhatikan penyusunan pertanyaan yang jelas, valid, dan reliabel agar dapat mengukur variabel yang diteliti secara akurat. Wawancara terdiri dari wawancara terstruktur dan wawancara tidak terstruktur. Meskipun wawancara sering dianggap sebagai metode pengumpulan data kualitatif, namun dapat digunakan dalam penelitian kuantitatif secara terstruktur dengan menggunakan pedoman wawancara yang telah disusun sebelumnya. Wawancara terstruktur melibatkan serangkaian pertanyaan yang sama diajukan kepada responden, sehingga memungkinkan data yang dikumpulkan lebih konsisten dan dapat dianalisis secara kuantitatif. Kelebihan wawancara dibanding kuesioner adalah peneliti dapat melakukan klarifikasi langsung jika ada pertanyaan yang tidak dipahami responden, serta menggali informasi lebih dalam jika diperlukan. Observasi terdiri dari observasi partisipan dan non partisipan yang dibagi menjadi observasi terstruktur dan tidak terstruktur. Observasi dalam penelitian kuantitatif biasanya dilakukan dengan menggunakan instrumen yang telah distandarkan, seperti lembar observasi, untuk mencatat perilaku atau fenomena yang diamati secara objektif. Teknik ini memungkinkan peneliti untuk mendapatkan data empiris langsung dari situasi atau responden yang sedang diamati. Observasi sering digunakan untuk mengukur variabel yang terkait dengan perilaku, tindakan, atau situasi yang tidak dapat dijelaskan secara verbal oleh responden. Observasi yang baik harus bersifat sistematis dan menggunakan kategori pengukuran yang jelas agar data yang dikumpulkan dapat dianalisis secara statistik.

Konsep Uji Validitas dan Reabilitas Instrumen Penelitian Kuantitatif

Validitas dalam penelitian kuantitatif merujuk pada tingkat keakuratan dan ketepatan sebuah instrumen dalam mengukur apa yang seharusnya diukur. Sebuah instrumen dikatakan valid jika data yang dihasilkannya mencerminkan konsep atau fenomena yang diukur secara tepat (Bushmakina & Cappelleri, 2022; Heale & Twycross, 2015; Lane & Pérez, 2023). Terdapat beberapa jenis validitas yang umum digunakan dalam penelitian kuantitatif, di antaranya validitas isi, validitas konstruk, dan validitas kriteria. Validitas isi berkaitan dengan sejauh mana item-item dalam instrumen secara komprehensif mencakup semua aspek konsep yang diukur. Validitas isi menjawab seberapa ketepatan isi instrumen, apakah instrumen sudah mengukur variabel yang diukur, seberapa ketepatan butir tes mewakili sampel materi dan seberapa ketepatan format instrumen. Isi dan format instrumen harus konsisten dengan definisi variabel dan sampel materi yang akan diukur. Hasil penilaian pakar terhadap validitas isi umumnya bersifat kualitatif. Gregory (2015) mengembangkan teknik pengujian validitas isi bersifat kuantitatif namun tetap menggunakan penilaian dua pakar yang hasil penilaiannya telah dikuantitatifkan. Mekanisme perhitungannya dimulai dengan melakukan penilaian terhadap instrumen perbutir menggunakan skala yang kemudian dikelompokkan menjadi kelompok skala kurang relevan dan kelompok skala sangat relevan. Hasil penilaian para pakar ditabulasi dalam bentuk matriks, ditabulasi silang dan kemudian dihitung untuk mendapatkan nilai kategori validitas isi. Lawshe (1975) juga memberikan panduan dalam melakukan validitas isi instrumen dimana para pakar memberikan penilaian terhadap instrumen secara kuantitatif dengan menilai setiap butir menggunakan kriteria 1 = butir esensial untuk mengukur indikator, 2 = butir diperlukan untuk mengukur indikator tetapi tidak esensial, dan 3 = butir tidak diperlukan untuk mengukur indikator. Hasil penilaian pakar dianalisis menggunakan analisis Lawshe ditabulasi untuk mendapatkan nilai *Content Validity Ratio* (CVR) sama dan lebih besar dari CVR tabel pada taraf signifikansi 0,05. Selain uji validitas tersebut, terdapat uji validitas muka yang mengacu pada penilaian subjektif peneliti terhadap penyajian dan relevansi alat ukur, apakah item-item dalam instrumen tampak relevan, masuk akal, tidak ambigu, dan jelas (Oluwatayo, 2012). Validitas ini merupakan penilaian subjektif pada

operasionalisasi konstruk yang mengukur sejauh mana suatu ukuran tampak terkait dengan konstruk tertentu, dalam penilaian orang yang tidak ahli. Oleh karena itu validitas muka dianggap sebagai bentuk validitas yang paling lemah dan dianggap bukan sebagai uji validitas sebenarnya.

