

## Tes Kemampuan Akademik (TKA) sebagai Inovasi Evaluasi Pendidikan: Sebuah Tinjauan Naratif

\*Agus Sukamto, Bahrani, Nanik Nurvayanti

Universitas Islam Negeri Samarinda, Samarinda, Indonesia

\*Email: [sukamto.agus1608@gmail.com](mailto:sukamto.agus1608@gmail.com) (Corresponding Author)



DOI: <https://doi.org/10.53621/jider.v6i1.695>

### Informasi Artikel

#### Riwayat Artikel:

Diterima: 5 Desember 2025

Revisi Akhir: 18 Desember 2025

Disetujui: 3 Februari 2026

Terbit: 28 Februari 2026

#### Kata Kunci:

Asesmen Objektif;

Inovasi Evaluasi Pendidikan;

Kebijakan Responsif;

Mutu Pendidikan;

Tes Kemampuan Akademik.



### ABSTRAK

Penelitian ini mengkaji secara kritis dan komprehensif peran Tes Kemampuan Akademik (TKA) sebagai inovasi evaluasi pendidikan dalam mewujudkan asesmen yang objektif, adil, dan terstandarisasi di Indonesia. TKA merupakan kebijakan baru yang menggantikan Ujian Nasional (UN) dan berfungsi sebagai asesmen opsional berbasis komputer yang tidak menentukan kelulusan, melainkan fokus pada pengukuran capaian akademik murid untuk seleksi prestasi, penyetaraan hasil belajar lintas jalur pendidikan (formal, non formal, informal), serta pengendalian mutu pendidikan. Metode penelitian menggunakan pendekatan kualitatif dengan desain tinjauan naratif literatur (*narrative literature review*), untuk menelaah literatur akademis, buku referensi, dokumen kebijakan resmi, dan laporan kajian. Kemudian data literatur diolah melalui kombinasi dua model analisis yang saling komplementer, yaitu model analisis interaktif dari Miles dan Huberman dan prosedur analisis tematik reflektif dari Braun dan Clarke. Hasil penelitian menunjukkan TKA dirancang secara objektif dan terstandarisasi dengan landasan psikometris IRT dan CAT, mengukur kemampuan HOTS seperti penalaran verbal, numerik, logika, dan spasial. Inovasi TKA yang inklusif menghadapi tantangan utama seperti kesenjangan digital di daerah terpencil, keterbatasan SDM, dan risiko *high-stakes test*. Solusi mitigasi mencakup pilot CAT-IRT, pelatihan guru via MoOC, dan dukungan infrastruktur afirmatif. Inovasi TKA diharapkan dapat merevolusi evaluasi pendidikan nasional menuju sistem yang adaptif, berbasis data, inklusif, dan berkeadilan.

## PENDAHULUAN

Pendidikan merupakan fondasi utama dalam pembangunan bangsa dan pengembangan sumber daya manusia yang berkualitas. Tujuan pendidikan nasional menurut Undang-Undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional adalah untuk mengembangkan potensi peserta didik agar menjadi manusia yang beriman dan bertakwa kepada Tuhan Yang Maha Esa, berakhlak mulia, sehat, berilmu, cakap, kreatif, mandiri, serta menjadi warga yang demokratis dan bertanggung jawab (Sukmayadi, V. & Yahya, A., 2020). Pendidikan sebagai proses pembentukan karakter dan kompetensi harus mampu menjawab tuntutan zaman yang terus berubah. Sependapat dengan Nikum, K. (2022) yang mengeksplorasi filosofi pendidikan secara keseluruhan, mulai dari evolusi Education 1.0 hingga 5.0, dengan penekanan pada praktik pembelajaran terbaru yang berorientasi pada kebutuhan sosial seperti inovasi, personalisasi, kolaborasi *multi-level*, dan pemberdayaan individu melalui teknologi canggih.

Peningkatan mutu pendidikan menjadi fokus strategis dalam upaya mewujudkan tujuan pendidikan nasional tersebut. Menurut Rodriguez, J. V., et al. (2022), mutu pendidikan tidak hanya berkaitan dengan fasilitas dan sarana, tetapi lebih pada kualitas proses pembelajaran dan hasil belajar yang mencerminkan penguasaan kompetensi secara menyeluruh. Peningkatan mutu pendidikan memerlukan evaluasi yang objektif dan berkelanjutan agar bisa memberikan gambaran yang akurat tentang capaian pendidikan dan area yang perlu diperbaiki. Diketahui bahwa siswa Indonesia mencetak skor rata-rata 379 pada matematika PISA 2018, dengan 71% di bawah kompetensi minimum, hal ini menandakan kesulitan besar dalam memecahkan masalah

numerik yang menitikberatkan penalaran bukan sekedar hafalan (dikdasmen, 2022). Data PISA 2022 skor rata-rata siswa Indonesia pada matematika mencapai 366 poin, turun 13 poin dari 379 pada 2018, dengan hanya 18,35% siswa mencapai level 2 atau jauh di bawah rata-rata OECD 68,91%. Meskipun Indonesia naik 5 posisi menjadi ke-69 dari 81 negara, namun skor absolut mengalami penurunan akibat *learning loss* pandemi (Pusmendik, 2023). Hasil temuan Wang, X. S., et al. (2023) mengungkapkan faktor-faktor yang memprediksi pencapaian matematika dalam *Programme for International Student Assessment (PISA)*, antara lain prediktor positif meliputi tingkat kelas siswa dan status sosial ekonomi (SES) keluarga secara keseluruhan. Prediktor negatif mencakup absensi siswa, ketidakrukunan, tingkat pengulangan dan putus sekolah, perilaku buruk siswa, kekurangan guru dan staf, serta pengajaran berpusat pada siswa. Dan didukung Nuratiqoh, N., et al. (2024) yang menyatakan bahwa proses berpikir matematis siswa saat menyelesaikan masalah numerasi dapat dilihat dari kecerdasan emosional mereka. Kecerdasan emosional dibagi menjadi tiga level, yaitu level rendah (skor <70), sedang (70-90), dan tinggi (>90), diukur melalui tes standar. Siswa dengan kecerdasan emosional tinggi cenderung lebih efektif dalam mengelola emosi selama pemecahan masalah, menghasilkan strategi kreatif dan refleksi mendalam. Sebaliknya, siswa level rendah sering terhambat oleh frustrasi atau kurangnya regulasi diri.

Dengan demikian perlunya pendekatan holistik dalam evaluasi mutu pendidikan di Indonesia, mengutamakan indikator seperti kompetensi guru, kualitas pembelajaran, dan partisipasi sosial sebagaimana tercantum dalam rapor pendidikan dapat diintegrasikan untuk pengukuran yang lebih akurat dan berkelanjutan. Hal ini membuktikan urgensi inovasi evaluasi yang dapat mengukur potensi akademik secara adaptif dan kontekstual, sehingga lebih mudah dipahami serta relevan untuk perbaikan pembelajaran. Mohan, R. (2023) menyoroti bahwa penilaian berkualitas tinggi, esensial untuk menentukan tujuan masa depan siswa, jalur karier, serta mempengaruhi semua pemangku kepentingan sistem sekolah, sejalan dengan Kebijakan Pendidikan Nasional dan diwujudkan dalam Permendikbud Nomor 23 Tahun 2020 yang mempercepat pergeseran dari tes konvensional ke asesmen kompetensi minimum pada literasi dan numerasi, menekankan keterampilan hidup, literasi digital, dan kesiapan kerja. Hal ini mengubah ketergantungan pada tes kertas-pensil semata, merekomendasikan teknik beragam untuk menilai siswa, kinerja guru, efektivitas sekolah, dan sistem pendidikan secara keseluruhan.

