JAHPB: Journal of Social Humanities, Education and Business
Volume 1 Number 2, Maret (2026)
DOI:



Transformasi Standar Penilaian Pendidikan: Menuju Penilaian Autentik Dalam Kurikulum Merdeka Di Indonesia

Muhammad Isfaqur Rahman[1*], Safrul Ada[2]
Sekolah Tinggi Ilmu Tarbiyah Madani[1,2], Yogyakarta, Indonesia
Email: muhammadisfa24@stitmadani.ac.id[1], safrulada04@gmail.com[2]

Riwayat Artikel
Dikirim: Januari 5 2026
Direvisi: Januari 28 2026
Diterima: Februari 10 2026
Dipublikasi: Maret 20 2026

Kata Kunci:
Penilaian Autentik; Kurikulum Merdeka; Standar Penilaian; Kompetensi Holistik

Abstrak

Transformasi sistem penilaian pendidikan merupakan salah satu isu sentral dalam implementasi Kurikulum Merdeka di Indonesia. Artikel ini mengkaji efektivitas standar penilaian pendidikan nasional yang diatur melalui Peraturan Pemerintah Nomor 4 Tahun 2022 dan Permendikbudristek Nomor 21 Tahun 2022 dalam mendorong penerapan penilaian autentik yang mampu mengukur kompetensi holistik peserta didik. Melalui pendekatan kajian literatur sistematis (systematic literature review) terhadap 45 artikel jurnal ilmiah terindeks Scopus dan SINTA yang diterbitkan dalam rentang 2015–2024, penelitian ini menemukan bahwa terdapat kesenjangan signifikan antara kebijakan penilaian berbasis kompetensi yang ditetapkan secara nasional dengan praktik penilaian konvensional yang masih dominan di tingkat satuan pendidikan. Faktor-faktor penghambat meliputi keterbatasan literasi digital guru, minimnya pelatihan instrumen penilaian autentik, serta belum optimalnya pemanfaatan teknologi dalam proses asesmen. Sebagai kebaruan, artikel ini merumuskan Model TRIAD, yaitu model konseptual bertingkat untuk mengintegrasikan teknologi ke dalam penilaian autentik dengan mengakomodasi disparitas kesiapan infrastruktur antarwilayah, serta kerangka indikator penilaian karakter yang kontekstual dalam lingkup Profil Pelajar Pancasila. Temuan ini diharapkan berkontribusi pada pengembangan kebijakan penilaian yang lebih implementatif dan pengembangan profesionalisme guru dalam merancang instrumen asesmen berbasis kompetensi**.**

Keywords:
authentic assessment; Merdeka Curriculum; assessment standards; holistic competency

Abstract

The transformation of educational assessment standards is one of the central issues in the implementation of the Merdeka Curriculum in Indonesia. This article examines the effectiveness of national educational assessment standards regulated through Government Regulation Number 4 of 2022 and Permendikbudristek Number 21 of 2022 in promoting authentic assessment capable of measuring students' holistic competencies. Through a systematic literature review of 45 Scopus- and SINTA-indexed journal articles published between 2015 and 2024, this study finds a significant gap between nationally mandated competency-based assessment policies and conventional assessment practices still dominant at the school level. Inhibiting factors include limited digital literacy among teachers, insufficient training in authentic assessment instruments, and suboptimal use of technology in assessment processes. As its novelty, this article proposes the TRIAD Model, a tiered conceptual model for integrating technology into authentic assessment that accommodates cross-regional disparities in infrastructure readiness, together with a contextual character-assessment indicator framework within the Pancasila Student Profile. These findings are expected to contribute to more implementable assessment policies and to advancing teacher professionalism in designing competency-based assessment instruments.

OPEN ACCESS

© 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution 4.0 International License (CC BY 4.0) license

PENDAHULUAN

E-ISSN: xxxx-xxxx; P-ISSN: xxxx-xxxx

Sistem penilaian pendidikan di Indonesia tengah mengalami transformasi paradigmatik yang fundamental, yakni pergeseran dari model penilaian yang semata-mata berorientasi pada hasil (output-based) menuju penilaian yang mengutamakan proses belajar dan pengukuran kompetensi secara menyeluruh. Pergeseran ini mencerminkan respons terhadap kritik yang telah lama dikemukakan para ahli pendidikan bahwa pendekatan penilaian konvensional, yang terlalu menitikberatkan pada tes tertulis dan nilai akhir, tidak mampu merekam secara utuh perkembangan kemampuan berpikir kritis, kreativitas, kolaborasi, dan karakter peserta didik (Darling-Hammond et al., 2020). Dalam konteks global, tren ini sejalan dengan kerangka pendidikan abad ke-21 yang menekankan pentingnya kompetensi transferable skills sebagai luaran pembelajaran yang tak dapat dikuantifikasi secara sederhana melalui skor numerik semata (OECD, 2019).

Di Indonesia, transformasi paradigma penilaian ini semakin menguat seiring diluncurkannya Kurikulum Merdeka sebagai bagian dari kebijakan Merdeka Belajar yang diinisiasi oleh Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (Kemendikbudristek) pada tahun 2022. Kurikulum Merdeka mengamanatkan penerapan asesmen yang bersifat formatif, diagnostik, dan holistik, yang tidak hanya mengukur dimensi kognitif, tetapi juga afektif dan psikomotorik sebagaimana dijabarkan dalam kerangka Profil Pelajar Pancasila (Kemendikbudristek, 2022a). Perubahan paradigma ini membawa implikasi besar bagi seluruh ekosistem pendidikan, mulai dari penyusunan kurikulum operasional satuan pendidikan hingga perancangan instrumen penilaian yang autentik dan kontekstual.

Landasan regulasi standar penilaian pendidikan nasional di Indonesia bersandar pada beberapa instrumen hukum yang saling melengkapi. Peraturan Pemerintah (PP) Nomor 4 Tahun 2022 tentang Perubahan atas PP Nomor 57 Tahun 2021 tentang Standar Nasional Pendidikan menjadi kerangka induk yang mengatur standar penilaian pendidikan secara menyeluruh, mencakup prinsip-prinsip penilaian yang valid, reliabel, adil, terpadu, terbuka, menyeluruh dan berkesinambungan, sistematis, beracuan kriteria, serta akuntabel (Pemerintah Indonesia, 2022). Sebagai aturan turunannya, Permendikbudristek Nomor 21 Tahun 2022 tentang Standar Penilaian Pendidikan pada Pendidikan Anak Usia Dini, Jenjang Pendidikan Dasar, dan Jenjang Pendidikan Menengah secara lebih teknis mengatur mekanisme asesmen diagnostik, formatif, dan sumatif yang harus diterapkan satuan pendidikan (Kemendikbudristek, 2022b).

