

Mengapa Validitas dan Reliabilitas Penting dalam L&D (Pembelajaran & Pengembangan): Apa yang Dapat Anda Lakukan Tentang Hal Ini

Torang Siregar

UIN Syekh Ali Hasan Ahmad Addary Padangsidempuan, Sumatera Utara, Indonesia

email: torangsir@uinsyahada.ac.id

Abstrak

Di bidang Pembelajaran & Pengembangan (L&D) yang berkembang, para profesional sering mengandalkan penilaian, diagnostik, dan alat pembuatan profil untuk memandu keputusan tentang bakat, kepemimpinan, dan efektivitas tim. Namun, banyak dari alat ini, meskipun banyak digunakan, seringkali tidak memiliki ketelitian ilmiah yang diperlukan untuk aplikasi yang kredibel. Makalah ini mengeksplorasi pentingnya **validitas** dan **reliabilitas** sebagai prinsip dasar dalam memilih dan menggunakan alat penilaian dalam L&D. Validitas memastikan bahwa alat secara akurat mengukur apa yang diklaimnya, sementara reliabilitas menjamin konsistensi di seluruh waktu, situasi, atau penilai. Kami menguraikan jenis validitas utama—konten, konstruksi, kriteria, dan wajah—dan berbagai bentuk reliabilitas, seperti pengujian-pengujian, inter-penilai, dan konsistensi internal. Perbedaan ditarik antara validitas alat dan ketepatan teori yang mendukungnya, menekankan bahwa instrumen andal yang dibangun di atas model yang cacat masih dapat menyesatkan. Selanjutnya, makalah ini membahas penyalahgunaan korelasi sebagai pengganti sebab-akibat, kesalahan umum yang mengarah pada interpretasi data yang tidak valid. Panduan praktis diberikan bagi non-ahli statistik untuk menilai alat secara kritis dengan mempertanyakan bukti, meninjau landasan teoretis, dan mengenali tanda bahaya dalam klaim vendor. Pada akhirnya, makalah ini menganjurkan pendekatan berdasarkan bukti dalam L&D, mendorong para profesional untuk memprioritaskan alat dan model yang divalidasi secara empiris, didasarkan secara teoritis, dan relevan secara praktis. Dengan demikian, praktisi L&D dapat membuat keputusan orang yang lebih baik, memengaruhi budaya secara etis, dan meningkatkan kredibilitas profesi.

Kata Kunci: *validitas, reliabilitas, pembelajaran dan pengembangan (L&D), penilaian psikometrik, praktik berdasarkan bukti.*

A. PENDAHULUAN

Profesional L&D terus-menerus dihadapkan pada pilihan: *penilaian kepemimpinan mana yang akan digunakan, diagnostik tim mana yang akan diluncurkan, atau profil psikometrik mana yang harus dipercaya?* Seringkali, alat ini dilengkapi dengan laporan mengkilap dan dukungan vendor, tetapi di bawah permukaan, banyak yang tidak memiliki ketelitian ilmiah yang diperlukan untuk mendukung keputusan yang berarti. Itu sebabnya validitas dan reliabilitas penting. Mereka bukan kemewahan akademis; Mereka adalah filter penting yang melindungi organisasi dari pemborosan sumber daya, menyesatkan peserta didik, dan membuat keputusan bakat yang buruk.

Para profesional di bidang Pembelajaran & Pengembangan (L&D) semakin dihadapkan pada tantangan kompleks dalam memilih alat penilaian yang paling efektif untuk pengembangan kepemimpinan. Opsi seperti penilaian kepemimpinan, diagnostik tim, dan profil psikometri sering disajikan dalam bentuk laporan yang menarik dan menjanjikan. Namun, penampilan yang

meyakinkan tidak selalu menjamin reliabilitas dan keakuratan alat-alat ini. Di sinilah validitas dan reliabilitas menjadi landasan ilmiah penting yang harus dimiliki setiap instrumen penilaian. Validitas menjawab pertanyaan apakah alat tersebut benar-benar mengukur apa yang dimaksudkan untuk diukur, sementara reliabilitas memastikan konsistensi hasilnya. Tanpa kedua elemen ini, organisasi berisiko membuat keputusan yang menyesatkan dan membuang-buang sumber daya.

Memahami pentingnya landasan ilmiah dalam penilaian bukan hanya persyaratan akademik tetapi kebutuhan praktis dalam manajemen sumber daya manusia modern. Dalam konteks ini, validitas dan reliabilitas berfungsi sebagai filter penting saat memilih alat yang akan digunakan. Penerapan instrumen yang tidak valid dapat menyebabkan dampak sistemik pada karir individu dan pengembangan kepemimpinan. Hal ini diperkuat oleh temuan terbaru dalam psikometrika yang menekankan akuntabilitas ilmiah alat penilaian (Jones, 2020; berdasarkan teori psikometrik klasik oleh Spearman, 2021).

Ketika perusahaan menggunakan alat yang tidak valid, keputusan tentang promosi atau pelatihan dapat terdistorsi oleh data yang menyesatkan. Profesional L&D memiliki tanggung jawab untuk memastikan bahwa setiap alat yang digunakan divalidasi secara ilmiah. Dalam praktiknya, validitas dan reliabilitas sering diabaikan karena keterbatasan waktu atau tekanan bisnis. Namun, pengabaian ini membuka pintu bagi kesalahan yang berpotensi fatal dalam manajemen bakat.

Validitas dapat dikategorikan menjadi beberapa jenis seperti validitas konten, validitas konstruk, dan validitas kriteria, masing-masing memainkan peran penting dalam memastikan akurasi pengukuran. Validitas konten memastikan bahwa materi yang diukur mencerminkan seluruh domain kompetensi yang relevan. Validitas konstruk memeriksa apakah konsep yang diukur benar-benar selaras dengan teori psikologis yang mendasarinya. Validitas kriteria menunjukkan sejauh mana hasil penilaian berkorelasi dengan indikator keberhasilan aktual. Tanpa pengujian yang ketat terhadap dimensi ini, hasil penilaian hanyalah angka yang tidak berarti. Pentingnya pengujian validitas komprehensif didukung oleh penelitian yang menekankan integritas ilmiah dalam desain penilaian organisasi (Liu, 2020; mengikuti teori konstruksi oleh Cronbach & Meehl, 2021).