Transformasi model evaluasi pendidikan selama dua dekade di Indonesia dimulai dari sistem ujian kertas-pensil (*paper-based test*) konvensional yang digunakan secara luas sejak lama. Peraturan Pemerintah (PP) No 19 Tahun 2005 menetapkan Ujian Nasional (UN) sebagai ujian kertas-pensil berskala nasional untuk mengukur standar kelulusan berbasis pada penguasaan materi kurikulum yang terstandar secara nasional, dengan fokus pada hafalan dan kemampuan menjawab soal pilihan ganda dan uraian secara manual. Kelemahan utamanya adalah kurang mampu mengukur keterampilan berpikir kritis dan aplikasi konsep, serta rawan kesalahan manusia dalam pengolahan dan penilaian hasil ujian (Kemendikbud, 2005). Selanjutnya, sebagai bagian modernisasi, diperkenalkan ujian berbasis komputer (*computer-based test - CBT*) dengan lembar ujian digital yang memanfaatkan teknologi informasi untuk pelaksanaan ujian yang disebut Ujian Nasional Berbasis Komputer (UNBK) yang telah dijabarkan dalam Permendikbud No 83 Tahun 2015 dan Permendikbud No 40 Tahun 2016. UNBK meningkatkan efisiensi pelaksanaan dan pengolahan hasil, mengurangi manipulasi, serta mempercepat *feedback*. Namun, UNBK masih sering bersifat format evaluasi tradisional yang kurang adaptif terhadap level kemampuan peserta dan tidak selalu mengukur kecakapan berpikir tingkat tinggi secara menyeluruh (Permendikbud, 2015; 2016).

Kemudian, muncul Asesmen Kompetensi Minimum (AKM) sebagai bagian dari kebijakan asesmen nasional yang menekankan pengukuran literasi membaca dan numerasi sebagai kemampuan dasar esensial siswa, diposisikan sebagai asesmen pembelajaran berbasis kompetensi dan keterampilan abad 21. Permendikbud No 23 Tahun 2020 memperkenalkan AKM sebagai bagian merdeka belajar, fokus literasi-numerasi berbasis kompetensi untuk diagnostik

mutu sekolah. AKM menggunakan instrumen berbasis komputer yang mengintegrasikan berbagai tipe soal, termasuk soal berbasis konteks dan soal simulasi, yang dirancang untuk mengukur pemahaman konsep dan penerapan, bukan hanya hafalan. AKM bertujuan memberikan gambaran capaian belajar siswa yang lebih holistik dan mendukung peningkatan mutu pembelajaran yang berkelanjutan (Permendikbud, 2020). Kebijakan terbaru Permendikdasmen No 9 Tahun 2025 menempatkan Tes Kemampuan Akademik (TKA) sebagai inovasi evaluasi asesmen opsional HOTS berbasis IRT-CAT untuk seleksi prestasi dan penyetaraan lintas jalur pendidikan (Kemendikdasmen, 2025).

Kebijakan pendidikan nasional inilah yang memegang peranan penting dalam mendorong perubahan tersebut. Standar nasional pendidikan dan kebijakan asesmen yang dirancang oleh pemerintah harus menjamin akses yang merata, keadilan, serta relevansi dengan kebutuhan lokal dan global (Sondergeld, T. A., et al., 2020). Implementasi kebijakan yang responsif terhadap dinamika pendidikan menjadi kunci keberhasilan transformasi sistem evaluasi. Chaemchoy, S., et al. (2022) dalam penelitiannya yang berfokus pada desain kebijakan transformasi sistem pembelajaran yang responsif terhadap perubahan global seperti disrupsi teknologi, perubahan iklim, dan tuntutan keterampilan abad 21, dengan menekankan pembelajaran seumur hidup, adaptabilitas, dan integrasi literasi digital serta keberlanjutan, mengusulkan kerangka kebijakan yang mencakup empat area prioritas diantaranya: (1) Pengembangan ide baru untuk sistem pendidikan adaptif, (2) Peningkatan keterampilan kritis dan literasi, (3) Efisiensi manajemen pendidikan untuk mengurangi ketimpangan, serta (4) Asesmen berbasis AI dan digital untuk mengukur hasil belajar holistik. Namun, kurangnya kemampuan menggambarkan kompetensi siswa secara holistik dan berkeadilan mendorong pemerintah dalam menentukan kebijakan untuk mencari alternatif asesmen yang lebih adaptif, inklusif, dan mampu mencerminkan *Higher Order Thinking Skills* (HOTS) serta kompetensi abad 21 yang relevan dengan tuntutan global dan lokal (Krstikj, A., et al., 2022).

Inovasi dalam evaluasi pendidikan menjadi sebuah keharusan untuk mengatasi keterbatasan evaluasi konvensional yang cenderung menekankan aspek hafalan. Evaluasi yang inovatif harus dapat mengukur kemampuan siswa dalam berpikir kritis, analitis, dan kreatif secara menyeluruh. Karena itu, saat ini dikembangkan berbagai model asesmen baru yang memanfaatkan teknologi dan pendekatan psikometrik modern agar penilaian menjadi lebih objektif, adil, dan terstandarisasi (Kwangmuang, P., et al., 2021; Bouckaert, M., 2023; Barorah, N. N., et al., 2025). Salah satu inovasi kebijakan yang diangkat adalah Tes Kemampuan Akademik (TKA) yang dirancang sebagai pengganti Ujian Nasional (UN) dengan pendekatan berbasis komputer. TKA bertujuan memberikan penilaian yang lebih valid dan reliabel terhadap kemampuan akademik siswa, serta memberikan akses yang lebih inklusif bagi berbagai jalur pendidikan formal, non formal, dan informal (Kemendikdasmen, 2025). Selain itu, TKA dirancang untuk memberikan data objektif capaian akademik yang digunakan sebagai validator rapor dalam seleksi perguruan tinggi (SNBP/SNBT), indikator penerimaan jenjang berikutnya, dan pemetaan mutu pendidikan daerah, dengan pelaksanaan bertahap mulai SMA/SMK November 2025 dan SD/SMP 2026. Asesmen ini bersifat opsional namun direkomendasikan sebagai nilai tambah, fokus pada HOTS (*Higher Order Thinking Skills*) verbal, numerik, logika, dan spasial, berbeda dari AKM atau Asesmen Sumatif Akhir Tahun. Dengan demikian, TKA diharapkan menjadi alat ukur prestasi yang objektif dan mendukung pemerataan mutu pendidikan.

Lebih lanjut penelitian ini akan menjabarkan beberapa hal meliputi kajian terhadap peran TKA sebagai inovasi evaluasi pendidikan, pelaksanaan TKA sebagai asesmen yang objektif dan terstandarisasi dan mengusung keadilan, serta dampaknya terhadap kebijakan pendidikan dan mutu pembelajaran. Meskipun demikian, adanya celah yang ditemukan terkait dengan implementasi TKA terutama pada aspek kesiapan teknologi di daerah terpencil, pelatihan tenaga pendidik, dan pengawasan pelaksanaan tes yang meritokratik. Hal ini menyebabkan ketimpangan akses dan potensi bias dalam hasil penilaian yang pada akhirnya mempengaruhi akuntabilitas mutu pendidikan (Iversen et al., 2021). Penguraian tersebut menjadi alasan kuat bagi penulis untuk mengangkat artikel mengenai pentingnya menelaah secara kritis dan

komprehensif bagaimana TKA sebagai inovasi evaluasi mampu mengatasi kelemahan evaluasi pendidikan sebelumnya, sekaligus memberikan solusi terhadap tantangan kesetaraan dan mutu pendidikan nasional. Penelitian ini diharapkan akan memberikan kontribusi signifikan dalam pengembangan kebijakan evaluasi pendidikan yang adaptif, inklusif, dan berkeadilan demi mendorong peningkatan kualitas sumber daya manusia Indonesia (Ossiannilsson, E., 2023; Kemendikdasmen, 2025).