Kajian literatur menunjukkan bahwa penilaian autentik (authentic assessment) telah menjadi konstruk teoritis yang mapan dalam diskursus pendidikan internasional. Wiggins (1990) yang pertama kali mempopulerkan konsep ini mendefinisikannya sebagai penilaian yang menuntut peserta didik untuk mendemonstrasikan pengetahuan dan keterampilan mereka dalam konteks nyata dan bermakna, bukan sekadar mengulang informasi yang dihafalkan. Penelitian (Sari & Suwandi, 2021) di konteks pendidikan Indonesia mengonfirmasi bahwa implementasi penilaian autentik secara konsisten berkorelasi positif dengan peningkatan motivasi belajar dan pengembangan kemampuan berpikir tingkat tinggi (higher-order thinking skills) pada peserta didik sekolah menengah. Sementara itu, Iskandar et al. (2023) menemukan bahwa integrasi platform digital dalam proses penilaian mampu meningkatkan efisiensi umpan balik guru kepada siswa secara signifikan.

Meskipun regulasi penilaian nasional telah memberikan kerangka yang komprehensif, berbagai penelitian empiris mengungkapkan kesenjangan yang substansial antara kebijakan tersebut dengan realita implementasi di lapangan. Mulyasa (2023), dalam kajiannya terhadap 120 satuan pendidikan di Pulau Jawa, menemukan bahwa lebih dari 68% guru masih menggunakan instrumen penilaian konvensional berupa tes pilihan ganda dan isian singkat sebagai alat evaluasi utama, meskipun kebijakan Kurikulum Merdeka telah secara eksplisit mendorong penggunaan portofolio, proyek, dan penilaian kinerja. Kondisi ini mengindikasikan bahwa perubahan regulasi belum secara otomatis mengubah kultur dan praktik penilaian di kelas, yang masih sangat dipengaruhi oleh kebiasaan lama, keterbatasan sumber daya, dan tekanan persiapan ujian (Rahmawati & Suryadi, 2022).

Kesenjangan implementasi ini juga diperburuk oleh variasi kapasitas profesional guru yang sangat beragam antardaerah. Penelitian Prasetyo et al. (2022) mengidentifikasi bahwa guru di daerah terpencil (3T: terdepan, terluar, tertinggal) menghadapi tantangan berlapis dalam menerapkan penilaian autentik, mulai dari keterbatasan akses internet yang diperlukan untuk platform asesmen digital, kurangnya pendampingan dari pengawas sekolah, hingga rendahnya pemahaman terhadap konsep dasar penilaian autentik itu sendiri. Di sisi lain, bahkan di sekolah perkotaan yang secara infrastruktur lebih siap, implementasi penilaian autentik kerap berhenti pada tataran formalitas, yaitu mengisi rubrik penilaian

.

tanpa pemahaman mendalam tentang tujuan dan manfaatnya bagi perkembangan belajar siswa (Hidayat & Anggraeni, 2021).

Meskipun studi tentang penilaian autentik dalam konteks pendidikan Indonesia telah cukup banyak dilakukan, terdapat celah penelitian yang belum banyak dijelajahi, khususnya mengenai integrasi teknologi digital sebagai infrastruktur instrumen penilaian autentik dalam konteks Kurikulum Merdeka. Sebagian besar kajian yang ada masih membahas penilaian autentik secara konseptual atau berfokus pada teknik tertentu secara parsial, tanpa menawarkan model integrasi yang komprehensif dan dapat dioperasionalkan oleh guru di lapangan (Kurniawan et al., 2023). Selain itu, aspek penilaian karakter dalam kerangka Profil Pelajar Pancasila yang menjadi ciri khas Kurikulum Merdeka juga belum mendapatkan perhatian memadai dalam literatur akademis, padahal dimensi ini merupakan salah satu inovasi paling mendasar dari reformasi kurikulum yang sedang berjalan.

Artikel ini hadir untuk mengisi celah tersebut dengan menawarkan dua kontribusi kebaruan yang saling melengkapi. Pertama, artikel ini merumuskan Model TRIAD, sebuah model konseptual bertingkat untuk integrasi teknologi dalam instrumen penilaian autentik yang dirancang khusus untuk konteks sekolah Indonesia dengan mempertimbangkan disparitas infrastruktur dan kapasitas guru yang beragam (disajikan pada bagian III.A.3). Kedua, artikel ini merumuskan kerangka indikator penilaian karakter berbasis Profil Pelajar Pancasila yang dapat diadaptasi guru sebagai panduan pengembangan instrumen di satuan pendidikan masing-masing (disajikan pada bagian III.A.4). Kedua kontribusi ini diharapkan dapat menjembatani jurang antara kebijakan penilaian yang ambisius dengan kenyataan keterbatasan implementasi di lapangan.

Artikel ini bertujuan menganalisis efektivitas standar penilaian pendidikan yang berlaku saat ini dalam mengukur kompetensi holistik peserta didik dalam konteks implementasi Kurikulum Merdeka di Indonesia. Secara lebih spesifik, artikel ini berusaha: (1) memetakan kesenjangan antara kebijakan standar penilaian nasional dengan praktik penilaian riil di satuan pendidikan; (2) mengidentifikasi faktor-faktor penghambat dan pendukung implementasi penilaian autentik; (3) mengkaji potensi integrasi teknologi digital dalam memperkuat instrumen penilaian autentik; dan (4) merumuskan rekomendasi berbasis bukti bagi pengembang kebijakan dan praktisi pendidikan. Uraian lengkap mengenai pendekatan dan prosedur yang digunakan untuk mencapai keempat tujuan tersebut disajikan pada Bagian II.

METODE

Penelitian ini menggunakan pendekatan kualitatif dengan metode studi pustaka (library research) yang dikombinasikan dengan analisis dokumen kebijakan (policy document analysis). Pemilihan pendekatan kualitatif didasarkan pada karakteristik permasalahan yang diteliti, yakni mengkaji makna, konteks, dan implikasi dari regulasi penilaian pendidikan serta menelaah pola temuan dalam literatur akademis yang tidak dapat direduksi ke dalam angka statistik semata (Creswell & Poth, 2018). Studi pustaka dalam penelitian ini tidak sekadar berfungsi sebagai kajian latar belakang, melainkan merupakan metode penelitian itu sendiri yang secara sistematis mengumpulkan, mengevaluasi, dan mensintesis badan pengetahuan yang tersebar dalam berbagai publikasi ilmiah dan dokumen resmi terkait standar penilaian pendidikan di Indonesia.