Selanjutnya dilakukan validitas konstruk yang mengacu pada sejauh mana instrumen dapat mengukur konsep teoretis yang mendasarinya, memastikan bahwa instrumen tersebut benar-benar menangkap esensi dari variabel abstrak yang diukur (Lane & Pérez, 2023). Validitas ini sering kali diperoleh melalui analisis faktor baik dengan pendekatan eksploratori maupun konfirmatori atau pengujian hipotesis teoretis. Terakhir adalah validitas kriteria yang mengukur hubungan antara hasil instrumen dengan ukuran atau kriteria eksternal yang relevan, baik yang bersifat konkuren (pengukuran pada waktu bersamaan) maupun prediktif (pengukuran di masa depan) (Bushmakini & Cappelleri, 2022). Pengujian validitas kriteria dapat menggunakan uji korelasi Pearson dimana jika nilai korelasi mendekati +1 menunjukkan korelasi positif yang kuat, mendekati -1 menunjukkan korelasi negatif yang kuat dan jika mendekati 0 maka menunjukkan tidak ada korelasi yang signifikan.

Instrumen yang telah terbukti memiliki validitas yang baik dilanjutkan dengan melakukan uji reliabilitas. Reliabilitas dalam penelitian kuantitatif merujuk pada konsistensi dan kestabilan hasil yang diperoleh dari suatu instrumen ketika digunakan dalam kondisi yang serupa di waktu yang berbeda. Sebuah instrumen yang reliabel akan memberikan hasil yang sama atau sangat mirip jika digunakan untuk mengukur hal yang sama dalam kondisi yang serupa (Heale & Twycross, 2015; Oluwatayo, 2012; Ramadhan et al., 2024; Taherdoost, 2017). Terdapat beberapa cara untuk menguji reliabilitas, yaitu: reliabilitas stabilitas (*test-retest reliability*), reliabilitas konsistensi internal (*internal consistency reliability*), reliabilitas bentuk paralel (*parallel-forms reliability*), dan reliabilitas antar penilai (*inter-rater reliability*). Reliabilitas test-retest mengukur stabilitas hasil pengukuran dari waktu ke waktu. Jika sebuah instrumen diuji pada dua kesempatan yang berbeda dengan jeda waktu tertentu, dan hasilnya tetap konsisten, maka instrumen tersebut memiliki reliabilitas yang tinggi. Teknik yang digunakan untuk pengujian ini diantaranya adalah Korelasi Pearson, dan Intraclass Correlation Coefficient (ICC) dengan nilai korelasi $\geq 0,70$ menunjukkan reliabilitas yang baik. Reliabilitas konsistensi internal mengukur sejauh mana item-item dalam satu instrumen saling berkorelasi, yang menunjukkan bahwa item-item tersebut secara konsisten mengukur konstruk yang sama. Pengujian ini mengukur konsistensi internal dalam satu tes atau satu kali pengukuran. Beberapa teknik yang digunakan untuk menguji konsistensi internal adalah Cronbach's Alpha (α), Kuder-Richardson Formula 20 (KR-20), Kuder-Richardson Formula 21 (KR-21), Split-Half Reliability (Konsistensi Paruh Belah), Average Inter-Item Correlation (Korelasi Antar Item Rata-rata) dan Analisis Varian Hoyt (Hoyt's Analysis of Variance). *Cronbach's Alpha* mengukur seberapa konsisten setiap butir dalam instrumen pengukuran dengan keseluruhan instrumen. Nilai *Cronbach's Alpha* berkisar antara 0 hingga 1, dengan nilai $> 0,70$ mengindikasikan bahwa indikator dalam instrumen tersebut memiliki konsistensi internal yang baik. Tes ini cenderung digunakan oleh peneliti untuk menguji instrumen yang menggunakan Skala Likert. Sedangkan instrumen dengan skala Guttman atau bersifat dikotomis cenderung menggunakan KR-20 atau KR-21. KR-20 lebih akurat jika tingkat kesulitan antar item bervariasi, sementara KR-21 memberikan estimasi yang lebih konservatif dengan asumsi varians yang seragam. *Split-Half Reliability* dilakukan dengan membagi instrumen menjadi dua bagian yang setara, seperti item ganjil dan genap, kemudian hasil kedua bagian tersebut dikorelasikan menggunakan rumus Spearman Brown. Jika hasil kedua bagian tersebut berkorelasi tinggi, maka reliabilitas instrumen dianggap baik. Sedangkan pada *Average Inter-Item Correlation* dilakukan dengan mengukur korelasi rata-rata antar item dalam satu instrumen. Jika korelasi rata-rata antar item tinggi, maka reliabilitas dianggap baik. Semakin tinggi korelasi antar item, semakin kuat hubungan antara item dan konstruk yang diukur. Analisis Varian Hoyt sering dihubungkan dengan konsep analisis varian dua arah (*two-way analysis of variance*) yang merupakan metode berbasis ANOVA (*Analysis of Variance*) untuk mengukur reliabilitas konsistensi internal, dengan membagi variansi total ke dalam beberapa komponen (variansi antar individu, antar item, dan error). Teknik ini bermanfaat terutama untuk instrumen pengukuran yang kompleks dengan banyak item, namun membutuhkan perhitungan yang lebih mendalam dibandingkan teknik lainnya dalam pengujian reliabilitas konsistensi internal.