## METODE PENELITIAN

Penelitian ini menggunakan pendekatan kualitatif dengan desain tinjauan naratif literatur (*narrative literature review*) (Kulczak-Dawkins, R. P., et al., 2018). Desain ini dipilih untuk mensintesis literatur akademis, buku referensi, dan dokumen kebijakan resmi terkait Tes Kemampuan Akademik (TKA) sebagai inovasi evaluasi pendidikan di Indonesia. Pencarian literatur dilakukan secara berulang dan adaptif dengan menggunakan sumber data dari empat kategori utama, yaitu: (1) jurnal akademis terindeks melalui Google Scholar, SINTA, dan Scopus; (2) buku referensi dan metodologi dari penerbit terpercaya seperti SAGE Publications, Routledge, dan Wiley; (3) dokumen kebijakan resmi termasuk Peraturan Pemerintah, Permendikbud, Permendikdasmen Nomor 9 Tahun 2025, dan Keputusan Menteri Pendidikan Dasar dan Menengah; serta (4) laporan kajian akademik dan *working papers* dari lembaga penelitian pendidikan terpercaya. Sumber-sumber ini dipilih untuk memastikan bahwa tinjauan literatur mencakup perspektif akademis, metodologis, dan kebijakan secara komprehensif Ferrari, R. (2015).

Selain itu, pencarian dengan menggunakan kata kunci utama (*primary keywords*) seperti: "Tes Kemampuan Akademik" OR "TKA", "inovasi evaluasi pendidikan", "penilaian objektif", "asesmen berbasis keadilan", "penilaian terstandarisasi", serta kata kunci sekunder (*secondary keywords*) seperti: "Item Response Theory", "Computer Adaptive Testing", "HOTS", "kesenjangan digital pendidikan", "pengendalian mutu pendidikan", "penyetaraan lintas jalur pendidikan", "pelatihan guru asesmen", dan "integritas akademik online". Proses pencarian diulang dan disesuaikan berdasarkan literatur yang ditemukan (*iterative and adaptive process*), dengan menggali referensi silang dari artikel relevan untuk identifikasi sumber tambahan.

Kemudian data literatur diolah melalui kombinasi dua model analisis yang saling komplementer, yaitu model analisis interaktif dari Miles dan Huberman (1994) dan prosedur analisis tematik reflektif dari Braun dan Clarke (2019, 2021). Data literatur yang dianalisis melalui model analisis interaktif dari Miles dan Huberman (1994), pada model ini melibatkan tiga tahap utama, yaitu: (1) reduksi data; (2) penyajian data; dan (3) penarikan kesimpulan.

Pada tahap reduksi data, literatur yang dikumpulkan dipilih secara selektif berdasarkan relevansi dan kontribusi substantif terhadap peran TKA sebagai inovasi evaluasi pendidikan. Setiap literatur yang masuk kategori inklusi diringkas secara sistematis dengan mengidentifikasi dan mendokumentasikan: (a) identitas lengkap pengarang dan tahun publikasi, (b) fokus penelitian atau topik utama yang dibahas, (c) metodologi atau pendekatan yang digunakan dalam literatur tersebut, (d) temuan atau kontribusi spesifik yang relevan dengan TKA, serta (e) keterkaitannya dengan aspek-aspek evaluasi pendidikan yang ingin dianalisis dalam penelitian ini. Selanjutnya dilakukan pengkodean awal (*initial coding*), di mana setiap literatur diberikan label atau kode yang merepresentasikan aspek-aspek utama yang dibahas. Pengkodean tidak hanya bersifat deskriptif (misalnya: "IRT Model", "CAT Mechanism") tetapi juga interpretatif (misalnya: "Objectivity Challenge", "Equity Barrier", "Infrastructure Gap"). Kode-kode awal yang dikembangkan ini kemudian diorganisir dan dikelompokkan ke dalam enam kategori tematik utama berdasarkan tema substantif yang muncul, seperti "Inovasi Evaluasi Pendidikan", "Asesmen Objektif dan Terstandarisasi", "Keadilan dan Inklusi Pendidikan", "Kebijakan Responsif dan Kontekstual", "Validitas dan Reliabilitas Pengukuran", "Implementasi dan Tantangan Teknis".

Tahapan selanjutnya adalah penyajian data, di mana data literatur disajikan dalam bentuk matriks dan tabel tematik untuk memudahkan identifikasi pola, hubungan, dan frekuensi dari

tema-tema yang muncul. Penyajian data mencakup kategorisasi literatur menurut tahun publikasi yang menunjukkan evolusi penelitian yang paralel dengan transformasi kebijakan evaluasi pendidikan dari 2015 hingga 2025, analisis literatur menurut enam aspek konten, dan matriks literatur tematik yang menampilkan hubungan antara sumber-sumber literatur dengan tema-tema utama, memudahkan pembaca untuk melihat koneksi antar perspektif. Penyajian visual ini membantu peneliti dan pembaca memahami *landscape* penelitian tentang TKA secara komprehensif dan sistematis, serta memberikan transparansi tentang *coverage* dari berbagai aspek yang dianalisis.

Tahapan akhir adalah penarikan kesimpulan yang dilakukan melalui analisis tematik reflektif (*reflexive thematic analysis*) Braun dan Clarke (2019, 2021) yang terdiri dari enam fase, yaitu: (1) familiarisasi data melalui pembacaan berulang abstrak, temuan, dan diskusi; (2) pengkodean awal secara induktif untuk mengidentifikasi pola makna (misalnya, "objektivitas IRT", "kesenjangan digital"); (3) pencarian tema dengan mengelompokkan kode menjadi tema utama seperti "landasan inovatif TKA" dan "mitigasi tantangan"; (4) review tema melalui verifikasi silang sumber untuk konsistensi; (5) definisi dan penamaan tema dengan narasi koheren; serta (6) produksi laporan yang mengintegrasikan temuan dengan konteks pendidikan nasional. Proses ini bersifat reflektif, di mana subjektivitas peneliti sebagai praktisi pendidikan dimanfaatkan untuk membangun interpretasi kontekstual, dengan transparansi tercatat dalam jurnal reflektif pribadi.

Verifikasi data dilakukan melalui triangulasi sumber dari empat kategori, yaitu jurnal terindeks (Google Scholar, SINTA, Scopus), buku referensi (SAGE, Routledge), dokumen kebijakan resmi (Permendikdasmen No. 9/2025), dan laporan kajian, dengan pencarian iteratif menggunakan kata kunci primer/sekunder untuk memastikan kelengkapan dan relevansi. Dan triangulasi metode analisis data yang menggabungkan model interaktif Miles-Huberman (reduksi, penyajian data via matriks tematik, kesimpulan) dengan analisis tematik reflektif Braun-Clarke (familiarisasi, pengkodean induktif, pencarian tema, review tema, penanaman tema, dan produksi laporan), termasuk jurnal reflektif peneliti untuk transparansi subjektivitas kontekstual. Proses ini menjamin validitas melalui kategorisasi enam tema utama (inovasi, asesmen, keadilan, kebijakan, pengukuran, dan implementasi). Verifikasi yang dilakukan bertujuan untuk mengurangi bias subjektif peneliti, memperkuat interpretasi melalui konfirmasi silang antar sumber (jurnal, dokumen kebijakan, buku, dan laporan kajian), serta menghasilkan pemahaman komprehensif tentang peran TKA sebagai inovasi evaluasi pendidikan.

## HASIL DAN PEMBAHASAN

### *Hasil*

#### *Definisi dan Tujuan*

Tes Kemampuan Akademik (TKA) merupakan proses evaluasi terhadap pencapaian kemampuan akademik siswa pada mata pelajaran tertentu yang sudah distandarisasi, bertujuan untuk seleksi pendidikan dan penyetaraan lintas jalur. Selain itu, TKA membantu meningkatkan kompetensi pendidik dalam menyusun penilaian yang bermutu serta menjaga kualitas pendidikan secara konsisten. Pelaksanaan TKA dilaksanakan melalui sistem komputer dengan metode tes secara daring penuh atau semi daring (Pusmendik, 2025).