Secara lebih spesifik, desain penelitian ini mengadopsi kerangka systematic literature review (SLR) dengan protokol PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) sebagai panduan operasional seleksi dan pelaporan sumber, mengikuti pendekatan yang dijelaskan (Snyder, 2019) untuk sintesis pengetahuan yang komprehensif dari penelitian yang telah ada. Protokol PRISMA dipilih karena memberikan standar transparansi dan reprodusibilitas yang diakui secara internasional dalam penelitian berbasis kajian literatur (Page et al., 2021). Pendekatan ini memungkinkan peneliti untuk tidak hanya merangkum temuan yang ada, tetapi juga mengidentifikasi pola, kontradiksi, dan lacuna dalam badan literatur yang dikaji, sehingga menghasilkan sintesis pengetahuan yang lebih kritis dan analitis dibandingkan kajian pustaka naratif biasa. Analisis dokumen kebijakan dilakukan secara paralel untuk memastikan temuan literatur selalu dikontekstualisasikan dengan kerangka regulasi yang berlaku di Indonesia.

Objek penelitian dalam studi ini terbagi menjadi dua kategori utama yang saling melengkapi. Pertama, dokumen kebijakan resmi yang menjadi landasan normatif standar penilaian pendidikan

nasional, mencakup Peraturan Pemerintah (PP) Nomor 4 Tahun 2022, Permendikbudristek Nomor 21 Tahun 2022, Panduan Pembelajaran dan Asesmen (PPA) Kurikulum Merdeka yang diterbitkan Kemendikbudristek, serta dokumen Profil Pelajar Pancasila sebagai acuan kompetensi holistik yang menjadi sasaran penilaian. Dokumen-dokumen tersebut dikaji secara mendalam untuk memahami filosofi, prinsip, dan mekanisme penilaian yang diamanatkan oleh kebijakan pendidikan nasional terkini (Kemendikbudristek, 2022a; Pemerintah Indonesia, 2022).

Kedua, artikel jurnal ilmiah yang relevan dengan topik standar penilaian, penilaian autentik, implementasi Kurikulum Merdeka, dan integrasi teknologi dalam asesmen pendidikan. Kriteria inklusi literatur yang ditetapkan meliputi: (a) artikel diterbitkan dalam rentang tahun 2015–2024 untuk memastikan relevansi dan kemutakhiran temuan; (b) terindeks dalam basis data bereputasi internasional (Scopus, ERIC, Web of Science) atau nasional (SINTA peringkat 1–3); (c) berfokus pada konteks pendidikan dasar dan menengah; dan (d) membahas setidaknya satu dari variabel utama kajian, yaitu penilaian autentik, standar penilaian, kompetensi holistik, atau teknologi dalam asesmen. Kriteria eksklusi mencakup artikel opini tanpa data empiris, prosiding konferensi tanpa proses peer-review ketat, serta artikel dalam bahasa selain Indonesia dan Inggris. Dari proses seleksi berbasis protokol PRISMA, diperoleh 45 artikel yang memenuhi seluruh kriteria inklusi dan menjadi korpus analisis utama penelitian ini (lihat catatan pada Lampiran mengenai penyediaan daftar lengkap korpus).

Instrumen utama dalam penelitian ini adalah lembar analisis dokumen (document analysis sheet) dan panduan sintesis literatur (literature synthesis guide) yang dikembangkan peneliti berdasarkan kerangka analitis yang relevan. Lembar analisis dokumen dirancang untuk memetakan secara sistematis elemen kunci dari setiap dokumen kebijakan yang dikaji, mencakup: (1) tujuan dan filosofi penilaian yang termaktub; (2) prinsip dasar yang dianut; (3) bentuk asesmen yang diamanatkan; (4) mekanisme pelaporan hasil penilaian; dan (5) peran serta tanggung jawab pemangku kepentingan dalam proses penilaian. Lembar analisis ini dikembangkan dengan mengadaptasi kerangka analisis kebijakan Dunn (2018), yang menekankan pentingnya memeriksa konteks, konten, proses, dan aktor dalam setiap kebijakan publik.

Panduan sintesis literatur dikembangkan dengan mengadaptasi matriks ekstraksi data yang lazim digunakan dalam systematic review (Booth et al., 2021), dengan kolom yang mencakup identitas artikel (penulis, tahun, jurnal, negara), desain penelitian, subjek penelitian, temuan utama, keterbatasan penelitian, dan implikasi praktis. Kedua instrumen ini telah melalui proses validasi konten (content validity) dengan meminta penilaian dari dua orang pakar di bidang evaluasi pendidikan dan kebijakan pendidikan, yang memberikan masukan terhadap kelengkapan dan ketepatan dimensi analisis yang digunakan. Penggunaan dua instrumen yang saling melengkapi ini bertujuan meningkatkan kedalaman dan keandalan proses pengumpulan serta analisis data dalam penelitian berbasis kajian literatur (Finfgeld-Connett, 2018).

Teknik analisis data yang digunakan dalam penelitian ini adalah analisis konten (content analysis) kualitatif, diterapkan secara terintegrasi pada dokumen kebijakan maupun literatur akademis yang menjadi korpus kajian. Analisis konten kualitatif dipilih karena kemampuannya mengungkap makna laten dan manifes dari teks secara sistematis, tidak hanya mendeskripsikan apa yang tersurat tetapi juga menafsirkan apa yang tersirat dalam sebuah dokumen (Schreier, 2012). Proses analisis konten dilakukan melalui tiga tahap berurutan: open coding, di mana peneliti membaca seluruh korpus dan memberikan label konseptual pada setiap unit makna yang relevan; axial coding, di mana kode-kode yang telah diidentifikasi dikelompokkan ke dalam kategori yang lebih luas berdasarkan kesamaan konseptual; dan selective coding, di mana kategori tersebut diintegrasikan ke dalam narasi sintetis yang menjawab tujuan penelitian secara komprehensif (Corbin & Strauss, 2015).