Reliabilitas bentuk paralel yang disebut juga sebagai tes ekuivalen melibatkan penggunaan dua bentuk instrumen yang berbeda tetapi dirancang untuk mengukur konstruk yang sama. Metode ini menguji reliabilitas dengan cara memberikan dua versi instrumen tersebut pada waktu yang sama

atau dalam waktu berbeda kepada kelompok yang sama. Reliabilitas dinilai dari korelasi antara dua bentuk instrumen tersebut melalui Koefisien Korelasi Pearson. Terakhir adalah reliabilitas antar penilai yang digunakan untuk mengukur konsistensi hasil pengukuran terhadap penilaian yang bersifat subjektif dan dilakukan oleh lebih dari satu penilai. Pengujian dilakukan dengan mengukur tingkat kesepakatan antara penilai yang berbeda menggunakan Kappa Statistic (Cohen's Kappa), atau Intraclass Correlation Coefficient (ICC) dengan nilai $\geq 0,70$ dianggap sebagai reliabilitas antar penilai yang baik.

Pengujian Ulang Validitas dan Reliabilitas Instrumen Baku

Instrumen baku yang telah teruji validitas dan reliabilitasnya secara empiris seringkali digunakan kembali dalam berbagai penelitian, terutama jika instrumen tersebut telah diakui secara luas sebagai alat ukur yang valid dan reliabel untuk mengukur konstruk tertentu (Taherdoost, 2017). Namun, meskipun instrumen baku memiliki rekam jejak yang baik, penting untuk mempertimbangkan apakah pengujian ulang validitas dan reliabilitas tetap diperlukan sebelum digunakan kembali. Keputusan untuk melakukan pengujian ulang bergantung pada beberapa faktor, seperti kesesuaian konteks penelitian, karakteristik populasi, ketelitian metodologis, dan kondisi yang sama dengan penelitian sebelumnya (Alcaraz-Ibáñez et al., 2022; Cruz et al., 2021; Stanger et al., 2020).