Tes Kemampuan Akademik (TKA) juga merupakan evaluasi yang dirancang untuk mengukur kemampuan akademik siswa secara menyeluruh, dengan fokus pada kemampuan berpikir kritis, analitis, dan pemecahan masalah. Berbeda dengan Ujian Nasional yang lebih menekankan penguasaan materi kurikulum, TKA menilai literasi, numerasi, dan keterampilan penalaran pada berbagai mata pelajaran sesuai jurusan siswa. Selanjutnya dalam Permendikdasmen Nomor 9 Tahun 2025 dijelaskan mengenai tujuan utama TKA adalah memberikan informasi capaian akademik yang lebih terstandar, menjamin pemenuhan akses murid pendidikan nonformal dan pendidikan informal terhadap penyetaraan hasil belajar, meningkatkan kualitas pembelajaran dan penilaian, serta pengendalian dan penjaminan mutu

pendidikan (Kemendikdasmen, 2025). TKA bersifat opsional, memberi fleksibilitas bagi siswa mengikuti tes sesuai kesiapan mereka.

### ***Landasan Teoritis (Psikometri)***

Secara psikometrik, TKA berdasar pada teori pengukuran yang memastikan validitas dan reliabilitas tes dalam mengukur kemampuan akademik. Psikometri menggunakan prinsip pengujian objektif serta model teori respons butir (*item response theory*) dan teori tes klasik untuk menghasilkan penilaian yang akurat dan adil (Wilson, M., 2023). Pendekatan ini memungkinkan penilaian kemampuan siswa secara konsisten dengan mempertimbangkan tingkat kesukaran soal dan kemampuan individu. Landasan ini menjamin bahwa hasil TKA valid untuk digunakan sebagai indikator prestasi akademik dan prediksi keberhasilan akademik pada tingkatan lebih lanjut (Castillo-Diaz, M. A., & Gomes, C. M. A., 2022).

Embretson dan Reise (2013) menyimpulkan bahwa IRT merevolusi psikometri dengan model probabilistik yang menghasilkan pengukuran kemampuan invarian terhadap sampel, mengatasi limitasi CTT melalui parameter butir spesifik mengatasi kelemahan seperti *difikultas*, *diskriminasi*, dan *guessing*. Prinsip dasar Teori Respons Butir (IRT) melalui model biner untuk jawaban benar/salah, model politomus untuk pilihan ganda bertingkat, serta skala trait laten yang mengukur kemampuan tersembunyi peserta. Proses analisis dalam Teori Respons Butir (IRT) mencakup penskoran peserta via estimasi maksimum kemungkinan, kalibrasi butir untuk menentukan kesulitan dan diskriminasi, dikombinasikan dengan uji kecocokan model CFA untuk memverifikasi unidimensi dan monotonicitas, mendukung pengembangan tes adaptif komputerisasi (CAT) yang reliabel dan invarian seperti pada TKA (Ferrando, P. J., & Lorenzo-Seva, U., 2018).

### ***Komponen dan Struktur***

TKA tersusun atas beberapa komponen utama yang mencakup subtes verbal, numerik, logika, dan spasial, dengan materi yang menyesuaikan kurikulum nasional dan program jurusan siswa seperti IPA, IPS, dan Bahasa sebagaimana telah dijabarkan dalam Keputusan Menteri Pendidikan Dasar dan Menengah Republik Indonesia Nomor 102/M/2025. Setiap komponen mengukur aspek spesifik dari kemampuan akademik, misalnya kemampuan bahasa, pemahaman konsep matematika, penalaran logis, dan kemampuan visual-spasial (Mercan, Z., & Kandir, A., 2024; Tiwari, S., et al., 2024). Struktur tes dirancang sistematis dengan kisi-kisi yang jelas untuk menjamin cakupan materi yang representatif dan evaluasi kemampuan yang menyeluruh.

Desain asesmen TKA mengintegrasikan tiga tipe soal utama, yaitu: (1) Pilihan Ganda (PG) Sederhana; (2) Pilihan Ganda Kompleks (PGK)-*Multiple Choice Multiple Answers* (MCMA); dan (3) Pilihan Ganda Kompleks (PGK)-Kategori, untuk merangsang inovasi metode belajar mengajar di kelas, terutama soal yang menuntut pemahaman mendalam bukan sekadar hafalan. Tiap domain ujian mencakup item independen (soal tunggal) dan item berbasis stimulus bersama (soal grup). Soal tunggal berdiri sendiri, sedangkan soal grup menghubungkan beberapa pertanyaan pada satu konteks stimulus. Tema soal dibuat kontekstual dengan realitas sehari-hari, seperti strategi penanggulangan bencana mitigasi bencana (geografi), pengolahan zat berbahaya pengelolaan limbah kimia (kimia), serta interaksi budaya masyarakat dinamika sosial budaya (sosiologi dan antropologi) (Handayani, H., 2025).

### ***Prinsip dan Standarisasi***

Pelaksanaan TKA berpedoman pada prinsip validitas, reliabilitas, dan standarisasi tinggi untuk memastikan hasil tes dapat dipercaya dan adil. Standarisasi mencakup prosedur pembuatan soal, pelaksanaan ujian, sampai penilaian hasil, dengan pengawasan ketat terhadap kualitas materi dan sarana pendukung. Prinsip transparansi dan akuntabilitas juga diutamakan agar setiap peserta mendapat perlakuan yang setara. Standar ini dimaksudkan untuk

menciptakan evaluasi akademik yang objektif serta mendukung pengambilan keputusan pendidikan berbasis data yang kredibel.

Kajian akademik Krisna., et al. (2025) secara eksplisit menguraikan prinsip-prinsip Tes Kemampuan Akademik (TKA) sebagai fondasi reformasi evaluasi pendidikan nasional, meliputi pengukuran capaian akademik individu yang terstandar untuk seleksi akademik, penyetaraan lintas jalur pendidikan, dan diagnostik mutu tanpa pengaruh kelulusan. Prinsip inti mencakup objektivitas melalui asesmen HOTS berbasis penalaran kritis bukan hafalan, keadilan akses dengan pelaksanaan gratis dan inklusif bagi disabilitas, serta integrasi dengan Asesmen Nasional (ANBK) demi umpan balik pembelajaran holistik. Pelaksanaan TKA menekankan kejujuran, kerahasiaan, akuntabilitas, dan partisipasi sukarela, selaras dengan Permendikdasmen No. 9/2025 untuk transformasi evaluasi multidimensi yang relevan secara global.

**Tabel 1.** Literatur yang Relevan dengan TKA

No	Penulis (Tahun)	Fokus Penelitian	Metode	Temuan yang Relevan dengan TKA
1	Schult, J., et al. (2015)	Mengkaji kecukupan tes kemampuan skolastik termasuk aspek penalaran verbal, numerik, dan figural secara komprehensif.	Kuantitatif dengan analisis SEM	TKA mencakup aspek penalaran seperti verbal, numerik, dan figural sebagaimana dibuktikan oleh analisis tes <i>scholastic aptitude</i> .
2	Coyle, T.R. (2015)	Menganalisis hubungan linear dan nonlinear antara kecerdasan umum, tes aptitudo (SAT, ACT, PSAT, ASVAB), dan IPK kuliah.	Analisis regresi linear dan kuadrat	TKA sebagai salah instrumen pengukuran kemampuan kognitif hubungannya dengan kecerdasan umum dan prestasi akademik.
3	Castillo-Diaz, M. A., & Gomes, C. M. A. (2022)	Menguji peran metakognitif monitoring dan kecerdasan sebagai prediktor prestasi akademik umum serta spesifik di pendidikan tinggi menggunakan ukuran standar seperti tes nasional masuk universitas.	kuantitatif korelasional dengan regresi hierarkis	Monitoring metakognitif dan kecerdasan memprediksi hasil TKA sebagai tolak ukur prestasi akademik umum dan spesifik di perguruan tinggi.