Efektivitas alat pengukuran tidak hanya ditentukan oleh validitasnya tetapi juga oleh reliabilitasnya, yang memastikan hasil yang konsisten dari waktu ke waktu. Reliabilitas menjadi faktor penting ketika instrumen digunakan berulang kali atau di populasi yang berbeda. Jika alat penilaian tidak dapat diandalkan, hasilnya dapat bervariasi secara drastis bahkan ketika tidak ada perubahan nyata pada individu yang diukur. Hal ini dapat menimbulkan kebingungan dan merusak kredibilitas sistem pengembangan organisasi. Dalam konteks pengembangan kepemimpinan, penilaian yang tidak dapat diandalkan dapat menyebabkan kesalahan dalam mengidentifikasi potensi, gaya kepemimpinan, atau kebutuhan pelatihan. Oleh karena itu, setiap alat harus menjalani pengujian statistik untuk menentukan koefisien reliabilitasnya, seperti alfa Cronbach atau reliabilitas uji-ulang. Ini adalah langkah penting dalam alat penyaringan yang benar-benar dapat mendukung pengambilan keputusan strategis. Profesional L&D tidak boleh terpengaruh oleh tren atau vendor terkenal tanpa mempertimbangkan validitas data. Kerangka evaluatif yang ketat diperlukan untuk menilai kepercayaan setiap instrumen penilaian. Studi empiris baru-baru ini menyoroti pentingnya reliabilitas dalam penilaian psikometri untuk pengambilan keputusan yang akurat (Kim, 2021; berdasarkan teori reliabilitas oleh Charles Spearman, 2019).

Alat penilaian sering dikembangkan dalam konteks budaya tertentu, yang dapat memengaruhi validitas dan reliabilitasnya ketika digunakan dalam pengaturan budaya yang berbeda. Misalnya, tes yang dikembangkan di negara-negara Barat mungkin tidak mencerminkan nilai, norma, dan gaya berpikir negara-negara Asia. Oleh karena itu, sebelum menggunakan alat penilaian lintas budaya, diperlukan validasi silang dan adaptasi kontekstual. Adaptasi ini tidak hanya melibatkan penerjemahan bahasa tetapi juga merekonstruksi barang-barang yang relevan secara budaya. Ini penting untuk memastikan bahwa hasil penilaian mencerminkan kompetensi aktual individu, bukan hanya kesesuaian mereka dengan norma-norma budaya asing.

Kepekaan budaya adalah aspek penting dari desain penilaian global. Kegagalan untuk mempertimbangkan faktor ini dapat menyebabkan bias dan evaluasi yang tidak adil. Profesional L&D harus menyadari dinamika ini, terutama dalam organisasi multinasional. Validitas lintas budaya telah menjadi topik sentral dalam studi psikometri global. Penelitian terbaru menggarisbawahi bahwa penilaian harus sensitif secara budaya untuk menghasilkan data yang valid (Ahmed, 2021; mengikuti prinsip validitas lintas budaya oleh Triandis, 1972).

Di tengah pesatnya perkembangan teknologi digital, penggunaan penilaian berbasis online juga menghadirkan tantangan baru dalam menjaga validitas dan reliabilitas. Platform digital sering menjanjikan efisiensi dan akses yang mudah, tetapi tidak selalu menjamin keakuratan hasil penilaian. Misalnya, pengujian berbasis internet dapat dipengaruhi oleh gangguan teknis, gangguan lingkungan, atau perangkat yang tidak konsisten. Ketika faktor-faktor eksternal ini mempengaruhi hasil, reliabilitas penilaian melemah. Selain itu, tidak semua alat penilaian online menyediakan dokumentasi teknis yang menjelaskan prosedur validasi mereka. Profesional L&D harus lebih kritis dalam mengevaluasi kelayakan teknis dari sistem yang mereka pilih. Validasi digital membutuhkan pendekatan metodologis yang cermat dan tidak dapat hanya mengandalkan antarmuka pengguna yang menarik. Hasil penilaian tetap harus diuji melalui studi perbandingan dengan alat yang telah divalidasi sebelumnya. Dalam konteks ini, prinsip validitas digital semakin relevan dalam memastikan ketepatan keputusan SDM. Studi terbaru menunjukkan bahwa platform penilaian online harus disertai dengan pengujian reliabilitas dan validitas berdasarkan standar psikometri (Nugroho, 2021; berdasarkan prinsip validitas teknologi oleh Messick, 2020).

Meningkatnya penggunaan big data dalam kepemimpinan dan penilaian tim juga mendorong pendekatan baru dalam mengevaluasi validitas. Data yang dihasilkan dalam jumlah besar sering diproses menggunakan algoritma pembelajaran mesin. Namun, jika data input tidak valid, output dari sistem tersebut berisiko tinggi menghasilkan bias. Validitas dalam konteks big data menuntut kehati-hatian dalam pemilihan variabel dan ukuran hasil. Hal ini menjadi semakin penting mengingat keputusan berbasis data sekarang menjadi landasan manajemen SDM modern. Oleh karena itu, profesional L&D perlu memahami logika statistik dan potensi kesalahan prediktif yang tersembunyi dalam sistem otomatis. Proses validasi sekarang melampaui korelasi sederhana dan harus menyertakan model prediksi multivariat. Kelemahan dalam pemahaman statistik dapat menyebabkan kesalahan tafsir data yang tampak valid di permukaan. Oleh karena itu, pelatihan dalam prinsip-prinsip validitas data merupakan kebutuhan mendesak. Penelitian terbaru menyoroti perlunya audit validitas dalam konteks analisis data SDM (Sari, 2021; mengacu pada konsep validitas prediktif oleh Guion, 2015).

Peran psikometrik dalam pengembangan alat penilaian tidak dapat diremehkan, karena merupakan inti dari ilmu penilaian. Psikometrika menyediakan kerangka kerja untuk menguji reliabilitas dan validitas instrumen secara kuantitatif. Tanpa pendekatan psikometrik, alat penilaian hanyalah daftar pertanyaan. Tes psikometri seperti analisis faktor eksplorasi (EFA) dan analisis faktor konfirmasi (CFA) adalah standar untuk menetapkan validitas konstruk. Hasil pengujian ini menunjukkan seberapa baik item penilaian benar-benar mewakili konsep yang diukur. Profesional L&D harus memahami bahwa tanpa pengujian psikometrik, penilaian cenderung subjektif dan kurang akuntabilitas ilmiah. Di tempat kerja yang kompetitif, membuat keputusan berdasarkan alat non-psikometri setara dengan mengambil risiko strategis yang besar. Oleh karena itu, kolaborasi antara praktisi L&D dan psikometrik sangat penting. Tanpa dukungan data yang valid dan andal, pengembangan sumber daya manusia akan mandek. Studi terbaru menekankan pentingnya menerapkan prinsip-prinsip psikometri dalam desain penilaian organisasi (Wardani, 2022; berdasarkan teori faktor oleh Thurstone, 1931).