Penelitian sebelumnya menunjukkan walaupun memiliki tantangan yang kompleks, pengujian ulang instrumen baku mungkin tidak selalu dilakukan jika konteks sosial, budaya, dan populasi penelitian tetap konsisten dengan penelitian sebelumnya (Gross et al., 2018). Konsep konsisten tersebut bersifat relatif dan perlu mendapat definisi jelas untuk memberikan justifikasi yang kuat. Perkembangan ilmu pengetahuan dari waktu ke waktu berdampak pada perubahan populasi yang mempengaruhi validitas dan reliabilitas instrumen. Selain itu, adanya perubahan dalam definisi konseptual atau teoritis variabel yang diukur membutuhkan evaluasi ulang pada validitas teoritis instrumen. Ini menunjukkan bahwa meskipun instrumen baku dapat digunakan kembali namun *pilot study* tetap disarankan untuk mengevaluasi kembali instrumen sebelum digunakan pada skala yang lebih besar pada penelitian utama (Stocker et al., 2020). Keputusan untuk melakukan pengujian ulang harus didasarkan pada evaluasi terhadap beberapa faktor, termasuk kesesuaian instrumen dengan tujuan penelitian, karakteristik sampel, perubahan dalam literatur dan pertimbangan etis penelitian.

Pengujian ulang validitas dan reliabilitas instrumen baku dapat mencakup seluruh jenis uji yang biasa dilakukan pada instrumen baru, seperti validitas isi, validitas konstruk, dan validitas kriteria. Namun, dalam praktiknya, pengujian ini dapat disesuaikan dengan tujuan penelitian dan kondisi penggunaan instrumen (Efsthathiou, 2019). Jika instrumen baku digunakan untuk mengukur variabel-variabel yang sama dalam populasi yang memiliki karakteristik serupa dengan penelitian sebelumnya, pengujian ulang validitas kriteria dan reliabilitas konsistensi internal sering kali dianggap cukup. Validitas kriteria dapat diuji melalui korelasi antara hasil instrumen dan kriteria eksternal yang relevan. Sedangkan, validitas konstruk perlu diuji kembali jika instrumen digunakan untuk mengukur konstruk atau teori yang baru atau jika terdapat perubahan dalam definisi variabel yang diukur. Teori yang mendasari konstruk dapat menjadi pedoman dalam merancang dan mengevaluasi instrumen. Uji reliabilitas cenderung mencakup uji reliabilitas konsistensi internal yang diuji menggunakan Cronbach's Alpha untuk mengevaluasi sejauh mana item-item dalam skala memberikan hasil yang konsisten. Untuk pengujian ulang tersebut, sampel yang digunakan harus representatif dari populasi utama penelitian. Sampel dapat diambil dari populasi utama atau dari populasi dengan karakteristik yang tidak jauh berbeda, namun penggunaan sampel dari populasi utama membutuhkan ketelitian dengan menghindari penggunaan berulang untuk menghindari bias.

SIMPULAN

Studi menegaskan pentingnya uji validitas dan reliabilitas instrumen pada penelitian kuantitatif. Studi menunjukkan bahwa pengujian validitas secara umum dilakukan melalui uji validitas isi, konstruk dan kriteria. Sedangkan pengujian reliabilitas dilakukan melalui reliabilitas stabilitas, konsistensi internal, bentuk parallel, dan antar penilai. Penerapan uji validitas dan reliabilitas dilakukan pada instrumen baru dan juga instrumen baku meskipun instrumen baku tersebut telah teruji sebelumnya. Sampel yang digunakan untuk pengujian instrumen sebaiknya berasal dari populasi yang berbeda tetapi memiliki karakteristik serupa dengan populasi utama, atau dapat juga berasal dari populasi

utama namun tidak digunakan kembali dalam penelitian untuk menghindari bias. Walaupun hasil penelitian ini berkontribusi dalam memberikan pemahaman terkait pengujian validitas dan reliabilitas, penelitian selanjutnya perlu mengkaji penerapan uji validitas dan reliabilitas secara praktis khususnya eksplorasi dampak dari perubahan dalam konteks sosial dan budaya terhadap validitas dan reliabilitas instrumen baku. Hal ini diharapkan dapat memberikan pemahaman yang lebih komprehensif mengenai implementasi pengujian dalam penelitian.