No	Penulis (Tahun)	Fokus Penelitian	Metode	Temuan yang Relevan dengan TKA
4	Kubiszyn, T., & Borich, G. D. (2024)	Mengkritik evaluasi pendidikan konvensional (NCLB) melalui ESSA dengan menekankan integrasi (MTSS) dan (RTI), untuk meningkatkan validitas dan keadilan pengukuran.	Bab Buku: norm-referenced testing (NRT) dan criterion-referenced testing (CRT)	TKA bersifat optional dan lebih berkeadilan sebagai asesmen berbasis kompetensi bukan sebagai penentu kelulusan mutlak menggantikan UN yang sering dikritik karena bertekanan tinggi dan sebagai satu-satunya penentu kelulusan.
5	Nye, C. D., Ma, J., & Wee, S. (2022)	Menguji validitas kemampuan kognitif sempit yang kurang berkorelasi dengan <i>general mental ability</i> (GMA) dalam memprediksi berbagai dimensi kinerja kerja.	meta-analisis untuk menguji validitas inkremental kemampuan kognitif sempit terhadap kinerja kerja di luar GMA.	TKA mengukur kemampuan kognitif sempit yang memvalidasi kinerja kerja, dapat digunakan dalam seleksi dan penerimaan karyawan dengan cara yang efektif dan sesuai prosedur.
6	Wilson, M. (2023)	Pendekatan <i>Item Response Theory</i> (IRT) membangun instrumen pengukuran akurat via estimasi parameter butir dan kemampuan.	Bab Buku: IRT (model Rasch) melalui empat blok: <i>construct map, desain item, outcome space, dan model statis</i>	Evaluasi penggunaan <i>Item Response Theory</i> (IRT) dalam standarisasi TKA memastikan konstruksi tes kemampuan akademik yang akurat.
7	Yan, D., et al. (2020)	Aplikasi praktis sistem penilaian otomatis untuk penilaian pembelajaran kompleks.	Tinjauan komprehensi f teori, dan studi kasus praktis.	TKA menyediakan kerangka penilaian otomatis untuk mengevaluasi respon tes kemampuan akademik secara objektif dan skalabel.

No	Penulis (Tahun)	Fokus Penelitian	Metode	Temuan yang Relevan dengan TKA
8	Opposs, D., et al. (2020)	Dampak struktur pemerintahan terhadap kekuasaan dan objektivitas penetapan standar penilaian pendidikan nasional.	Studi kasus berganda holistic	TKA yang dikelola pusat oleh Kemendikdasmen, menjadikannya sebagai asesmen yang objektif dan terstandarisasi.
9	Huda, A., et al. (2024)	Membuktikan CAT berbasis IRT 3-PL sangat praktis (91% konten, 90% tampilan) untuk asesmen adaptif.	model 4-D dengan IRT 3-PL untuk CAT yang merandomisasi soal adaptif hingga $SE \leq 0.01$ .	CAT unggul bagi TKA dengan efisiensi waktu tes (rata-rata 12 CAT vs 34 CBT), akurasi tinggi via IRT, minim <i>guessing</i> , dan aman dari kecurangan.
10	Cumming, T., & Miller, M. D. (Eds.). (2023)	Penerapan prinsip psikometri praktis untuk meningkatkan validitas dan reliabilitas asesmen di pendidikan tinggi melalui analisis dan standarisasi tes yang mendukung keputusan berbasis bukti.	Bab Buku: Metode psikometri dalam asesmen pendidikan melibatkan pengembangan tes dan analisis statistik	Prinsip psikometri praktis untuk meningkatkan validitas, reliabilitas, dan standarisasi asesmen akademik, yang mendukung desain TKA sebagai tes terstandar nasional berbasis HOTS.
11	Pambudi, N. A., & Harjanto, B. (2020)	Salah satu uraiannya mengenai tantangan Pendidikan vokasi di Indonesia meliputi infrastruktur yang tidak lengkap, terutama daerah terpencil tanpa internet/komputer, kurikulum teoritis, minim kemitraan industri, guru kurang kompeten teknologi, lulusan tidak siap kerja.	literature review komprehensif dengan analisis induktif, deduktif, dan meta-analysis	Tantangan infrastruktur vokasi (kekurangan lab/internet di daerah terpencil) dan kurikulum teoritis menghambat kesiapan siswa menghadapi asesmen adaptif TKA yang menuntut kompetensi praktis.

No	Penulis (Tahun)	Fokus Penelitian	Metode	Temuan yang Relevan dengan TKA
12	Iversen, V., et al. (2021)	Sintesis multidisiplin yang mempengaruhi mobilitas sosial di negara berkembang, termasuk celah pengetahuan, hambatan dan peluang untuk meningkatkan kesempatan naik secara sosial dan mengurangi ketidaksetaraan.	Bab Buku: sintesis literatur multidisiplin dengan analisis komparatif data survei	TKA sebagai asesmen standar nasional berpotensi memfasilitasi mobilitas sosial vertikal, naik melalui akses pendidikan dan seleksi prestasi yang inklusif, sambil mengurangi risiko penurunan status akibat ketidaksetaraan akses Pendidikan.

## Pembahasan

### Asesmen yang Objektif dan Terstandarisasi

Dalam dokumen Permendikdasmen Nomor 9 Tahun 2025 tentang Tes Kemampuan Akademik (TKA), objektivitas diwujudkan dengan prinsip pelaksanaan penilaian yang bersifat netral, adil, dan bebas dari intervensi subjektif (Sireci, S. G., & Randall, J., 2021; Baniyadi, A., et al., 2023). TKA dirancang sebagai tes standar nasional yang menggunakan sistem berbasis komputer dengan pengawasan ketat untuk memastikan hasil penilaian mencerminkan capaian akademik murid secara valid dan dapat dipertanggungjawabkan (Yan, D., et al., 2020). Pelaksanaan TKA dijaga integritasnya melalui pengawas silang, pengawas daring dari perguruan tinggi (Cagala, T., et al., 2024), dan infrastruktur teknologi yang memadai. Selain itu, hasil TKA tidak menjadi faktor kelulusan, sehingga menghindari tekanan untuk memanipulasi hasil, melainkan berfungsi sebagai bahan refleksi dan perbandingan capaian murid di tingkat nasional. Dengan demikian, TKA memberikan gambaran objektif tentang kemampuan akademik siswa yang dapat digunakan bagi pihak sekolah, pemerintah daerah, dan kementerian untuk memperbaiki mutu pembelajaran dan penilaian secara keseluruhan.

Meskipun TKA dirancang tidak wajib dan bukan syarat kelulusan, realitasnya hasil TKA kerap berperan dalam seleksi masuk perguruan tinggi jalur prestasi sehingga menjadikannya *high-stakes test* (Kubiszyn, T., & Borich, G. D., 2024). Oleh karena itu, perlu penguatan pelatihan, penyesuaian bertahap, dan sosialisasi transparan agar TKA menjadi evaluasi akademik nasional yang adil, valid, bermanfaat, dan mampu meningkatkan mutu pendidikan secara menyeluruh.

### Tantangan Keadilan

Tes Kemampuan Akademik (TKA) menuju asesmen yang adil diwujudkan melalui prinsip-prinsip penyelenggaraan yang inklusif dan setara (Scott, S., et al., 2014; Barroilhet, A., et al., 2025), di mana TKA dapat diikuti oleh murid dari jalur pendidikan formal, non formal, serta informal termasuk paket A/B/C, pesantren, dan sekolah rumah, sehingga menjamin akses penyeteraan hasil belajar bagi semua peserta didik tanpa diskriminasi jenjang atau latar belakang, sebagaimana tercantum dalam Pasal 3 huruf b dan Pasal 8 dalam Permendikdasmen nomor 9 tahun 2025. Keadilan lebih lanjut ditegakkan melalui standarisasi nasional yang objektif, dengan Kementerian menetapkan pedoman, sistem, dan kerangka asesmen TKA untuk memastikan pengukuran capaian akademik murid terstandar di seluruh jenjang, menghindari variasi lokal yang berpotensi tidak adil, serta memungkinkan hasil TKA digunakan untuk seleksi akademik

seperti PPDB jalur prestasi dan penerimaan perguruan tinggi secara transparan (Hossler, D., 2022).