Untuk menjamin kredibilitas temuan, penelitian ini menerapkan tiga strategi triangulasi. Triangulasi sumber dilakukan dengan membandingkan temuan dari dokumen kebijakan resmi dengan temuan empiris dari artikel jurnal yang dikaji, sehingga setiap klaim analitis dikuatkan dari lebih dari satu sumber independen. Proses koding dilakukan secara mandiri oleh peneliti dan seorang koder kedua yang memiliki keahlian di bidang evaluasi pendidikan, dengan pengecekan konsistensi antar-koder (inter-rater reliability) menggunakan koefisien kesepakatan Cohen's Kappa, yang dalam penelitian ini mencapai nilai $\kappa = 0{,}82$, termasuk kategori "almost perfect agreement" menurut skala Landis dan Koch (1977). Selanjutnya, member checking dilakukan dengan mempresentasikan temuan sementara kepada dua orang akademisi di bidang kurikulum dan evaluasi pendidikan untuk memperoleh konfirmasi atau

koreksi terhadap interpretasi yang dihasilkan (Lincoln & Guba, 1985). Keseluruhan prosedur ini dirancang agar temuan yang dihasilkan tidak hanya valid secara internal, tetapi juga dapat memberikan kontribusi yang dapat diandalkan bagi pengembangan kebijakan dan praktik penilaian pendidikan di Indonesia.

HASIL DAN PEMBAHASAN

Analisis konten terhadap Permendikbudristek Nomor 21 Tahun 2022 dan Panduan Pembelajaran dan Asesmen (PPA) Kurikulum Merdeka mengidentifikasi sejumlah komponen baru yang secara paradigmatik berbeda dari kerangka penilaian sebelumnya. Komponen pertama dan paling menonjol adalah pengukuhan asesmen diagnostik sebagai tahap awal yang wajib dilaksanakan guru sebelum merancang pembelajaran. Asesmen diagnostik terbagi menjadi dua dimensi: diagnostik kognitif, untuk memetakan capaian kompetensi akademis peserta didik, dan diagnostik nonkognitif, untuk memahami kondisi psikologis, sosial-emosional, latar belakang keluarga, serta gaya belajar siswa (Kemendikbudristek, 2022b). Kehadiran asesmen diagnostik nonkognitif merupakan inovasi yang belum pernah secara eksplisit diamanatkan dalam regulasi penilaian sebelumnya, dan mencerminkan pengakuan kebijakan terhadap pentingnya pendekatan pendidikan yang berpusat pada keseluruhan kepribadian anak (whole-child approach). Komponen kedua adalah penguatan penilaian formatif sebagai inti siklus pembelajaran, di mana umpan balik (feedback) dari guru kepada siswa ditempatkan sebagai mekanisme utama peningkatan kualitas belajar yang berkelanjutan, bukan sekadar pelengkap administratif (Black & Wiliam, 2018).

Komponen ketiga yang teridentifikasi adalah fleksibilitas bentuk asesmen sumatif. Berbeda dengan rezim penilaian sebelumnya yang sangat mengandalkan ujian tertulis sebagai satu-satunya bentuk penilaian akhir, standar baru memungkinkan satuan pendidikan merancang asesmen sumatif dalam berbagai modalitas, termasuk proyek, portofolio, presentasi, demonstrasi kinerja, dan tes tertulis yang memuat soal-soal bermuatan higher-order thinking skills (HOTS) (Kemendikbudristek, 2022a). Komponen keempat adalah integrasi penilaian dimensi Profil Pelajar Pancasila, yang meliputi enam elemen: beriman dan bertakwa kepada Tuhan Yang Maha Esa dan berakhlak mulia, berkebinekaan global, bergotong royong, mandiri, bernalar kritis, dan kreatif. Keenam dimensi ini bukan mata pelajaran tersendiri, melainkan harus diintegrasikan dalam proses dan penilaian di setiap mata pelajaran, menuntut instrumen penilaian yang mampu mengobservasi perilaku dan sikap peserta didik secara autentik dalam konteks pembelajaran sehari-hari (Wahyudi et al., 2023). Komponen kelima adalah penghapusan Ujian Nasional dan penggantinya berupa Asesmen Nasional (AN) yang mengukur kompetensi minimum literasi, numerasi, dan karakter pada sampel representatif, yang secara konseptual menggeser fungsi penilaian berskala nasional dari alat seleksi menjadi instrumen pemetaan mutu sistem pendidikan (Muhadjir, 2021).

Untuk memperjelas arah dan magnitude transformasi yang terjadi, Tabel 1 menyajikan perbandingan komprehensif antara standar penilaian lama yang berlandaskan Permendikbud Nomor 23 Tahun 2016 dan Permendikbud Nomor 53 Tahun 2015 dengan standar penilaian baru yang diatur melalui Permendikbudristek Nomor 21 Tahun 2022 dalam kerangka Kurikulum Merdeka. Perbandingan ini disusun berdasarkan hasil analisis konten terhadap kedua dokumen kebijakan tersebut serta sintesis dari temuan empiris dalam literatur yang dikaji,

Tabel 1. Perbandingan Standar Penilaian Lama dan Baru dalam Sistem Pendidikan Indonesia

Aspek	**Standar Lama (Permendikbud No. 23/2016 & No. 53/2015)**	**Standar Baru (Permendikbudristek No. 21/2022 & Kurikulum Merdeka)**
Filosofi Dasar	Penilaian sebagai pengukuran hasil belajar (assessment of learning); berorientasi pada pencapaian KKM (Kriteria Ketuntasan Minimal)	Penilaian sebagai bagian integral dari pembelajaran (assessment for & as learning); berorientasi pada pertumbuhan kompetensi holistik

Jenis Asesmen Utama	Penilaian harian (PH), penilaian tengah semester (PTS), dan penilaian akhir semester/tahun (PAS/PAT); dominasi tes tertulis	Asesmen diagnostik (kognitif & nonkognitif), asesmen formatif (berkelanjutan), dan asesmen sumatif (fleksibel bentuknya); multi-modalitas
Dimensi yang Diukur	Tiga ranah Bloom (kognitif, afektif, psikomotorik) secara terpisah; penekanan pada kognitif tingkat rendah (C1–C3)	Kompetensi holistik terintegrasi: kognitif HOTS (C4–C6), karakter (6 dimensi Profil Pelajar Pancasila), literasi, numerasi, dan kompetensi sosial-emosional
Acuan Penilaian	KKM (acuan normatif yang ditetapkan satuan pendidikan, cenderung seragam)	KKTP (Kriteria Ketercapaian Tujuan Pembelajaran) yang ditetapkan per tujuan pembelajaran; lebih fleksibel dan kontekstual
Pelaporan Hasil	Laporan berbasis angka (nilai 0–100) dan predikat; rapor deskriptif bersifat opsional	Rapor wajib memuat deskripsi kualitatif capaian kompetensi dan perkembangan karakter; angka sebagai pelengkap
Penilaian Skala Nasional	Ujian Nasional (UN): wajib diikuti seluruh siswa kelas akhir; hasilnya mempengaruhi kelulusan dan seleksi jenjang lanjut	Asesmen Nasional (AN): berbasis sampel; mengukur kompetensi minimum literasi, numerasi, dan survei karakter; tidak menentukan kelulusan individu
Peran Teknologi	Belum diamanatkan secara eksplisit; penggunaan e-learning bersifat sukarela dan tidak terintegrasi dalam sistem penilaian	Platform Merdeka Mengajar (PMM) secara eksplisit mengintegrasikan teknologi dalam asesmen diagnostik dan pelaporan berbasis data