DAFTAR PUSTAKA

- Agung, A. A. G. (2018). *Quantitative Research Methodology (Education Management Perspective)*. Universitas Pendidikan Ganesha.
- Alcaraz-Ibáñez, M., Paterna, A., Sicilia, Á., & Griffiths, M. D. (2022). Examining the reliability of the scores of self-report instruments assessing problematic exercise: A systematic review and meta-analysis. *Journal of Behavioral Addictions*, 11(2), 326–347. <https://doi.org/10.1556/2006.2022.00014>
- Amir Behghadami, M., Janati, A., Sadeghi-Bazargani, H., Gholizadeh, M., Rahmani, F., & Arab-Zozani, M. (2019). Developing and validating an instrument to assess non-hospital health centers' preparedness to provide initial emergency care: a study protocol. *BMJ Open*, 9(7), e026651. <https://doi.org/10.1136/bmjopen-2018-026651>
- Andrade, C. (2021). A Student's Guide to the Classification and Operationalization of Variables in the Conceptualization and Design of a Clinical Study: Part 1. *Indian Journal of Psychological Medicine*, 43(2), 177–179. <https://doi.org/10.1177/0253717621994334>
- Anthoine, E., Moret, L., Regnault, A., Sébille, V., & Hardouin, J.-B. (2014). Sample size used to validate a scale: a review of publications on newly-developed patient reported outcomes measures. *Health and Quality of Life Outcomes*, 12(1), 2. <https://doi.org/10.1186/s12955-014-0176-2>
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27–40. <https://doi.org/10.3316/QRJ0902027>
- Bushmakina, A. G., & Cappelleri, J. C. (2022). Construct Validity and Criterion Validity. In *A Practical Approach to Quantitative Validation of Patient-Reported Outcomes* (pp. 151–249). <https://doi.org/https://doi.org/10.1002/9781119376354.ch5>
- Candrawati, S. A. K., Sriani, N. K. A., Suabhaktiyasa, P. G., Andini, N. K. S., Putri, N. L. N. D. D., Citrawati, N. K., & Andriany, M. (2024). Spiritual Leadership Enhances Caring Behaviour: The Mediating Role of Calling. *Nurse Media Journal of Nursing*, 14(2). <https://doi.org/https://doi.org/10.14710/nmjn.v14i2.58964>
- Cha, S., Erar, B., Niaura, R. S., & Graham, A. L. (2016). Baseline Characteristics and Generalizability of Participants in an Internet Smoking Cessation Randomized Trial. *Annals of Behavioral Medicine*, 50(5), 751–761. <https://doi.org/10.1007/s12160-016-9804-x>
- Cisneros-Barahona, A. S., Marqués-Molíás, L., Samaniego-Erazo, N., Mejía-Granizo, C., & De la Cruz-Fernández, G. (2023). Multivariate data analysis: Validation of an instrument for the evaluation of teaching digital competence. *F1000Research*, 12(May), 1–24. <https://doi.org/10.12688/f1000research.135194.2>
- Clark, L. A., & Watson, D. (2019). Constructing validity: New developments in creating objective measuring instruments. *Psychological Assessment*, 31(12), 1412–1427. <https://doi.org/10.1037/pas0000626>
- Cruz, M. L., Bogaard, M. E. D. van den, Saunders-Smiths, G. N., & Groen, P. (2021). Testing the Validity and Reliability of an Instrument Measuring Engineering Students' Perceptions of Transversal Competency Levels. *IEEE Transactions on Education*, 64(2), 180–186. <https://doi.org/10.1109/TE.2020.3025378>
- Dalati, S. (2018). *Measurement and Measurement Scales BT - Modernizing the Academic Teaching and Research Environment: Methodologies and Cases in Business Research* (J. Marx Gómez & S. Mouselli (eds.); pp. 79–96). Springer International Publishing. https://doi.org/10.1007/978-3-319-74173-4_5
- Ediyanto, E., Sunandar, A., Ramadhani, R. S., & Aqilah, T. S. (2022). Sustainable Instrument Development in Educational Research. *Discourse and Communication for Sustainable Education*, 13(1), 37–47. <https://doi.org/doi:10.2478/dcse-2022-0004>