Prinsip kejujuran, kerahasiaan, dan akuntabilitas pada Pasal 2, menjadi pondasi keadilan operasional, di mana pelanggaran seperti kebocoran soal dicegah untuk menjaga integritas proses bagi semua peserta (Holden, O. L., et al., 2021; Alessio, H. M., & Messinger, J. D., 2021; Garg, M., & Goel, A., 2022). Sementara pengecualian bagi murid berkebutuhan khusus dengan disabilitas intelektual memastikan aksesibilitas yang proporsional, sehingga TKA menjadi instrumen pengendalian mutu pendidikan yang berkeadilan nasional.

Meskipun demikian, ketimpangan infrastruktur dan sumber daya antar sekolah menyebabkan peluang akses yang tidak setara (Pambudi, N. A., & Harjanto, B., 2020), kesiapan tenaga pendidik dan operator yang belum merata, serta risiko penyalahgunaan hasil tes dalam seleksi tanpa mempertimbangkan konteks sosial (Iversen, V., et al., 2021), dan prestasi non-akademik siswa menjadi tantangan mewujudkan keadilan dalam pelaksanaan TKA. Selain itu, pengelolaan data dan pengawasan yang belum optimal dapat mengganggu integritas (Reisenwitz, T. H., 2020) dan objektivitas pelaksanaan sehingga membutuhkan perbaikan kebijakan, transparansi, serta dukungan afirmatif bagi daerah tertinggal agar TKA benar-benar menjadi alat evaluasi yang adil dan merata.

### ***Inovasi Evaluasi Pendidikan***

Tes Kemampuan Akademik (TKA) merupakan inovasi evaluasi tercermin dari penggantian Ujian Nasional (UN) yang sebelumnya traumatis (Von der Embse, N., 2018; Tan, S. H., & Pang, J. S., 2023) dan berisiko kelulusan, TKA di desain sebagai asesmen opsional berbasis standar nasional yang tidak menentukan kelulusan, melainkan fokus pada pengukuran capaian akademik murid secara objektif untuk seleksi prestasi, penyeteraan hasil belajar nonformal/informal, dan pengendalian mutu pendidikan. Inovasi ini diwujudkan melalui pendekatan kolaboratif (Zamjani, I., 2022; Cahyanto, I., 2025), di mana Kementerian menyusun soal untuk jenjang SMA/MA/SMK/MAK serta mengelola data secara terpusat (Opposs, D., et al., 2020; ), sementara pemerintah daerah kabupaten/kota menyusun soal SD/MI/SMP/MTs dengan penjaminan mutu provinsi, didukung sistem berbasis komputer untuk akuntabilitas dan inklusivitas bagi semua jalur pendidikan, termasuk pesantren dan sekolah rumah.

TKA menjadi terobosan dengan prinsip kejujuran, kerahasiaan, dan akuntabilitas (Pasal 2), menghasilkan Sertifikat Hasil TKA (SHTKA) sebagai referensi nasional untuk PPDB jalur prestasi dan seleksi perguruan tinggi, mendorong peningkatan kapasitas pendidik dalam penilaian berkualitas serta memberikan data empiris untuk kebijakan mutu pendidikan yang adaptif dan berkeadilan (Palanisamy, P., 2021). Namun, tantangan implementasi seperti kesenjangan infrastruktur digital, keterbatasan SDM, gangguan jaringan, tekanan psikis siswa, dan potensi *teaching to the test* tetap ada, sehingga memerlukan pelatihan intensif, serta penyesuaian bertahap agar TKA benar-benar adil dan efektif.

### ***Implikasi Pengembangan TKA***

Implementasi TKA membawa implikasi strategis bagi ekosistem pendidikan nasional, diantaranya: (1) diagnostik mutu pembelajaran melalui data empiris sekolah ke daerah lanjut ke pusat untuk reformasi kurikulum berbasis kebutuhan (Oo, C. Z., Alonzo, D., & Davison, C., 2023); (2) peningkatan daya saing lulusan via Sertifikat Hasil TKA (SHTKA) untuk PPDB prestasi, beasiswa, dan rekrutmen kerja (Pucciarelli, F., & Kaplan, A., 2016); 3) mengurangi kecemasan tes melalui format adaptif *non-high-stakes* (Akhtar, H., 2023); serta (4) transformasi guru menjadi fasilitator berbasis data dengan pelatihan asesmen IRT dan STEAM untuk keterampilan abad-21 (Trantham, P. S., 2022; Wu, Z., 2024). Dengan demikian, TKA berpotensi merevolusi pendidikan Indonesia dari evaluasi sumatif menjadi sistem pembelajaran adaptif yang inklusif dan berkeadilan.

Namun, implikasi TKA juga menuntut mitigasi risiko seperti kesenjangan digital antar daerah, tekanan psikis siswa akibat asesmen berlapis, dan potensi *teaching to the test* yang

mengurangi kreativitas pembelajaran. Oleh karena itu, roadmap implementasi harus mencakup pilot CAT-IRT di wilayah prioritas (Weiss, D. J., & Sahin, A., 2024), pelatihan nasional guru via MoOC (*Massive Open Online Course*) (Ahmad, I., 2022), serta dashboard analitik *real-time* untuk *feedback* adaptif sekolah ke provinsi dan lanjut ke pusat (Hasnine, M. N., 2023), memastikan TKA berevolusi menjadi ekosistem pembelajaran inklusif yang berbasis data dan berorientasi siswa.

Dengan demikian, TKA bukan sekadar pengganti UN, melainkan transformasi paradigma evaluasi pendidikan nasional menuju sistem adaptif, inklusif, dan berbasis data yang mendorong pemerataan mutu pembelajaran serta daya saing global generasi muda Indonesia (Kemendikdasmen, 2025). Keberhasilan implementasinya bergantung pada sinergi lintas sektor, pemantauan berkelanjutan, dan adaptasi berbasis *feedback* empiris, menjadikan TKA sebagai tonggak revolusi pendidikan berkualitas bagi semua.

## KESIMPULAN

Tes Kemampuan Akademik (TKA) berperan sebagai inovasi evaluasi pendidikan dengan menggantikan Ujian Nasional (UN) melalui asesmen opsional berbasis komputer (IRT-CAT), fokus pada pengukuran HOTS verbal, numerik, logika, dan spasial untuk seleksi prestasi serta penyetaraan lintas jalur pendidikan. Sebagai asesmen objektif dan terstandarisasi, TKA mengusung keadilan melalui prinsip netralitas, inklusivitas bagi formal/nonformal/informal, dan integritas tanpa menentukan kelulusan, sehingga mengurangi bias subjektif serta tekanan *high-stakes*. TKA berdampak pada kebijakan pendidikan dengan menyediakan data empiris untuk diagnostik mutu, reformasi kurikulum, dan peningkatan pembelajaran adaptif, sekaligus mengatasi keterbatasan evaluasi sebelumnya seperti hafalan *oriented* pada UN/UNBK/AKM yang kurang adaptif terhadap kemampuan individu. Solusi kesetaraan dan mutu pendidikan dicapai melalui pilot bertahap, pelatihan guru MoOC, dukungan infrastruktur afirmatif, serta *dashboard real time*, memastikan akses merata meskipun tantangan kesenjangan digital tetap ada.