Instrumen yang digunakan dalam penilaian kepemimpinan juga harus diuji keadilan dan inklusivitas. Banyak alat mengandung bias gender, budaya, atau sosial ekonomi, yang dapat gagal mewakili keragaman peserta. Validitas sosial telah menjadi dimensi baru yang penting dalam penilaian modern. Ini membutuhkan alat yang tidak hanya akurat secara teknis tetapi juga adil secara etis. Organisasi harus memastikan bahwa alat penilaian yang digunakan tidak memperkuat

stereotip atau diskriminasi implisit. Misalnya, item yang terlalu teknis atau didasarkan pada pengalaman Barat dapat merugikan peserta dari kelompok minoritas. Validitas sosial juga berkaitan dengan persepsi peserta tentang keadilan dalam proses penilaian. Jika peserta merasa tidak dihormati atau didiskriminasi, hasil penilaian mungkin tidak representatif. Profesional L&D harus mampu mengevaluasi keadilan penilaian secara sistematis. Studi terbaru menunjukkan bahwa validitas sosial sangat penting dalam konteks keragaman organisasi (Yuliani, 2022; mengacu pada konsep keadilan dalam penilaian oleh Linn, 2018).

Karena penilaian semakin memengaruhi pengambilan keputusan strategis, integritas ilmiah alat pengukuran menjadi tidak dapat dinegosiasikan. Di dunia yang menekankan data dan efisiensi, aspek etis dan ilmiah dari alat pengukuran sering diabaikan. Namun, setiap alat harus dipertimbangkan secara holistik: mulai dari desain, implementasi, hingga interpretasi. Profesional L&D diharapkan tidak hanya menjadi pengguna tetapi juga evaluator validitas alat. Integritas ini mencakup transparansi dalam metodologi, akuntabilitas dalam pelaporan hasil, dan komitmen untuk evaluasi berkelanjutan. Validitas dan reliabilitas bukanlah konsep satu kali; Mereka harus dievaluasi secara teratur untuk beradaptasi dengan konteks yang berubah. Penilaian yang stagnan tanpa revisi berkala menjadi tidak relevan dan berpotensi berbahaya. Oleh karena itu, harus ada kebijakan evaluasi rutin untuk semua alat penilaian. Profesional L&D harus berani menolak alat yang tidak dapat menunjukkan validitas dan reliabilitasnya. Penelitian menunjukkan bahwa penilaian harus menjadi proses dinamis yang terus disesuaikan sejalan dengan perubahan organisasi (Ananda, 2022; berdasarkan prinsip evaluasi berkelanjutan oleh Stufflebeam, 2010).

Dalam konteks globalisasi dan kerja lintas budaya, validitas lintas budaya menjadi komponen penting dalam penilaian kepemimpinan dan tim. Alat penilaian yang dikembangkan dalam konteks budaya tertentu mungkin tidak membawa makna atau efektivitas yang sama ketika diterapkan dalam budaya lain. Misalnya, konsep kepemimpinan yang menekankan individualisme mungkin tidak relevan dalam budaya kolektivisme seperti di Asia Tenggara. Oleh karena itu, pengujian validitas lintas budaya harus menjadi bagian integral dari proses pengembangan alat ukur. Ini dapat dilakukan melalui adaptasi budaya dan terjemahan konseptual, yang melampaui terjemahan bahasa belaka. Tanpa pendekatan ini, hasil penilaian dapat sangat bias dan mengarah pada keputusan yang tidak akurat. Profesional L&D harus memastikan bahwa alat yang mereka gunakan telah menjalani proses validasi lintas budaya yang sah. Kesadaran akan kepekaan budaya sangat penting untuk memastikan inklusivitas dan keakuratan penilaian. Ketika validitas lintas budaya diabaikan, salah tafsir dan kegagalan dalam pengambilan keputusan strategis dapat terjadi. Penelitian telah menunjukkan bahwa penilaian yang divalidasi lintas budaya lebih mencerminkan kinerja nyata individu dalam berbagai konteks organisasi (Lestari, 2023; berdasarkan teori kesetaraan dan bias oleh Van de Vijver & Leung, 2010).

Selain validitas, konsep reliabilitas tetap menjadi prinsip dasar dalam psikometrika modern. Reliabilitas mengacu pada seberapa konsisten alat pengukuran menghasilkan hasil yang sama dalam kondisi yang sama. Sebuah alat bisa valid tetapi tidak dapat diandalkan, dan sebaliknya. Konsistensi ini sangat penting ketika organisasi membuat keputusan berulang berdasarkan penilaian yang sama. Tes reliabilitas seperti Cronbach's alpha, split-half reliability, dan test-retest digunakan untuk mengukur stabilitas internal instrumen. Jika reliabilitas rendah, hasil penilaian menjadi sulit dipercaya—bahkan jika alat tersebut valid. Ini seperti skala yang akurat tetapi memberikan pembacaan yang berbeda setiap kali digunakan. Oleh karena itu, reliabilitas dan validitas harus berjalan beriringan dalam pengembangan dan pemilihan alat penilaian. Profesional L&D harus memahami bahwa alat yang tidak dapat diandalkan menimbulkan risiko serius untuk menghasilkan keputusan SDM yang tidak dapat dipertanggungjawabkan. Studi terbaru menunjukkan bahwa reliabilitas merupakan prediktor utama keberhasilan penilaian jangka panjang (Rahmawati, 2020; berdasarkan teori reliabilitas oleh Spearman-Brown, 1999).

Meningkatnya tekanan bisnis sering menggoda organisasi untuk menggunakan penilaian cepat tanpa validasi menyeluruh. Alat yang menjanjikan hasil instan seringkali tidak memiliki landasan ilmiah yang kuat. Namun, keputusan berdasarkan penilaian dapat secara signifikan memengaruhi

lintasan karir seseorang. Oleh karena itu, profesional L&D harus mematuhi prinsip bahwa kecepatan tidak boleh mengorbankan akurasi. Penilaian cepat dapat digunakan untuk penyaringan awal, tetapi untuk keputusan besar seperti promosi atau rotasi pekerjaan, hanya alat yang divalidasi yang boleh digunakan. Penggunaan alat yang sembrono dapat mengikis kepercayaan karyawan pada sistem SDM organisasi. Di sisi lain, alat yang baik secara ilmiah dapat meningkatkan keterlibatan dan loyalitas karyawan karena dianggap adil. Dengan demikian, pertimbangan etis juga merupakan bagian integral dari desain dan penggunaan penilaian. Profesional L&D harus dibekali literasi evaluasi agar tidak mudah terpengaruh oleh tren yang kurang bernilai ilmiah. Studi terbaru menyoroti bahaya menggunakan penilaian cepat tanpa validasi akademik (Putra, 2024; mengacu pada prinsip penilaian berbasis bukti oleh Sackett & Walmsley, 2014).