- Efstathiou, G. (2019). *Translation, Adaptation and Validation Process of Research Instruments BT - Individualized Care: Theory, Measurement, Research and Practice* (R. Suhonen, M. Stolt, & E. Papastavrou (eds.); pp. 65–78). Springer International Publishing. https://doi.org/10.1007/978-3-319-89899-5_7
- Eisend, M., & Kuss, A. (2019). *Obtaining Data for Theory Testing: Operationalization, Measurement, and Data Collection BT - Research Methodology in Marketing: Theory Development, Empirical Approaches and Philosophy of Science Considerations* (M. Eisend & A. Kuss (eds.); pp. 123–150). Springer International Publishing. https://doi.org/10.1007/978-3-030-10794-9_6
- Fávero, L. P., Belfiore, P., & de Freitas Souza, R. (2023). *Chapter 3 - Types of variables, measurement scales, and accuracy scales** (L. P. Fávero, P. Belfiore, & R. B. T.-D. S. de Freitas Souza Analytics and Machine Learning with R (eds.); pp. 29–36). Academic Press. <https://doi.org/https://doi.org/10.1016/B978-0-12-824271-1.00001-9>
- Fenn, J., Tan, C.-S., & George, S. (2020). Development, validation and translation of psychological tests. *BJPsych Advances*, 26(5), 306–315. <https://doi.org/DOI: 10.1192/bja.2020.33>
- Flannelly, L. T., Flannelly, K. J., & Jankowski, K. R. B. (2014). Independent, Dependent, and Other Variables in Healthcare and Chaplaincy Research. *Journal of Health Care Chaplaincy*, 20(4), 161–170. <https://doi.org/10.1080/08854726.2014.959374>
- Gärtner, F. R., Bomhof-Roordink, H., Smith, I. P., Scholl, I., Stiggelbout, A. M., & Pieterse, A. H. (2018). The quality of instruments to assess the process of shared decision making: A systematic review. In *PLoS ONE* (Vol. 13, Issue 2). <https://doi.org/10.1371/journal.pone.0191747>
- Gregory, R. J. (2015). *Psychological Testing: History, Principles, and Applications* (7th ed.). Pearson Education Limited.
- Gross, A. L., Chu, N., Anderson, L., Glymour, M. M., Jones, R. N., & Diseases, C. A. M. (2018). Do people with Alzheimer's disease improve with repeated testing? Unpacking the role of content and context in retest effects. *Age and Ageing*, 47(6), 866–871. <https://doi.org/10.1093/ageing/afy136>
- Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-Based Nursing*, 18(3), 66–67. <https://doi.org/10.1136/eb-2015-102129>
- Hidayat, A. A. (2021). *Menyusun instrumen penelitian dan uji validitas-reliabilitas*. Health Books Publishing.
- Holton III, E. F., Bates, R. A., Seyler, D. L., & Carvalho, M. B. (1997). Toward construct validation of a transfer climate instrument. *Human Resource Development Quarterly*, 8(2), 95–113. <https://doi.org/https://doi.org/10.1002/hrdq.3920080203>
- Jin, S., Noh, M., Yang-Wallentin, F., & Lee, Y. (2021). Robust nonlinear structural equation modeling with interaction between exogenous and endogenous latent variables. *Structural Equation Modeling: A Multidisciplinary Journal*, 28(4), 547–556. <https://doi.org/10.1080/10705511.2020.1857255>
- Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Journal of Health-System Pharmacy*, 65(23), 2276–2284. <https://doi.org/10.2146/ajhp070364>
- Kittur, J. (2023). Conducting Quantitative Research Study: A Step-by-Step Process. *Journal of Engineering Education Transformations*, 36(4), 100–112. <https://doi.org/10.16920/jeet/2023/v36i4/23120>
- Lane, S., & Pérez, A. (2023). *Validity* (R. J. Tierney, F. Rizvi, & K. B. T.-I. E. of E. (Fourth E. Ercikan (eds.); pp. 35–44). Elsevier. <https://doi.org/https://doi.org/10.1016/B978-0-12-818630-5.10005-3>
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563–575. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- Little, T. D., Stickley, Z. L., Rioux, C., & Wu, W. (2024). *Quantitative research methods* (W. Troop-Gordon & E. W. B. T.-E. of A. (Second E. Neblett (eds.); pp. 403–417). Academic Press. <https://doi.org/https://doi.org/10.1016/B978-0-323-96023-6.00095-6>
- Liu, S., Xu, S., Li, Q., Xiao, H., & Zhou, S. (2023). Development and validation of an instrument to assess students' science, technology, engineering, and mathematics identity. *Physical Review Physics Education Research*, 19(1), 10138. <https://doi.org/10.1103/PhysRevPhysEducRes.19.010138>