## DAFTAR PUSTAKA

- Ahmad, I., Sharma, S., Singh, R., Gehlot, A., Priyadarshi, N., & Twala, B. (2022). MOOC 5.0: A Roadmap to the Future of Learning. *Sustainability*, 14(18), 11199. <https://doi.org/10.3390/su141811199>
- Akhtar, H., Silfiasari, Vekety, B., & Kovacs, K. (2023). The effect of computerized adaptive testing on motivation and anxiety: a systematic review and meta-analysis. *Assessment*, 30(5), 1379-1390. <https://doi.org/10.1177/10731911221100995>
- Alessio, H. M., & Messinger, J. D. (2021, April). Faculty and student perceptions of academic integrity in technology-assisted learning and testing. In *Frontiers in Education* (Vol. 6, p. 629220). Frontiers Media SA. <https://doi.org/10.3389/feduc.2021.629220>
- Baniasadi, A., Salehi, K., Khodaie, E., Bagheri Noaparast, K., & Izanloo, B. (2023). Fairness in classroom assessment: A systematic review. *The Asia-Pacific Education Researcher*, 32(1), 91-109. <https://doi.org/10.1007/s40299-021-00636-z>
- Barorah, N. N., Ika, V. E., Nasrun, & Kadek, S. W. N. (2025). Holistic Approach in Education to Promote Critical Thinking and Creativity. *The New Educational Review*, 79, 162-170. <http://doi.org/10.15804/tner.2025.79.1.11>
- Barroilhet, A., Silva, M., & Geisinger, K. F. (2025). The Right to be Fairly Assessed. *Higher Education Policy*, 38(2), 408-425. <https://doi.org/10.1057/s41307-024-00352-5>
- Bouckaert, M. (2023). The assessment of students' creative and critical thinking skills in higher education across OECD countries: A review of policies and related practices. *OECD Education Working Papers*, (293), 0\_1-46. <https://dx.doi.org/10.1787/35dbd439-en>
- Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health*, 11(4), 589-597. <https://doi.org/10.1080/2159676X.2019.1628806>
- Braun, V., & Clarke, V. (2021). Thematic analysis: A practical guide. <https://www.torrossa.com/it/publishers/sage-publications.html>

- Cagala, T., Glogowsky, U., & Rincke, J. (2024). Detecting and preventing cheating in exams: Evidence from a field experiment. *Journal of Human Resources*, 59(1), 210-241. <https://doi.org/10.3368/jhr.0620-10947R1>
- Cahyanto, I. (2025). Designing a stakeholder engagement model in educational policy decision-making. *Edum: Jurnal Pendidikan*, 11(2), 327-340. <https://doi.org/10.31943/edumjournal.v8i1.327>
- Castillo-Diaz, M. A., & Gomes, C. M. A. (2022). Monitoring and intelligence as predictors of a standardized measure of general and specific higher education achievement. *Trends in Psychology*, 30(4), 821-841. <http://doi.org/10.1007/s43076-022-00160-z>
- Chaemchoy, S., Siribanpitak, P., Usaho, C., Koraneekij, P., Petpon, P., & Chaimongkol, N. (2022). Policy design for transforming learning systems responsive to future global changes in Thailand 2040. *Kasetsart Journal of Social Sciences*, 43(2), 509-516. retrieved from <https://so04.tci-thaijo.org/index.php/kjss/article/view/258512>
- Coyle, T. R. (2015). Relations among general intelligence (g), aptitude tests, and GPA: Linear effects dominate. *Intelligence*, 53, 16-22. <https://doi.org/10.1016/j.intell.2015.08.005>
- Cumming, T., & Miller, M. D. (Eds.). (2023). *Enhancing assessment in higher education: Putting psychometrics to work*. Taylor & Francis.
- Deb, S., Strodl, E., & Sun, H. (2015). Academic stress, parental pressure, anxiety and mental health among Indian high school students. *International Journal of Psychology and Behavioral Science*, 5(1), 26-34. <https://doi.org/10.5923/j.ijpbs.20150501.04>
- Embretson, S. E., & Reise, S. P. (2013). *Item response theory for psychologists*. Psychology Press. <https://doi.org/10.4324/9781410605269>
- Ferrando, P. J., & Lorenzo-Seva, U. (2018). Assessing the quality and appropriateness of factor solutions and factor score estimates in exploratory item factor analysis. *Educational and psychological measurement*, 78(5), 762-780. <https://doi.org/10.1177/0013164417719308>
- Ferrari, R. (2015). Writing narrative style literature reviews: Definitions, approaches, evaluations. *Journal of the European Medical Writers Association*, 24(4), 230-235. <https://doi.org/10.1179/2047480615Z.000000000290>
- Garg, M., & Goel, A. (2022). A systematic literature review on online assessment security: Current challenges and integrity strategies. *Computers & Security*, 113, 102544. <https://doi.org/10.1016/j.cose.2021.102544>
- Handayani, H. (2025). Kerangka Asesmen Tes Kompetensi Akademik (TKA). <https://id.scribd.com/document/946939798/Kerangka-Asesmen-Tes-Kompetensi-Akademik-TKA>
- Hasnine, M. N., Nguyen, H. T., Tran, T. T. T., Bui, H. T., Akçapınar, G., & Ueda, H. (2023). A real-time learning analytics dashboard for automatic detection of online learners' affective states. *Sensors*, 23(9), 4243. <https://doi.org/10.3390/s23094243>
- Holden, O. L., Norris, M. E., & Kuhlmeier, V. A. (2021). Academic integrity in online assessment: A research review. In *Frontiers in education* (Vol. 6, p. 639814). Frontiers Media SA. <https://doi.org/10.3389/feduc.2021.639814>
- Hossler, D. (2022). The Interconnectedness of issues confronting four-year colleges in the area of admissions, transparency, and equity: Ethical considerations. In *College Admissions and Admissions Testing in a Time of Transformational Change* (pp. 32-52). Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9781003326823-3/interconnectedness-issues-confronting-four-year-colleges-area-admissions-transparency-equity-hossler>
- <https://gurudikdas.dikdasmen.go.id/news/mengkaji-kembali-hasil-pisa-sebagai-pendekatan-inovasi-pembelajaran--untuk-peningkatan-kompetensi-li>
- <https://lldikti13.kemdikbud.go.id/wp-content/uploads/2020/03/Salinan-Keputusan-Menteri-Pendidikan-dan-Kebudayaan-8203.pdf>
- <https://peraturan.bpk.go.id/Details/108758/permendikbud-no-40-tahun-2016>
- <https://peraturan.bpk.go.id/Details/163751/permendikbud-no-23-tahun-2020>