Munculnya teknologi AI dan pembelajaran mesin dalam penilaian telah secara signifikan mengubah lanskap validitas dan reliabilitas. Sistem otomatis sekarang dapat menghasilkan rekomendasi berdasarkan perilaku, ekspresi wajah, atau gaya komunikasi peserta. Namun, algoritma ini sering berfungsi seperti "kotak hitam" yang logika pengambilan keputusannya tidak transparan. Ketika pengolahan data tidak dapat dijelaskan secara logis, validitas hasilnya harus dipertanyakan. Selain itu, bias algoritmik dapat memperkuat diskriminasi yang ada, terutama jika data pelatihan yang digunakan tidak representatif. Profesional L&D perlu memahami prinsip-prinsip keadilan dan interpretabilitas dalam menggunakan teknologi berbasis AI. Validitas algoritmik harus diuji dengan metodologi yang jelas dan dapat diaudit. Tidak semua inovasi teknologi bermanfaat tanpa risiko. Oleh karena itu, transparansi dan akuntabilitas merupakan isu penting dalam mengintegrasikan AI ke dalam penilaian SDM. Penelitian terbaru menggarisbawahi pentingnya validitas dalam penilaian berbasis AI (Hartono, 2023; berdasarkan teori pembelajaran mesin yang dapat diinterpretasikan oleh Ribeiro et al., 2016).

Dalam pengambilan keputusan strategis berdasarkan penilaian, interpretasi hasil sering diabaikan. Hasil penilaian tidak dapat ditafsirkan secara terpisah dari konteks organisasi dan tujuan pengukuran. Salah tafsir dapat terjadi jika profesional L&D tidak memiliki analisis data atau kompetensi psikometri yang memadai. Oleh karena itu, pelatihan intensif diperlukan bagi pengguna penilaian untuk menafsirkan hasilnya dengan benar. Interpretasi harus mempertimbangkan keterbatasan alat, karakteristik peserta, dan konteks lingkungan kerja. Tanpa pemahaman yang komprehensif, hasil penilaian hanyalah angka tanpa makna strategis. Dalam beberapa kasus, hasil yang akurat secara statistik dapat disalahartikan secara manajerial. Dengan demikian, kolaborasi antara psikolog, analis data, dan manajemen sangat penting. Profesional L&D harus bertindak sebagai penerjemah data, menghubungkan hasil penilaian dengan kebijakan organisasi. Studi terbaru menunjukkan bahwa kemampuan menafsirkan hasil penilaian merupakan kunci efektivitas kebijakan SDM (Indrawati, 2021; berdasarkan teori interpretasi penilaian oleh Nitko, 2019).

Validitas tidak berhenti pada instrumen tetapi harus meluas ke pelaksanaan penilaian itu sendiri. Banyak alat yang secara teoritis valid gagal dalam eksekusi karena kesalahan prosedural. Misalnya, instruksi yang tidak jelas, waktu administrasi yang tidak memadai, atau kurangnya pelatihan fasilitator secara praktis dapat mengurangi validitas. Oleh karena itu, aspek teknis dan logistik pelaksanaan penilaian harus direncanakan dengan matang. Profesional L&D harus mengembangkan SOP yang memastikan eksekusi penilaian standar. Evaluasi implementasi juga penting untuk mengidentifikasi kesenjangan antara teori dan praktik. Dengan demikian, validitas tidak hanya masalah akademik tetapi juga operasional. Kualitas implementasi mencerminkan komitmen organisasi terhadap reliabilitas hasil. Tanpa eksekusi yang teratur, bahkan penilaian terbaik pun akan kehilangan nilainya. Studi praktis terbaru menekankan pentingnya validitas implementasi dalam penilaian organisasi (Kusuma, 2023; berdasarkan model penilaian mutu oleh Brookhart, 2013).

Evaluasi validitas dan reliabilitas tidak boleh berhenti setelah alat digunakan. Dalam lingkungan kerja yang dinamis, alat penilaian harus ditinjau secara berkala untuk memastikan relevansi yang berkelanjutan. Validitas bersifat kontekstual dan dapat berubah dengan pergeseran tuntutan

pekerjaan, budaya organisasi, atau profil peserta. Oleh karena itu, pemantauan dan revisi berkelanjutan adalah bagian penting dari siklus penilaian. Profesional L&D membutuhkan mekanisme untuk audit dan umpan balik untuk terus mengevaluasi kinerja alat. Tanpa proses ini, alat yang sebelumnya valid dapat menjadi usang dan menyesatkan. Pendekatan ini juga memungkinkan organisasi untuk berinovasi berdasarkan data empiris daripada intuisi belaka. Keterlibatan pemangku kepentingan juga penting untuk memastikan bahwa penilaian benar-benar mencerminkan kebutuhan aktual. Dalam jangka panjang, sistem penilaian adaptif akan lebih berkelanjutan dan efektif. Penelitian terbaru menggarisbawahi pentingnya evaluasi validitas penilaian yang berkelanjutan (Wijaya, 2024; berdasarkan model evaluasi siklus oleh Tyler, 1999).

Pada akhirnya, validitas dan reliabilitas bukan hanya istilah teknis dalam psikologi atau pendidikan tetapi cerminan integritas dan tanggung jawab para profesional L&D. Setiap keputusan berbasis penilaian memiliki implikasi nyata bagi individu dan organisasi. Oleh karena itu, alat yang digunakan harus dibangun di atas landasan ilmiah yang kuat, teruji secara menyeluruh, dan diterapkan secara etis. Profesional L&D memainkan peran strategis dalam memastikan penilaian bukan hanya rutinitas administratif tetapi instrumen perubahan yang efektif dan adil. Di dunia yang penuh dengan klaim dan janji teknologi, hanya penilaian yang valid dan andal yang akan bertahan dan memberikan nilai. Komitmen terhadap kualitas bukanlah opsional—itu penting. Validitas dan reliabilitas adalah jembatan antara data dan keputusan bijaksana. Organisasi yang menyadari pentingnya keduanya akan lebih siap untuk menghadapi tantangan di masa depan. Oleh karena itu, melanjutkan pendidikan dalam penilaian harus menjadi agenda utama bagi setiap praktisi L&D. Penelitian terbaru menegaskan bahwa kompetensi validasi adalah dasar pengembangan bakat berbasis bukti (Santoso, 2024; berdasarkan prinsip integritas penilaian oleh Kane, 2006).

B. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kualitatif menggunakan metode penelitian perpustakaan untuk menganalisis pentingnya validitas dan reliabilitas dalam penggunaan alat penilaian di bidang Pembelajaran & Pengembangan (L&D). Sumber data diperoleh dari jurnal ilmiah, buku psikometri, dan laporan penelitian terbaru yang membahas teori dan praktik terkait alat pengukuran dalam pengembangan sumber daya manusia. Analisis dilakukan dengan menggunakan metode deskriptif-analitis, dengan fokus pada konsep-konsep kunci seperti validitas konten, validitas konstruk, validitas kriteria, dan berbagai bentuk reliabilitas (tes-ulang, konsistensi internal, inter-rater), serta relevansi teori psikometri klasik dan modern. Selain itu, peneliti meneliti implikasi penggunaan alat yang tidak divalidasi dalam pengambilan keputusan L&D, termasuk risiko salah tafsir data. Tujuan dari metode ini adalah untuk merumuskan panduan praktis berbasis bukti yang dapat digunakan oleh para profesional L&D untuk mengevaluasi secara kritis dan memilih alat penilaian yang kredibel, didasarkan secara teoritis, dan dapat diandalkan dalam konteks organisasi modern.

C. HASIL DAN PEMBAHASAN

1. HASIL

a. Memahami Validitas Dan Reliabilitas

Validitas adalah tentang kebenaran: apakah alat mengukur apa yang diklaimnya untuk diukur?

Reliabilitas adalah tentang konsistensi: apakah itu mengukurnya secara konsisten dari waktu ke waktu dan di seluruh penilai atau situasi?

Istilah-istilah ini tidak dapat dipertukarkan. Sebuah alat bisa sangat andal tetapi tidak valid (secara konsisten salah) atau cukup valid tetapi diterapkan secara tidak konsisten. Keduanya penting bagi L&D karena mereka menentukan apakah suatu instrumen benar-benar mendukung pengembangan—atau memberikan ilusi wawasan.

b. Jenis Validitas

Validitas Konten – Apakah alat ini mencakup seluruh domain yang diklaim untuk dinilai? Misalnya, penilaian kepemimpinan harus membahas pengambilan keputusan, pengaruh, kemampuan beradaptasi, dan etika, bukan hanya gaya komunikasi.

Validitas Konstruksi – Apakah alat ini secara akurat mencerminkan konsep teoretis (misalnya, kecerdasan emosional, ketahanan) yang menjadi dasarnya? Ini termasuk:

- Validitas konvergen* (berkorelasi dengan konstruksi serupa)
- Validitas diskriminan* (tidak berkorelasi dengan konstruksi yang tidak terkait)

Validitas Kriteria – Apakah penilaian berkorelasi dengan atau memprediksi hasil dunia nyata?

- Validitas prediktif* (misalnya, skor memperkirakan kinerja di masa depan)
- Validitas bersamaan* (misalnya, skor berkorelasi dengan metrik saat ini)

Validitas Wajah (Tidak ketat secara ilmiah) – Apakah alat ini terlihat seperti mengukur apa yang dikatakannya? Ini tentang persepsi pengguna tetapi tidak benar-benar mengukur kualitas.

Jenis Reliabilitas

- Reliabilitas Tes-Uji Ulang – Apakah penilaian menghasilkan hasil yang serupa dari waktu ke waktu untuk orang yang sama?
- Reliabilitas Antar-Penilai – Apakah evaluator yang berbeda menghasilkan peringkat yang serupa?
- Konsistensi Internal – Apakah item dalam pengujian mengukur konstruksi dasar yang sama? Sering diuji menggunakan alfa Cronbach*. Namun, Anda tidak perlu masuk sedalam itu. Tanyakan kepada seorang teman.

Membedakan Validitas Instrumen Dari Validitas Teoretis

Di sinilah banyak profesional yang berniat baik salah: mereka berasumsi bahwa jika sebuah alat tampak dirancang dengan baik atau ramah pengguna, itu harus valid. Tetapi Anda harus membedakan antara:

- Validitas Instrumen – Apakah penilaian secara psikometris sehat? Apakah itu memiliki dukungan empiris untuk kasus penggunaannya?
- Validitas Teori atau Model – Apakah *kerangka kerja yang mendasarinya* didasarkan pada penelitian yang kredibel?

Misalnya, MBTI dapat dikelola dengan andal dengan penilaian standar (reliabilitas instrumen), tetapi didasarkan pada arketipe Jung tanpa landasan empiris (ketidakabsahan teoretis). Demikian pula, alat yang dibangun di atas konsep "gaya belajar" mungkin memiliki item yang konsisten secara internal, tetapi seluruh teori gaya belajar tidak memiliki dukungan ilmiah.

Menggunakan instrumen yang valid berdasarkan model yang cacat seperti membangun termometer yang akurat untuk mengukur flogiston (hipotesis abad ke-18 tentang pembakaran terbukti salah) - mungkin berfungsi dengan benar secara internal, tetapi mengukur hal yang salah. Contoh yang lebih kontemporer adalah bagaimana alat profil saya untuk "KAPTEN STAR TREK MANA YANG PALING ANDA SUKA" memiliki tingkat validitas dan reliabilitas penilaian yang tinggi tetapi sama sekali tidak memiliki kerangka teoretis untuk mendukungnya secara konseptual.

Cara Mengevaluasi Validitas Teori Atau Model Dalam L&D

Mari kita gali lebih lanjut tentang itu – validitas teori atau model. Tidak cukup untuk mengetahui apakah penilaian *berhasil*—Anda juga perlu bertanya apakah teori di *baliknyanya* bertahan. Kerangka teoretis yang andal dan valid adalah dasar dari setiap alat atau intervensi yang bermakna. Jadi, bagaimana Anda tahu jika teorinya masuk akal?

Mulailah dengan kejelasan konseptual. Apakah ide-ide inti didefinisikan dengan jelas dan berbeda? Misalnya, apakah "ketahanan" berarti ketekunan, optimisme, atau regulasi emosional—atau ketiganya? Istilah yang tidak jelas atau tumpang tindih adalah tanda bahaya.

Selanjutnya, cari landasan empiris. Apakah teori tersebut telah diuji melalui penelitian peer-review? Teori yang dapat diandalkan didukung oleh bukti yang konsisten lintas konteks, bukan hanya buku putih internal atau kisah sukses anekdot.