- Miksza, P., Shaw, J. T., Kapalka Richerme, L., Hash, P. M., Hodges, D. A., & Cassidy Parker, E. (2023). 309Measurement. In P. Miksza, J. T. Shaw, L. Kapalka Richerme, P. M. Hash, & D. A. Hodges (Eds.), *Music Education Research: An Introduction* (p. 0). Oxford University Press. <https://doi.org/10.1093/oso/9780197639757.003.0015>
- Mthembu, P. (2018). *Methodology and the Operationalisation of Variables BT - China and India's Development Cooperation in Africa: The Rise of Southern Powers* (P. Mthembu (ed.); pp. 75–110). Springer International Publishing. https://doi.org/10.1007/978-3-319-69502-0_5
- Nermend, K. (2023). *Initial Procedure for Data Analysis BT - Multi-Criteria and Multi-Dimensional Analysis in Decisions : Decision Making with Preference Vector Methods (PVM) and Vector Measure Construction Methods (VMCM)* (K. Nermend (ed.); pp. 37–94). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-40538-9_2
- Oluwatayo, J. A. (2012). Validity and Reliability Issues in Educational Research. *Journal of Educational and Social Research*, 2(May), 391–400. <https://doi.org/10.5901/jesr.2012.v2n2.391>
- Peeters, M. J., & Harpe, S. E. (2020). Updating conceptions of validity and reliability. *Research in Social and Administrative Pharmacy*, 16(8), 1127–1130. <https://doi.org/https://doi.org/10.1016/j.sapharm.2019.11.017>
- Pelletier, J. A., & Sharma, N. (2015). *Investigating the Impact of Student Samples on Data Quality and Research Outcomes BT - Ideas in Marketing: Finding the New and Polishing the Old* (K. Kubacki (ed.); p. 230). Springer International Publishing.
- Pittman, J., & Bakas, T. (2010). Measurement and Instrument Design. *Journal of Wound Ostomy & Continence Nursing*, 37(6). https://journals.lww.com/jwocnonline/fulltext/2010/11000/measurement_and_instrument_design.8.aspx
- Ramadhan, M. F., Siroj, R. A., & Afgani, M. W. (2024). Validitas and Reliabilitas. *Journal on Education*, 6(2), 10967–10975. <https://doi.org/10.31004/joe.v6i2.4885>
- Rossi, J. S., & Merenda, P. F. (2010). Variables in Research. In *The Corsini Encyclopedia of Psychology* (pp. 1–2). <https://doi.org/https://doi.org/10.1002/9780470479216.corpsy1020>
- Salmia, S. S. (2023). Development of Quality Instruments and Data Collection Techniques. *Jurnal Pendidikan Dan Pengajaran Guru Sekolah Dasar (JPPGuseda)*, 6(1), 119–124. <https://doi.org/10.55215/jppguseda.v6i1.7527>
- Selvi, H., & Özge, A. (2023). *Reliability and Validity of Clinical Scales Measurement BT - Clinical Scales for Headache Disorders* (P. Yalınay Dikmen & A. Özge (eds.); pp. 45–60). Springer International Publishing. https://doi.org/10.1007/978-3-031-25938-8_3
- Slife, B. D., Wright, C. D., & Yanchar, S. C. (2016). Using Operational Definitions in Research: A Best-Practices Approach. *The Journal of Mind and Behavior*, 37(2), 119–139. <http://www.jstor.org/stable/44631540>
- Stanger, N., Whitaker, L., & Backhouse, S. H. (2020). Testing the validity and reliability of the doping willingness in sport scale. *Scandinavian Journal of Medicine and Science in Sports*, 30(8), 1529–1543. <https://doi.org/10.1111/sms.13683>
- Stocker, M., Darroch, L., Krahl, R., Habermann, T., Devaraju, A., Schwardmann, U., D'onofrio, C., & Häggström, I. (2020). Persistent identification of instrument. *Data Science Journal*, 19(1), 1–12. <https://doi.org/10.5334/dsj-2020-018>
- Stuart, E. A., Ackerman, B., & Westreich, D. (2017). Generalizability of Randomized Trial Results to Target Populations: Design and Analysis Possibilities. *Research on Social Work Practice*, 28(5), 532–537. <https://doi.org/10.1177/1049731517720730>
- Subhaktiyasa, P. G. (2024a). Menentukan Populasi dan Sampel: Pendekatan Metodologi Penelitian Kuantitatif dan Kualitatif. *Jurnal Ilmiah Profesi Pendidikan*.
- Subhaktiyasa, P. G. (2024b). PLS-SEM for Multivariate Analysis: A Practical Guide to Educational Research using SmartPLS. *EduLine: Journal of Education and Learning Innovation*, 4(3). <https://doi.org/https://doi.org/10.35877/454RI.eduline2861>
- Subhaktiyasa, P. G., Agung, A. A. G., Jampel, I. N., & Dantes, K. R. (2024). Spiritual Leadership and Lecturer Performance: Mediating Role of Work Motivation. *International Journal of Evaluation and Research in Education*, 13(6), 3653–3662. <https://doi.org/10.11591/ijere.v13i6.29175>
- Subhaktiyasa, P. G., Andriana, K. R. F., Sintari, S. N. N., Shinta, W., Sumaryani, N. P., & Lede, Y. U.