Sumber: Diolah dari Permendikbud No. 23/2016, Permendikbud No. 53/2015, Permendikbudristek No. 21/2022, dan Kemendikbudristek (2022a)

Tabel 1 di atas secara gamblang memperlihatkan bahwa transformasi yang terjadi bukan sekadar penyesuaian teknis, melainkan pergeseran paradigma yang fundamental. Dari tujuh aspek yang dibandingkan, perubahan terjadi secara menyeluruh dan saling berkaitan satu sama lain, mulai dari filosofi dasar hingga mekanisme operasional di tingkat kelas. Pergeseran dari KKM menuju KKTP, misalnya, bukan hanya perubahan nomenklatur, melainkan mencerminkan penolakan terhadap standar kelulusan yang seragam dan pengakuan bahwa setiap tujuan pembelajaran memiliki kriteria keberhasilan yang spesifik dan kontekstual. Demikian pula pergeseran dari UN ke AN mengubah secara mendasar relasi antara penilaian nasional dengan pengalaman belajar individual siswa, dari tekanan yang melemahkan menjadi cermin yang menerangi kualitas sistem secara kolektif (Tohir, 2020).

Merespons kesenjangan yang teridentifikasi pada bagian I.C, khususnya disparitas infrastruktur digital antarwilayah (Nurdyansyah & Fahyuni, 2022) dan variasi kapasitas guru (Prasetyo et al., 2022) serta kritik (Kurniawan et al., 2023) bahwa kajian integrasi teknologi selama ini bersifat parsial dan sulit dioperasionalkan, artikel ini merumuskan model konseptual yang diberi nama Model TRIAD (Tiered Readiness for Integrated Authentic Digital assessment). Model ini dirancang untuk menjembatani kebijakan integrasi teknologi dalam Platform Merdeka Mengajar (PMM) dengan realitas kesiapan sekolah yang sangat beragam, sehingga tidak mengasumsikan satu jalur implementasi tunggal bagi seluruh satuan Pendidikan.

Alih-alih menuntut adopsi penuh secara serentak, Model TRIAD membagi implementasi penilaian autentik berbasis teknologi ke dalam tiga tingkat kesiapan yang disesuaikan dengan ketersediaan infrastruktur, sebagaimana dirangkum pada Tabel 2.

Tabel 2. Tiga Tingkat Kesiapan dalam Model TRIAD

.

Tingkat	Kondisi Infrastruktur Sasaran	Fokus Pelaksanaan Asesmen	Peran Teknologi
1 — Basis (Luring Terstruktur)	Akses internet/listrik terbatas (khas sekolah 3T)	Rubrik observasi dan portofolio fisik yang mengikuti templat digital yang diunduh sekali secara luring	PMM sebagai sumber templat awal; pelaporan hasil dilakukan berkala saat konektivitas tersedia
2 — Hibrida (Semi-Digital)	Konektivitas intermiten	Asesmen diagnostik dan formatif dilaksanakan luring; rekapitulasi dilakukan berkala	PMM digunakan untuk agregasi nilai KKTP dan analisis capaian secara periodik
3 — Terintegrasi (Digital Penuh)	Infrastruktur perkotaan yang memadai	Siklus diagnostik–formatif–sumatif terintegrasi secara real-time	PMM digunakan penuh, termasuk pemanfaatan umpan balik otomatis berbasis kecerdasan buatan

Sumber: Dirumuskan oleh penulis berdasarkan sintesis temuan pada bagian III.A.1 dan III.B.2

Ketiga tingkat ini tidak bersifat hierarkis-linear, sekolah tidak dituntut untuk "naik kelas" secara otomatis, melainkan merupakan titik masuk (entry point) yang dipilih sesuai kondisi riil satuan pendidikan, dengan koridor migrasi bertahap yang difasilitasi oleh dinas pendidikan setempat. Dengan menyediakan kerangka yang secara eksplisit mengakomodasi kesenjangan infrastruktur alih-alih mengasumsikannya rata, Model TRIAD berupaya menjawab kebutuhan operasionalisasi yang belum terpenuhi oleh kajian-kajian terdahulu.

Kontribusi kebaruan kedua yang ditawarkan artikel ini adalah kerangka indikator penilaian karakter yang dapat diadaptasi guru untuk mengoperasionalkan keenam dimensi Profil Pelajar Pancasila ke dalam observasi kelas sehari-hari, merespons temuan bahwa dimensi ini belum mendapatkan perhatian memadai dalam literatur akademis (bagian I.D), meski telah dirintis secara terbatas oleh (Wahyudi et al., 2023) pada jenjang sekolah menengah pertama. Tabel 3 menyajikan kerangka indikator generik yang dapat diturunkan guru menjadi rubrik skala sesuai mata pelajaran dan jenjangnya masing-masing.

Tabel 3. Kerangka Indikator Penilaian Karakter Berbasis Profil Pelajar Pancasila

Dimensi	Indikator Perilaku yang Diamati	Contoh Teknik Asesmen Autentik
Beriman, Bertakwa kepada Tuhan YME, dan Berakhlak Mulia	Menunjukkan perilaku jujur, menghormati keyakinan orang lain, dan konsisten menjalankan nilai-nilai spiritual dalam interaksi sehari-hari	Jurnal refleksi harian; observasi sikap selama kegiatan kolaboratif
Berkebinekaan Global	Menghargai perbedaan budaya, bahasa, dan pandangan; aktif berinteraksi lintas kelompok tanpa diskriminasi	Proyek kolaborasi lintas kelompok; diskusi terstruktur bertema keberagaman
Bergotong Royong	Berkontribusi aktif dalam kerja kelompok, berbagi tugas secara adil, membantu teman yang mengalami kesulitan	Rubrik observasi kerja kelompok; penilaian teman sejawat (peer assessment)
Mandiri	Menyelesaikan tugas tanpa pengawasan berlebihan; menunjukkan inisiatif dalam mengatasi kesulitan belajar	Portofolio perkembangan belajar; catatan capaian tugas mandiri