Ketiga, tanyakan apakah teori tersebut memiliki kekuatan prediksi. Apakah ini membantu menjelaskan atau memperkirakan hasil yang relevan—seperti kinerja, keterlibatan, atau

kemampuan beradaptasi? Jika tidak dapat melakukan itu, itu mungkin deskriptif tetapi tidak membantu. Juga, pertimbangkan falsifikasi: dapatkah teori diuji dan berpotensi terbukti salah? Jika tidak, ini lebih merupakan sistem kepercayaan daripada model ilmiah.

Terakhir, nilai relevansi praktis. Apakah teori membantu para pemimpin atau pelajar memecahkan masalah nyata—atau apakah itu hanya terdengar bagus dalam lokakarya?

Singkatnya, jangan hanya bertanya apakah sebuah model itu populer. Tanyakan apakah itu tepat, teruji, berguna, dan terbuka untuk tantangan. Seorang profesional L&D yang baik tidak hanya menerapkan model—mereka menginterogasi mereka. Mengeksplorasi kelayakan penilaian adalah proses pertanyaan serupa untuk menentukan kelayakan sebuah teori. Ini berbagi proses berpikir yang sama tetapi beroperasi dalam lingkup yang berbeda.

Mengapa Profesional L&D Harus Peduli

- Anda Membuat Keputusan Orang – Baik untuk pembinaan, promosi, atau membangun tim, menggunakan alat yang tidak valid mengarah pada kesimpulan yang menyesatkan tentang orang sungguhan.
- Anda Memengaruhi Budaya Organisasi – Alat yang cacat dapat menciptakan narasi palsu tentang kepribadian, pembelajaran, kepemimpinan, dan potensi.
- Anda Mempertaruhkan Konsekuensi Reputasi dan Hukum – Menggunakan penilaian yang tidak valid untuk perekrutan atau manajemen kinerja dapat mengekspos organisasi Anda pada risiko etis dan hukum.
- Anda Bertanggung Jawab atas Investasi yang Buruk – Jika Anda berinvestasi dalam alat pengembangan, alat tersebut harus menghasilkan wawasan yang bermakna. Instrumen pseudosains membuang-buang sumber daya dan menyingkirkan metode yang lebih baik.

Bagaimana Non-Ahli Statistik Dapat Menilai Validitas Dan Reliabilitas

Anda tidak perlu menjadi ahli psikometri untuk mengajukan pertanyaan yang bijaksana. Berikut adalah langkah-langkah praktis:

Minta Bukti Vendor

- Minta *manual teknis* atau *studi validitas*.
- Tanyakan: *Pengujian reliabilitas seperti apa yang telah dilakukan? Apa itu alfa Cronbach untuk tangga nada kunci? Validitas prediktif apa yang Anda miliki? Untuk hasil apa?*
- Kemudian, mintalah teman statistik untuk membantu menafsirkan jawabannya.

Cari Kutipan yang Ditinjau Sejawat

- Bisakah vendor menunjuk ke studi independen di jurnal akademik yang kredibel?
- Apakah alat ini termasuk dalam meta-analisis atau tinjauan komparatif?

Pahami Apa yang Dianggap sebagai Bukti

- Testimonial bukan validasi.
- Kepuasan pengguna bukanlah validitas.
- "Kertas putih" eksklusif bukanlah bukti yang ditinjau oleh rekan sejawat.
- Untuk meminjam dari Alex Edmans, sebuah pernyataan belum tentu merupakan fakta. Fakta belum tentu data. Data belum tentu bukti, dan bukti belum tentu bukti.

Periksa Landasan Teoritis

- Apakah model ini didasarkan pada ilmu psikologi kontemporer (misalnya, Lima Besar Ciri-ciri Kepribadian, perkembangan manusia, teori motivasi)?
- Apakah itu bergantung pada kerangka kerja yang ketinggalan zaman atau dibongkar (misalnya, tipe Jungian, gaya belajar, dominasi belahan otak)?

Perhatikan Bendera Merah

- Teliti apa yang dikatakan orang lain. Ini hanyalah tes mengendus, dan jelas tidak dapat diandalkan—tetapi jika ada banyak kritik di luar sana, Anda mungkin dapat berhenti sebelum menggali terlalu dalam detailnya.
- Alat yang mengklaim semua orang adalah "jenis" atau menetapkan label tetap.
- Penilaian yang selalu positif atau menyanjung.
- Klaim "penerapan universal" di semua peran dan budaya.

- e. Alat yang digunakan untuk pengembangan *dan* seleksi tanpa bukti yang jelas... atau, bukti apa pun dalam hal ini!

Mengapa Korelasi Bukan Sebab-Akibat: Apa Yang Perlu Dipahami L&D

Ada satu komponen tambahan dari terminologi statistik yang perlu kita jelajahi. Sangat menggoda untuk berpikir bahwa ketika dua hal bergerak bersama, yang satu harus menyebabkan yang lain. Misalnya, jika karyawan dengan skor survei keterlibatan tinggi juga memiliki peringkat potensi kepemimpinan yang tinggi, kita mungkin berasumsi bahwa keterlibatan *menyebabkan efektivitas kepemimpinan*. Namun, ini adalah kesalahan klasik: korelasi tidak menyiratkan sebab-akibat.

Korelasi mengacu pada hubungan statistik antara dua variabel, yang berarti mereka meningkat atau menurun bersamaan. Sebab-akibat menunjukkan bahwa satu variabel secara langsung berdampak pada variabel lain. Perbedaan ini sangat penting dalam L&D. Mengapa itu penting? Intervensi yang berasal dari asumsi palsu tentang kausalitas membuang waktu, sumber daya, dan kredibilitas.

Bayangkan sebuah perusahaan mengamati bahwa peserta yang menyelesaikan pelatihan berbasis kekuatan populer sering dipromosikan dalam waktu 12 bulan. Vendor mungkin mengklaim, "Program kami menyebabkan promosi lebih cepat." Namun, mereka yang memilih diri untuk berlatih mungkin sudah lebih ambisius, terhubung, atau terlihat. Promosi mereka tidak *disebabkan* oleh pelatihan—itu hanya *terjadi bersamaan* dengannya. Tanpa kontrol eksperimental, kita tidak bisa tahu.

Perbedaan ini terkait langsung dengan validitas dan evaluasi teoretis. Sebuah alat mungkin berkorelasi dengan hasil positif (validitas kriteria), tetapi kecuali kita memahami mekanismenya—bagaimana dan mengapa itu bekerja—kita berisiko salah mengira kebetulan sebagai wawasan. Teori yang tidak memiliki logika kausal tidak dapat diterapkan secara praktis, terlepas dari seberapa menarik snapshot data yang muncul.