- <https://peraturan.bpk.go.id/Details/39691/permendikbud-no-83-tahun-2015>  
<https://peraturan.bpk.go.id/Details/40249/pp-no-19-tahun-2005>  
<https://pusmendik.kemdikbud.go.id/konten/perilisan-hasil-pisa-2022-peringkat-indonesia-naik-5-6-posisi>  
<https://pusmendik.kemendikdasmen.go.id/tka/page/download>
- Huda, A., Firdaus, F., Irfan, D., Hendriyani, Y., Almasri, A., & Sukmawati, M. (2024). Optimizing educational assessment: The practicality of computer adaptive testing (CAT) with an item response theory (IRT) approach. *JOIV: International Journal on Informatics Visualization*, 8(1), 473-480. <https://dx.doi.org/10.62527/joiv.8.1.2217>
- Iversen, V., Krishna, A., & Sen, K. (2021). *Social mobility in developing countries: Concepts, methods, and determinants* (p. 512). Oxford University Press. <https://doi.org/10.1093/oso/9780192896858.001.0001>
- Krisna, F. N. A., Ulumudin, I., Handayani, M., Purnama, J., Krismahardhika, B. G., Rachmat, H., ... & Prabawa, H. A. (2025). Kajian Akademik Kebijakan Evaluasi Pendidikan Indonesia. <https://pskp.kemendikdasmen.go.id/analisis-kebijakan/risalah-kebijakan>
- Krstikj, A., Sosa Godina, J., García Bañuelos, L., González Peña, O. I., Quintero Milián, H. N., Urbina Coronado, P. D., & Vanoye García, A. Y. (2022). Analysis of Competency Assessment of Educational Innovation in Upper Secondary School and Higher Education: A Mapping Review. *Sustainability*, 14(13), 8089. <https://doi.org/10.3390/su14138089>
- Kubiszyn, T., & Borich, G. D. (2024). *Educational testing and measurement*. John Wiley & Sons.
- Kulczak-Dawkins, R. P., Piroozfar, P. A., & Harder, M. K. (2018). A grounded thematic narrative literature review method. *Building Meaningful Futures*. <https://buildingmeaningfulfutures.org/wp-content/uploads/2024/05/kulczak-dawkins-et-al-2018a-meth1-v2.3-gtn-lit-review-method.pdf>
- Kwangmuang, P., Jarutkamolpong, S., Sangboonraung, W., & Daungtod, S. (2021). The development of learning innovation to enhance higher order thinking skills for students in Thailand junior high schools. *Heliyon*, 7(6). <https://doi.org/10.1016/j.heliyon.2021.e07309>
- Mercan, Z., & Kandır, A. (2024). The effect of the Early STEAM Education Program on the visual-spatial reasoning skills of children: research from Turkey. *Education 3-13*, 52(2), 123-153. <https://doi.org/10.1080/03004279.2022.2075906>
- Miles, M. B. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks.
- Mohan, R. (2023). *Measurement, evaluation and assessment in education*. PHI Learning Pvt. Ltd.
- Nikum, K. (2022). Answers to the Societal Demands with Education 5.0 : Indian Higher Education System. *Journal of Engineering Education Transformations*, 115-127. Retrieved from <https://journaleet.in/index.php/jeet/article/view/250>
- Nuratiqoh, N., Hidayanto, E., & Chandra, T. D. (2024). Deskripsi Berpikir Matematis Siswa Pada Penyelesaian Masalah AKM Numerasi Ditinjau Berdasarkan Level Kecerdasan Emosional [Description Of Students' Mathematical Thinking In Solving AKM Numeracy Problems Based On Emotional Intelligence Levels]. *JOHME: Journal of Holistic Mathematics Education*, 8(1), 18-44. <https://doi.org/10.19166/johme.v8i1.8002>
- Nye, C. D., Ma, J., & Wee, S. (2022). Cognitive ability and job performance: Meta-analytic evidence for the validity of narrow cognitive abilities. *Journal of Business and Psychology*, 37(6), 1119-1139. <https://doi.org/10.1007/s10869-022-09796-1>
- Oo, C. Z., Alonzo, D., & Davison, C. (2023). Using a Needs-Based Professional Development Program to Enhance PreService Teacher Assessment for Learning Literacy. *International Journal of Instruction*, 16(1), 781-800. Retrieved from <https://eiji.net/ats/index.php/pub/article/view/232>
- Opposs, D., Baird, J. A., Chankseliani, M., Stobart, G., Kaushik, A., McManus, H., & Johnson, D. (2020). Governance structure and standard setting in educational assessment. *Assessment in Education: Principles, Policy & Practice*, 27(2), 192-214. Yan, D., Rupp, A. A., & Foltz, P. W.

- (Eds.). (2020). *Handbook of automated scoring: Theory into practice*. CRC Press. <https://doi.org/10.1080/0969594X.2020.1730766>
- Ossiannilsson, E. (2023). Open educational resources (OER) and some of the United Nations sustainable development goals. *The International Journal of Information and Learning Technology*, 40(5), 548-561. <https://doi.org/10.1108/IJILT-01-2023-0002>
- Palanisamy, P., Thilarajah, S., & Chen, Z. (2021). Providing equitable education through personalised adaptive learning and assessment. *ASCILITE Publications*, 202-207. <https://doi.org/10.14742/ascilite2021.0129>
- Pambudi, N. A., & Harjanto, B. (2020). Vocational education in Indonesia: History, development, opportunities, and challenges. *Children and Youth Services Review*, 115, 105092. <https://doi.org/10.1016/j.childyouth.2020.105092>
- Pucciarelli, F., & Kaplan, A. (2016). Competition and strategy in higher education: Managing complexity and uncertainty. *Business horizons*, 59(3), 311-320. <https://doi.org/10.1016/j.bushor.2016.01.003>
- Reisenwitz, T. H. (2020). Examining the necessity of proctoring online exams. *Journal of Higher Education Theory and Practice*, 20(1), 118-124. Examining the necessity of proctoring online exams. *Journal of Higher Education Theory and Practice*, 20(1), 118-124.
- Rodríguez, J. V., Rodado, D. N., Borrero, T. C., & Parody, A. (2022). Multidimensional indicator to measure quality in education. *International Journal of Educational Development*, 89, 102541. <https://doi.org/10.1016/j.ijedudev.2021.102541>
- Schult, J., Fischer, F. T., & Hell, B. (2015). Tests of scholastic aptitude cover reasoning facets sufficiently. *European Journal of Psychological Assessment*. <https://doi.org/10.1027/1015-5759/a000247>
- Scott, S., Webber, C. F., Lupart, J. L., Aitken, N., & Scott, D. E. (2014). Fair and equitable assessment practices for all students. *Assessment in Education: Principles, Policy & Practice*, 21(1), 52-70. <https://doi.org/10.1080/0969594X.2013.776943>
- Sireci, S. G., & Randall, J. (2021). Evolving notions of fairness in testing in the United States. In *The history of educational measurement* (pp. 111-135). Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9780367815318-6/evolving-notions-fairness-testing-united-states-stephen-sireci-jennifer-randall>
- Sondergeld, T. A., Stone, G. E., & Kruse, L. M. (2020). Objective standard setting in educational assessment and decision making. *Educational Policy*, 34(5), 735-759. <https://doi.org/10.1177/0895904818802115>
- Sukmayadi, V. & Yahya, A. (2020). Indonesian Education Landscape and the 21st Century Challenges. *Journal of Social Studies Education Research*, 11(4), 219-234. <https://www.learntechlib.org/p/218538/>.
- Tan, S. H., & Pang, J. S. (2023). Test anxiety: An integration of the test anxiety and achievement motivation research traditions. *Educational Psychology Review*, 35(1), 13. <https://doi.org/10.1007/s10648-023-09737-1>
- Tiwari, S., Shah, B., & Muthiah, A. (2024). A global overview of SVA – Spatial-Visual ability. *Applied System Innovation*, 7(3), 48. <https://doi.org/10.3390/asi7030048>
- Trantham, P. S., Sikorski, J., De Ayala, R. J., & Doll, B. (2022). An item response theory and Rasch analysis of the NUDKS: a data literacy scale. *Educational Assessment, Evaluation and Accountability*, 34(1), 113-135. <https://doi.org/10.1007/s11092-021-09372-w>
- Von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of affective disorders*, 227, 483-493. <https://doi.org/10.1016/j.jad.2017.11.048>
- Wang, X.S., Perry, L.B., Malpique, A. et al. Factors predicting mathematics achievement in PISA: a systematic review. *Large-scale Assess Educ* 11, 24 (2023). <https://doi.org/10.1186/s40536-023-00174-8>
- Weiss, D. J., & Sahin, A. (2024). *Computerized adaptive testing: From concept to implementation*. Guilford Publications.

- Wilson, M. (2023). *Constructing measures: An item response modeling approach*. <https://doi.org/10.4324/9781003286929>
- Wu, Z. (2024). *Developing a Digital Solution Framework through a Digital Community of Practice System to Support Teachers' Implementation of STEAM* (Doctoral dissertation, RMIT University). <https://doi.org/10.25439/rmt.26046901>
- Zamjani, I. (2022). *The politics of educational decentralisation in Indonesia. A Quest for Legitimacy*. Springer Nature. <https://link.springer.com/book/10.1007/978-981-16-6901-9>