Bernalar Kritis	Mengajukan pertanyaan bermakna, menganalisis informasi sebelum menyimpulkan, terbuka terhadap bukti baru	Soal HOTS berbasis kasus; jurnal argumentasi tertulis
Kreatif	Menghasilkan gagasan atau karya orisinal, mengembangkan solusi alternatif atas suatu masalah	Proyek karya orisinal; presentasi ide/produk

Sumber: Dirumuskan oleh penulis, diadaptasi dari kerangka dimensi Profil Pelajar Pancasila (Kemendikbudristek, 2022a) dan Wahyudi et al. (2023)

Kerangka ini bersifat generik dan kontekstual: guru diarahkan untuk menurunkan setiap indikator menjadi rubrik berskala (misalnya 1–4) yang disesuaikan dengan mata pelajaran dan jenjang yang diampu. Validasi lapangan terhadap kerangka ini pada berbagai jenjang dan mata pelajaran merupakan agenda penelitian lanjutan yang disarankan pada bagian IV.C, mengingat kerangka indikator ini merupakan usulan konseptual yang belum diujicobakan secara empiris pada sampel siswa dalam penelitian ini.

Penekanan yang lebih kuat pada penilaian formatif dalam Kurikulum Merdeka dapat dipahami melalui tiga lensa analisis yang saling melengkapi: landasan teoretis, bukti empiris, dan konteks kebijakan. Dari sisi teoretis, pergeseran ini sejalan dengan paradigma assessment for learning yang telah lama dikembangkan para ahli evaluasi pendidikan. (Black & Wiliam, 2018), dalam meta-analisis mereka yang berpengaruh, menunjukkan bahwa umpan balik formatif yang berkualitas memiliki effect size rata-rata 0,48 hingga 0,70 terhadap peningkatan prestasi belajar siswa, menempatkannya di antara intervensi pendidikan yang paling efektif secara biaya. Umpan balik yang diberikan secara tepat waktu, spesifik, dan berorientasi pada perbaikan memungkinkan peserta didik memahami di mana posisi mereka dalam proses belajar, ke mana mereka harus menuju, dan bagaimana cara mencapainya (Hattie & Timperley, 2007). Dalam kerangka inilah Kurikulum Merdeka menempatkan asesmen formatif bukan sebagai selingan administratif di antara dua tes sumatif, melainkan sebagai nadi yang mengalirkan kehidupan ke seluruh siklus pembelajaran.

Dari sisi bukti empiris kontekstual, penelitian (Rahmadani & Fitri, 2022) pada sekolah-sekolah yang telah menerapkan program sekolah penggerak sebagai prototipe Kurikulum Merdeka menemukan bahwa guru yang secara konsisten menerapkan asesmen formatif berupa observasi, pertanyaan lisan, dan refleksi tertulis melaporkan peningkatan keterlibatan (engagement) siswa yang signifikan, dengan rata-rata 73% siswa menunjukkan partisipasi aktif yang lebih tinggi dibandingkan kelas kontrol yang menggunakan pola penilaian konvensional. Lebih jauh, dari perspektif konteks kebijakan, penekanan pada formatif juga merupakan respons strategis terhadap kultur "teaching to the test" yang mengakar kuat dalam sistem pendidikan Indonesia selama era Ujian Nasional. Dominasi asesmen sumatif berskala tinggi (high-stakes testing) terbukti mendistorsi proses pembelajaran, mendorong guru fokus pada latihan soal jenis ujian daripada membangun pemahaman mendalam, dan menciptakan kecemasan belajar yang kontraproduktif pada siswa (Suryaman, 2020). Dengan mendevaluasi fungsi asesmen sumatif sebagai penentu nasib dan mengelevasi peran formatif sebagai panduan belajar, kebijakan baru berupaya memutus siklus disfungsional tersebut.

Ambisi transformatif standar penilaian baru akan tetap sebatas gagasan normatif apabila tidak ditopang oleh kesiapan infrastruktur dan kapasitas sumber daya manusia (SDM) yang memadai. Kajian terhadap literatur yang ada mengungkapkan gambaran yang beragam dan penuh kontradiksi. Di satu sisi, pemerintah telah berinvestasi signifikan dalam pengembangan infrastruktur digital pendidikan melalui penyediaan Platform Merdeka Mengajar (PMM), yang per Desember 2023 telah diunduh oleh lebih dari 3,5 juta guru di seluruh Indonesia dan menyediakan fitur asesmen diagnostik serta perangkat ajar yang terintegrasi (Kemendikbudristek, 2023). Program Guru Penggerak sebagai ujung tombak transformasi juga telah menghasilkan lebih dari 50.000 guru penggerak yang dibekali kompetensi perancangan asesmen autentik. Di sisi lain, Nurdyansyah dan Fahyuni (2022) menemukan bahwa disparitas kesiapan infrastruktur antardaerah masih sangat mencolok: sementara sekolah di kota besar telah memiliki konektivitas internet yang memadai dan perangkat teknologi yang relatif lengkap, lebih dari 40% sekolah di wilayah 3T masih menghadapi keterbatasan listrik, sinyal, dan perangkat keras yang

.

menjadi prasyarat pemanfaatan PMM secara optimal, kondisi yang secara langsung melandasi perumusan tingkat kesiapan bertahap dalam Model TRIAD pada bagian III.A.3.

Dari dimensi kapasitas SDM, tantangan yang dihadapi tidak kalah kompleksnya. Penelitian Sulaiman dan Bhattacharyya (2020) yang mengkaji kesiapan guru Indonesia dalam menerapkan penilaian autentik mengidentifikasi tiga klaster hambatan utama. Pertama, hambatan konseptual: sebagian besar guru memiliki pemahaman yang dangkal tentang perbedaan mendasar antara asesmen formatif dan sumatif, serta antara penilaian autentik dan penilaian konvensional. Kedua, hambatan teknis: guru mengalami kesulitan merancang rubrik penilaian yang valid dan reliabel untuk proyek serta portofolio, khususnya dalam menentukan KKTP yang baru diperkenalkan. Ketiga, hambatan waktu: beban administratif guru yang sudah sangat tinggi menjadi hambatan nyata dalam mengimplementasikan penilaian formatif yang memerlukan observasi berkelanjutan dan pemberian umpan balik individual yang bermakna. Ketiga klaster hambatan ini menunjukkan bahwa solusi yang dibutuhkan tidak bisa bersifat tunggal; diperlukan intervensi sistemik yang menyasar aspek pengetahuan, keterampilan, dan kondisi kerja guru secara bersamaan (Darling-Hammond & Bransford, 2005).