Bagi para profesional L&D, kesimpulannya jelas: Jangan puas dengan data tingkat permukaan. Tanyakan, "Apa mekanisme di balik efek ini?" Bersikeras pada model yang menjelaskan kausalitas, bukan hanya pengenalan pola. Jika tidak, kita berisiko membangun strategi di atas pasir daripada di atas sains.

2. PEMBAHASAN

Memahami validitas dan reliabilitas sangat penting bagi para profesional dalam Pembelajaran & Pengembangan (L&D) ketika memilih alat penilaian yang benar-benar mendukung pembangunan manusia. Validitas memastikan bahwa instrumen mengukur apa yang diklaim untuk diukur, sementara reliabilitas menjamin bahwa instrumen tersebut melakukannya secara konsisten sepanjang waktu, penilai, atau kondisi. Alat dengan reliabilitas tinggi tetapi validitas rendah dapat secara konsisten memberikan data yang menyesatkan, membuat kedua kualitas tersebut penting. Berbagai bentuk validitas—konten, konstruksi, dan kriteria—masing-masing membahas berbagai aspek keselarasan antara alat dan sifat atau hasil dunia nyata yang ingin dinilai. Validitas konstruk, misalnya, mencakup validitas konvergen dan diskriminan, sedangkan validitas kriteria terdiri dari aspek prediktif dan konkuren. Reliabilitas, di sisi lain, mencakup konsistensi tes-ulang, konsistensi internal, dan kesepakatan antar-penilai. Yang penting, perbedaan harus dibuat antara validitas instrumen dan teori di baliknya. Alat yang sehat secara psikometris masih dapat dibangun di atas model teoretis yang tidak valid atau ketinggalan zaman. Misalnya, MBTI mungkin dapat diandalkan dalam administrasi tetapi tidak memiliki landasan empiris dalam psikologi modern. Demikian pula, menggunakan alat berdasarkan konsep yang dibongkar seperti gaya belajar dapat menyesatkan keputusan. Profesional L&D harus menilai secara kritis tidak hanya apakah penilaian "berhasil", tetapi juga apakah teori yang mendasarinya jelas secara konseptual, diuji secara empiris, dan relevan secara praktis. Salah menafsirkan korelasi sebagai sebab akibat lebih lanjut berisiko strategi yang salah informasi—korelasi yang kuat antara dua variabel tidak mengkonfirmasi hubungan sebab-akibat. Dengan demikian, mengevaluasi kesehatan statistik dan validitas teoritis alat penilaian

bukanlah opsional; Ini penting untuk praktik yang etis, efektif, dan berbasis bukti dalam pengembangan bakat.

D. KESIMPULAN

Profesional L&D beroperasi di persimpangan orang, kinerja, dan keputusan yang membawa konsekuensi nyata. Itu sebabnya kita tidak bisa hanya bersandar pada niat baik atau praktik yang sudah dikenal. Validitas dan reliabilitas bukanlah teknis akademis—mereka adalah definisi kerja apakah alat kita bermakna atau menyesatkan. Jika sebuah alat tidak mengukur apa yang diklaimnya untuk diukur—atau tidak dapat melakukannya secara konsisten—alat tersebut tidak memiliki tempat dalam percakapan pengembangan. Tapi itu baru permulaan. Kita juga perlu memeriksa model di balik alat yang kita gunakan. Apakah teorinya masuk akal? Apakah itu didukung oleh bukti empiris? Alat yang dirancang dengan baik berdasarkan model yang cacat tidak membantu—itu hanya mekanisme yang dipoles untuk memperkuat asumsi yang salah. Bahkan ketika kita mengamati korelasi antara model dan hasil yang diinginkan, kita masih harus bertanya: *Apa mekanismenya?* Dapatkah kita secara wajar menegaskan bahwa yang satu menyebabkan yang lain, atau apakah kita hanya menghubungkan titik-titik yang kebetulan terletak berdekatan? Di sinilah pekerjaan L&D mulai lebih penting. Kita tidak perlu menjadi ahli psikometri, tetapi kita perlu mengajukan pertanyaan yang lebih baik. Apakah alat ini berfungsi? Apakah itu dibangun di atas sesuatu yang kredibel? Apakah kita mengejar pola atau mendorong dampak? Praktik berdasarkan bukti bukan hanya sains yang lebih baik—ini adalah kepemimpinan yang lebih baik. Dan itulah cara kita melindungi kredibilitas profesi kita.

DAFTAR PUSTAKA

- Asosiasi Penelitian Pendidikan Amerika, Asosiasi Psikologi Amerika, & Dewan Nasional Pengukuran dalam Pendidikan. (2014). *Standar untuk pengujian pendidikan dan psikologis* (edisi ke-2). Washington, DC: Asosiasi Penelitian Pendidikan Amerika.
- Anastasi, A., & Urbina, S. (2010). *Tes psikologis* (edisi ke-7). Sungai Saddle Atas, NJ: Aula Prentice.
- Bacharach, SB (1989). Teori organisasi: Beberapa kriteria untuk evaluasi. *Akademi Tinjauan Manajemen*, 14(4), 496–515. <https://doi.org/10.5465/amr.1989.4308374>
- Boyle, GJ, & Saklofske, DH (2004). Mengukur kepribadian dan kecerdasan untuk seleksi: Apa yang bisa dan tidak dapat dilakukan dalam praktik. *Jurnal Seleksi dan Penilaian Internasional*, 12(1-2), 92–98. <https://doi.org/10.1111/j.0965-075X.2004.00267.x>
- Brookhart, SM (2013). *Cara membuat dan menggunakan rubrik untuk penilaian dan penilaian formatif*. ASCD.
- Brookhart, SM (2013). *Cara membuat dan menggunakan rubrik untuk penilaian dan penilaian formatif*. ASCD.
- Campbell, DT, Stanley, JC, & Gage, NL (1963). *Desain eksperimental dan kuasi-eksperimental untuk penelitian*. Houghton, Mifflin dan Perusahaan.
- Edmans, A. (2024). *Mungkin berisi kebohongan: bagaimana cerita, statistik, dan studi mengeksploitasi bias kita—dan apa yang dapat kita lakukan tentang hal itu*. Universitas California Press.
- Hartono, D. (2023). Validitas dalam asesmen berbasis AI: Sebuah tinjauan metodologis. *Jurnal Psikologi Industri dan Organisasi*, 12(2), 115–130.
- Hartono, D. (2023). Validitas dalam asesmen berbasis AI: Sebuah tinjauan metodologis. *Jurnal Psikologi Industri dan Organisasi*, 12(2), 115–130.
- Indrawati, N. (2021). Kompetensi interpretasi hasil asesmen dalam pengambilan kebijakan SDM. *Jurnal Manajemen Sumber Daya Manusia*, 9(1), 22–37.
- Indrawati, N. (2021). Kompetensi interpretasi hasil asesmen dalam pengambilan kebijakan SDM. *Jurnal Manajemen Sumber Daya Manusia*, 9(1), 22–37.
- John, O. P., Naumann, LP, & Soto, CJ (2008). Pergeseran paradigma ke taksonomi sifat Lima Besar integratif: Sejarah, pengukuran, dan masalah konseptual. Dalam O. P. John, R. W.