- (2023). The Effect of Transformational Leadership, Servant Leadership, and Spiritual Leadership on Organizational Citizenship Behavior. *Jurnal Organisasi Dan Manajemen*, 19(1), 224–238. <https://doi.org/10.33830/jom.v19i1.3695.2023>
- Subhaktiyasa, P. G., & Sintari, S. N. N. (2024). Elevating Performance : Spiritual Leadership and Mitigation of Counterproductive Work Behavior in Higher Education. *Indonesian Journal of Educational Research and Review*, 7(2). <https://doi.org/10.23887/ijerr.v7i2.76966>
- Subhaktiyasa, P. G., Sintari, S. N. N., Andriana, K. R. F., Sumaryani, N. P., Werang, B. W., & Sudiarta, I. N. (2024). Tri Hita Karana and Organizational Culture in Society 5.0: Effect on Adaptability, Consistency, Involvement, and Mission. *Jurnal Ilmiah Manajemen Dan Bisnis*, 9(1), 1–10. <https://doi.org/https://doi.org/10.38043/jimb.v7i2.4679>
- Subhaktiyasa, P. G., Sutrisna, I. G. P. A. F., Sumaryani, N. P., & Sunita, N. W. (2024). Entrepreneurial Intentions among Medical Laboratory Technology Students : Effect of Education and Self-Efficacy. *Journal of Education and Learning*, 18(3), 719–726. <https://doi.org/https://doi.org/10.11591/edulearn.v18i3.21467>
- Sugiono. (2021). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D* (2nd ed.). Alfabeta.
- Taherdoost, H. (2017). Validity and Reliability of the Research Instrument ; How to Test the Validation of a Questionnaire / Survey in a Researchfile. *International Journal of Academic Research in Management*, 5(3), 27–36. <https://hal.science/hal-02546799/document>
- Thomas, S. (2022). Robert F. DeVellis & Carolyn T. Thorpe. (2022) Scale development: Theory and applications. California: SAGE Publications, Inc., 320 pages, \$60.00 softcover. *Personnel Psychology*, 75(1), 243–244. <https://doi.org/https://doi.org/10.1111/peps.12499>
- Thompson, B., & Daniel, L. G. (1996). Factor Analytic Evidence for the Construct Validity of Scores: A Historical Overview and Some Guidelines. *Educational and Psychological Measurement*, 56(2), 197–208. <https://doi.org/10.1177/0013164496056002001>
- Turcios, M. E., Agarwal, N., & Watkins, L. (2012). Library & information science literature: How much of it is research? *Proceedings of the American Society for Information Science and Technology*, 49(1), 1–4. <https://doi.org/https://doi.org/10.1002/meet.14504901320>
- Wang, R., & Krosnick, J. A. (2020). Middle alternatives and measurement validity: a recommendation for survey researchers. *International Journal of Social Research Methodology*, 23(2), 169–184. <https://doi.org/10.1080/13645579.2019.1645384>
- Winkler, C. E., & Berenbon, R. F. (2021). Validation of a survey for measuring scientists' attitudes toward data reuse. *Journal of the Association for Information Science and Technology*, 72(4), 449–453. <https://doi.org/https://doi.org/10.1002/asi.24412>
- Yusup, F. (2018). Uji Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif. *Jurnal Tarbiyah: Jurnal Ilmiah Kependidikan*, 7(1), 17–23. <https://doi.org/10.21831/jorpres.v13i1.12884>