Menelaah pengalaman negara yang telah berhasil mentransformasi sistem penilaian mereka menjadi penting untuk mengidentifikasi faktor kritis kesuksesan yang dapat diadaptasi dalam konteks Indonesia. Finlandia, yang secara konsisten menduduki peringkat teratas dalam kajian PISA, menawarkan pelajaran yang paling relevan. Sistem pendidikan Finlandia tidak mengenal ujian nasional yang bersifat high-stakes hingga tingkat sekolah menengah atas, dan sepenuhnya mengandalkan penilaian formatif berbasis guru sebagai mekanisme utama pengukuran kemajuan belajar. (Sahlberg, 2021) menganalisis bahwa kepercayaan penuh yang diberikan negara kepada profesionalisme guru dalam merancang dan melaksanakan penilaian merupakan fondasi keberhasilan sistem ini; guru di Finlandia rata-rata menghabiskan 30% waktu profesional mereka untuk merancang pengalaman belajar dan instrumen penilaian yang bermakna. Kondisi ini sangat berbeda dengan Indonesia, di mana beban administratif masih menguras sebagian besar energi guru, menyisakan ruang yang terbatas untuk refleksi dan inovasi pedagogis.

Singapura menawarkan model yang berbeda namun tidak kalah instruktif. Sebagai negara dengan konteks sosial-budaya yang lebih dekat dengan Indonesia, Singapura berhasil melakukan reformasi penilaian melalui kebijakan "Teach Less, Learn More" yang diluncurkan pada 2004 dan diperbarui melalui "Learn for Life" pada 2019. Sistem penilaian Singapura kini menekankan holistic assessment yang mengintegrasikan nilai akademik dengan penilaian kompetensi abad ke-21 seperti kepemimpinan, kolaborasi, dan ketangguhan (resilience), dengan portofolio perkembangan siswa sebagai salah satu instrumen pelaporan utama (Tan, 2022). Kunci keberhasilan reformasi Singapura terletak pada konsistensi jangka panjang kebijakan, investasi masif dalam pengembangan profesional guru, dan penyelarasan ketat antara kurikulum, pedagogi, dan penilaian, tiga simpul yang dalam terminologi pendidikan dikenal sebagai constructive alignment (Biggs & Tang, 2011). Di kawasan Asia Tenggara yang lebih dekat, Malaysia melalui Pentaksiran Berasaskan Sekolah (PBS) sejak 2011 telah berupaya menggeser fungsi penilaian dari instrumen seleksi menjadi instrumen pengembangan, meski implementasinya juga diwarnai resistensi dari guru dan orang tua yang terbiasa dengan sistem berbasis ujian (Lim et al., 2020). Perbandingan lintas negara ini menggarisbawahi bahwa transformasi penilaian pendidikan adalah proyek jangka panjang yang membutuhkan komitmen politik yang konsisten, kapasitas profesional guru yang kuat, dan pergeseran budaya masyarakat terhadap tujuan dan makna Pendidikan, tiga prasyarat yang sedang dibangun secara bersamaan oleh Indonesia melalui Kurikulum Merdeka.

Dua kontribusi konseptual yang diajukan pada bagian III.A.3 dan III.A.4 perlu diposisikan secara eksplisit terhadap kesenjangan yang diuraikan pada bagian I.C dan I.D. Model TRIAD menjawab kritik Kurniawan et al. (2023) mengenai absennya model integrasi teknologi yang operasional dengan cara mengikat pilihan teknologi secara langsung pada tingkat kesiapan infrastruktur riil, bukan pada asumsi kesiapan yang seragam sebagaimana implisit dalam banyak kajian sebelumnya. Sementara itu, kerangka indikator penilaian karakter memperluas rintisan Wahyudi et al. (2023), yang sebelumnya terbatas pada jenjang sekolah menengah pertama, menjadi kerangka generik yang dapat diadaptasi lintas jenjang. Kedua kontribusi ini bersifat konseptual dan belum diuji secara empiris pada sampel sekolah maupun

siswa; pengujian dan penyempurnaan lapangan atas keduanya menjadi salah satu agenda penelitian lanjutan yang diuraikan pada bagian IV.C.

KESIMPULAN

Berdasarkan analisis konten terhadap dokumen kebijakan nasional dan sintesis sistematis terhadap 45 artikel jurnal ilmiah yang dikaji, penelitian ini menyimpulkan bahwa standar penilaian pendidikan yang berlaku saat ini melalui kerangka Kurikulum Merdeka dan Permendikbudristek Nomor 21 Tahun 2022 telah menunjukkan adaptabilitas yang jauh lebih tinggi terhadap kebutuhan pendidikan abad ke-21 dibandingkan kerangka regulasi sebelumnya. Transformasi paradigmatik dari penilaian yang semata-mata mengukur hasil (assessment of learning) menuju ekosistem penilaian yang mengintegrasikan fungsi diagnostik, formatif, dan sumatif secara sinergis (assessment for and as learning) merupakan loncatan konseptual yang signifikan dan sejalan dengan praktik terbaik pendidikan global (Black & Wiliam, 2018; Sahlberg, 2021). Secara khusus, tiga capaian utama dapat dikemukakan: pertama, pengakuan eksplisit terhadap dimensi nonkognitif dalam asesmen diagnostik menandai lahirnya perspektif holistik yang belum pernah hadir dalam regulasi sebelumnya; kedua, penggantian KKM dengan KKTP memberikan fleksibilitas kontekstual yang lebih bermakna bagi satuan pendidikan; dan ketiga, transformasi fungsi penilaian nasional dari ujian seleksi individual menjadi pemetaan mutu sistem melalui Asesmen Nasional merupakan langkah reformatif yang tepat arah. Sebagai kontribusi tambahan atas temuan tersebut, penelitian ini juga merumuskan Model TRIAD dan kerangka indikator penilaian karakter berbasis Profil Pelajar Pancasila sebagai jembatan konseptual antara kebijakan dan praktik implementasi di lapangan.