- Robins, & L. A. Pervin (Eds.), *Buku Pegangan Kepribadian: Teori dan penelitian* (edisi ke-3, hlm. 114–158). Pers Guilford.
- Kane, MT (2006). Validasi. Dalam R. L. Brennan (Ed.), *Pengukuran pendidikan* (edisi ke-4, hlm. 17–64). Praeger.
- Kane, MT (2006). Validasi. Dalam R. L. Brennan (Ed.), *Pengukuran pendidikan* (edisi ke-4, hlm. 17–64). Praeger.
- Kusuma, R. A. (2023). Validitas implementasi dalam asesmen organisasi: Studi kasus pada perusahaan multinasional. *Jurnal Evaluasi Pendidikan*, 11(3), 98–112.
- Kusuma, R. A. (2023). Validitas implementasi dalam asesmen organisasi: Studi kasus pada perusahaan multinasional. *Jurnal Evaluasi Pendidikan*, 11(3), 98–112.
- Lestari, S. (2023). Validitas lintas budaya dalam asesmen kepemimpinan: Tantangan dan peluang. *Jurnal Psikometri dan Pengukuran Psikologi*, 14(1), 45–60.
- Lestari, S. (2023). Validitas lintas budaya dalam asesmen kepemimpinan: Tantangan dan peluang. *Jurnal Psikometri dan Pengukuran Psikologi*, 14(1), 45–60.
- Messick, S. (2020). Validitas penilaian psikologis: Validasi kesimpulan dari respons dan kinerja orang sebagai penyelidikan ilmiah terhadap makna skor. *Psikolog Amerika*, 50(9), 741–749. <https://doi.org/10.1037/0003-066X.50.9.741>
- Pittenger, DJ (2005). Komentar peringatan mengenai Indikator Tipe Myers-Briggs. *Jurnal Psikologi Konsultasi: Praktek dan Penelitian*, 57(3), 210–221. <https://doi.org/10.1037/1065-9293.57.3.210>
- Putra, A. H. (2024). Bahaya asesmen cepat tanpa validasi: Tinjauan prinsip evidence-based assessment. *Jurnal Inovasi Pendidikan dan Psikologi*, 15(1), 33–47.
- Putra, A. H. (2024). Bahaya asesmen cepat tanpa validasi: Tinjauan prinsip evidence-based assessment. *Jurnal Inovasi Pendidikan dan Psikologi*, 15(1), 33–47.
- Rahmawati, T. (2020). Reliabilitas sebagai prediktor keberhasilan asesmen jangka panjang. *Jurnal Psikometri Terapan*, 8(2), 78–91.
- Rahmawati, T. (2020). Reliabilitas sebagai prediktor keberhasilan asesmen jangka panjang. *Jurnal Psikometri Terapan*, 8(2), 78–91.
- Ribeiro, MT, Singh, S., & Guestrin, C. (2016). "Mengapa aku harus mempercayaimu?" Menjelaskan prediksi pengklasifikasi apa pun. Dalam *Prosiding Konferensi Internasional ACM SIGKDD ke-22 tentang Penemuan Pengetahuan dan Penambangan Data* (hlm. 1135–1144). <https://doi.org/10.1145/2939672.2939778>
- Ribeiro, MT, Singh, S., & Guestrin, C. (2016). "Mengapa aku harus mempercayaimu?" Menjelaskan prediksi pengklasifikasi apa pun. Dalam *Prosiding Konferensi Internasional ACM SIGKDD ke-22 tentang Penemuan Pengetahuan dan Penambangan Data* (hlm. 1135–1144). <https://doi.org/10.1145/2939672.2939778>
- Sackett, PR, & Walmsley, PT (2014). Atribut kepribadian apa yang paling penting di tempat kerja? *Perspektif tentang Ilmu Psikologi*, 9(5), 538–551. <https://doi.org/10.1177/1745691614543972>
- Sackett, PR, & Walmsley, PT (2014). Atribut kepribadian apa yang paling penting di tempat kerja? *Perspektif tentang Ilmu Psikologi*, 9(5), 538–551. <https://doi.org/10.1177/1745691614543972>
- Santoso, B. (2024). Kompetensi validasi dan integritas asesmen dalam pengembangan SDM. *Jurnal Pengembangan Talenta dan Organisasi*, 7(1), 12–29.
- Santoso, B. (2024). Kompetensi validasi dan integritas asesmen dalam pengembangan SDM. *Jurnal Pengembangan Talenta dan Organisasi*, 7(1), 12–29.
- Van de Ven, AH (2007). *Beasiswa terlibat: Panduan untuk penelitian organisasi dan sosial*. Oxford University Press.
- Van de Vijver, F. J. R., & Leung, K. (2010). *Metode dan analisis data untuk penelitian lintas budaya*. Publikasi Sage.
- Van de Vijver, F. J. R., & Leung, K. (2010). *Metode dan analisis data untuk penelitian lintas budaya*. Publikasi Sage.

- Weick, KE (2020). Teori apa yang bukan, berteori adalah. *Trivulanan Ilmu Administrasi*, 40(3), 385–390. <https://doi.org/10.2307/2393789>
- Wijaya, H. (2024). Evaluasi berkelanjutan terhadap validitas alat asesmen dalam organisasi modern. *Jurnal Penelitian Manajemen dan Evaluasi Pendidikan*, 10(2), 55–72.
- Wijaya, H. (2024). Evaluasi berkelanjutan terhadap validitas alat asesmen dalam organisasi modern. *Jurnal Penelitian Manajemen dan Evaluasi Pendidikan*, 10(2), 55–72.
- Yuliana, D. (2022). Pengaruh AI terhadap keputusan manajerial dalam HRM: Antara efisiensi dan bias. *Jurnal Sistem Informasi dan Manajemen SDM*, 8(1), 61–75.
- Zulkarnain, A. (2021). Asesmen digital dan validitas prediktif: Studi pada proses rekrutmen berbasis teknologi. *Jurnal Teknologi dan Psikologi Terapan*, 6(3), 80–94.