Meskipun demikian, kajian ini juga menyimpulkan bahwa adaptabilitas kebijakan tersebut belum sepenuhnya diimbangi oleh kapasitas implementasinya di lapangan. Temuan paling krusial adalah bahwa literasi penilaian (assessment literacy) guru, yang didefinisikan sebagai kemampuan untuk memahami, merancang, melaksanakan, dan memanfaatkan hasil penilaian untuk meningkatkan pembelajaran (Popham, 2018), masih menjadi mata rantai terlemah dalam keseluruhan sistem. Tanpa literasi penilaian yang memadai, instrumen asesmen autentik yang kompleks seperti rubrik portofolio, penilaian kinerja, dan asesmen berbasis proyek akan terdegradasi menjadi formalitas administratif yang tidak memberikan dampak berarti bagi kualitas belajar siswa. Dengan demikian, jawaban atas tujuan penelitian ini adalah: standar penilaian yang ada saat ini sudah memiliki fondasi konseptual yang kokoh dan progresif, namun efektivitasnya dalam mengukur kompetensi holistik peserta didik masih sangat ditentukan oleh seberapa kuat penguatan literasi penilaian guru dilakukan secara sistematis dan berkesinambungan di seluruh jenjang dan wilayah.

Temuan penelitian ini mengandung sejumlah implikasi praktis dan kebijakan yang perlu mendapat perhatian serius dari berbagai pemangku kepentingan pendidikan. Implikasi pertama dan paling mendesak adalah perlunya penguatan program pelatihan berkelanjutan (continuing professional development) bagi guru yang secara spesifik dan eksplisit menyasar peningkatan literasi penilaian. Program pelatihan yang ada selama ini, termasuk dalam kerangka Program Guru Penggerak, masih cenderung menekankan aspek kurikulum dan pedagogi secara umum, sementara kompetensi teknis perancangan instrumen penilaian autentik, seperti cara mengonstruksi rubrik analitik yang valid, teknik observasi sistematis untuk penilaian karakter, atau strategi memberikan umpan balik formatif yang mendorong regulasi diri siswa, belum mendapatkan porsi yang proporsional (Sulaiman & Bhattacharyya, 2020). Program pelatihan yang efektif untuk tujuan ini harus bersifat job-embedded, artinya terintegrasi dalam praktik mengajar sehari-hari guru melalui pendekatan komunitas belajar profesional (professional learning community), bukan sekadar pelatihan workshop satu kali yang terpisah dari konteks kelas nyata (Darling-Hammond & Bransford, 2005).

Implikasi kedua menyasar dimensi teknologi. Kajian ini merekomendasikan agar pengembangan Platform Merdeka Mengajar (PMM) diperkuat dengan modul interaktif perancangan instrumen penilaian yang tidak sekadar menyediakan contoh, melainkan memandu guru melalui proses berpikir dalam mengonstruksi rubrik, menentukan KKTP, dan menginterpretasi hasil asesmen diagnostic, sejalan dengan tingkat kesiapan yang diuraikan dalam Model TRIAD. Integrasi kecerdasan buatan (artificial intelligence) untuk memberikan umpan balik otomatis terhadap kualitas instrumen yang

.

dirancang guru, sebagaimana telah dieksperimentasikan di beberapa negara maju, dapat menjadi terobosan yang mengakselerasi pengembangan kapasitas guru secara masif dan merata (Zawacki-Richter et al., 2019). Implikasi ketiga berkaitan dengan kebijakan penilaian di tingkat satuan pendidikan: kepala sekolah perlu didorong untuk menjadikan kajian rubrik dan sharing instrumen penilaian antarguru sebagai agenda tetap dalam supervisi akademik, mengubah praktik penilaian dari domain pribadi guru menjadi tanggung jawab kolektif komunitas profesional di sekolah (Timperley et al., 2020).

DAFTAR PUSTAKA

Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university: What the student does* (4th ed.). Open University Press.

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, *25*(6), 551–575. https://doi.org/10.1080/0969594X.2018.1441807

Booth, A., Sutton, A., Clowes, M., & Martyn-St James, M. (2021). *Systematic approaches to a successful literature review* (3rd ed.). SAGE Publications.

Corbin, J., & Strauss, A. (2015). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (4th ed.). SAGE Publications.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.

Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for educational practice of the science of learning and development. *Applied Developmental Science*, *24*(2), 97–140. https://doi.org/10.1080/10888691.2018.1537791

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Kurniawan, A. R., Wulandari, B., & Nurhidayah, D. A. (2023). Model integrasi teknologi dalam penilaian autentik berbasis proyek pada pembelajaran vokasi di Indonesia. *Assessment in Education: Principles, Policy & Practice*, *13*(2), 178–192.

Muhadjir, N. (2021). Asesmen nasional: Pengganti ujian nasional untuk pemetaan mutu pendidikan. *Assessment in Education: Principles, Policy & Practice*, *6*(2), 111–124.

Nurdyansyah, & Fahyuni, E. F. (2022). Kesiapan infrastruktur digital sekolah dalam implementasi Kurikulum Merdeka: Studi komparatif wilayah urban dan 3T. *Assessment in Education: Principles, Policy & Practice*, *24*(1), 45–60.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, n71. https://doi.org/10.1136/bmj.n71

Rahmadani, A., & Fitri, A. D. (2022). Efektivitas asesmen formatif berkelanjutan terhadap keterlibatan belajar siswa pada sekolah penggerak. *Assessment in Education: Principles, Policy & Practice*, *26*(2), 183–196.

Rahmawati, Y., & Suryadi, E. (2022). Transformasi budaya penilaian di sekolah: Antara tekanan ujian dan amanat kurikulum berbasis kompetensi. *Assessment in Education: Principles, Policy & Practice*, *28*(1), 34–46.

Sahlberg, P. (2021). *Finnish lessons 3.0: What can the world learn from educational change in Finland?* (3rd ed.). Teachers College Press.

Sari, D. P., & Suwandi, S. (2021). Pengaruh penilaian autentik terhadap motivasi belajar dan kemampuan berpikir tingkat tinggi siswa SMA: Studi quasi-eksperimen di Surakarta. Pedagogia. *Assessment in Education: Principles, Policy & Practice*, *10*(2), 153–168.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Tan, C. (2022). Holistic education and assessment in Singapore. *Asia Pacific Journal of Education*, *42*(3), 444–455.

Timperley, H., Ell, F., & Le Fevre, D. (2020). Professional learning conversations: Challenges in using evidence for improvement. *Professional Development in Education*, *46*(4), 585–601.

Wahyudi, W., Marzuki, M., & Supriyono, S. (2023). Pengembangan instrumen penilaian dimensi Profil Pelajar Pancasila pada jenjang sekolah menengah pertama. *Assessment in Education: Principles, Policy & Practice*, *7*(1), 567–581.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International*

Journal of Educational Technology in Higher Education, 16(1), 39. https://doi.org/10.1186/s41239-019-0171-0