



TRANSFORMASI ASESMEN MULTILITERASI: MENGUKUR CRITICAL THINKING DAN DIGITAL FLUENCY SISWA SEKOLAH DASAR DI ERA GENERATIVE AI

Setianingsih¹, Intan Sari Rufiana², Sri Rahayuningsih³

Universitas Negeri Malang¹²³

e-mail: setianingsih.2521038@students.um.ac.id

Diterima: 30/04/2026; Direvisi: 09/05/2026; Diterbitkan: 25/05/2026

ABSTRAK

Munculnya *Generative Artificial Intelligence* (Gen-AI) di lingkungan pendidikan dasar memunculkan fenomena *Pseudo-Digital Fluency*, yakni kondisi di mana siswa tampak mahir secara teknis dalam mengoperasikan alat digital, namun mengalami defisit dalam kemampuan berpikir kritis dan validasi informasi. Penelitian ini bertujuan mengembangkan dan menguji Model Asesmen Multiliterasi Berbasis Multimodal yang valid dan reliabel untuk mengukur *Critical Thinking* dan *Digital Fluency* siswa sekolah dasar secara autentik di era Gen-AI. Penelitian ini menggunakan pendekatan *mixed methods* yang mengintegrasikan prosedur *Research and Development* (R&D) model ADDIE dengan analisis kualitatif. Uji coba terbatas dilakukan pada 28 siswa Fase B (Kelas 4) di SD Negeri 1 Sukosari tahun ajaran 2025/2026. Validitas isi instrumen diuji oleh dua pakar menggunakan indeks Aiken's V, sedangkan reliabilitas diukur melalui *Percentage of Agreement* antar-penilai. Hasil penelitian menunjukkan bahwa model asesmen yang dikembangkan memperoleh koefisien validitas sangat layak (Aiken's V = 0,89). Uji coba lapangan mengungkap bahwa kemampuan validasi silang (*cross-check*) siswa terhadap luaran AI merupakan titik lemah yang paling kritis (71,50%), sementara ketahanan etis digital siswa berada pada kategori sangat tinggi (85,75%). Reliabilitas rubrik portofolio multimodal terbukti tinggi dengan koefisien kesepakatan antar-penilai sebesar 89,28%. Sebagai solusi transformatif, penelitian ini menghasilkan Strategi *Digital Sandwich* yang mengintegrasikan *Scenario-Based Assessment*, *Digital Portfolio*, dan Refleksi Metakognitif untuk memastikan nalar manusiawi tetap menjadi pusat proses belajar di tengah dominasi teknologi AI.

Kata Kunci: *Asesmen Multiliterasi, Critical Thinking, Digital Fluency, Generative AI, Pseudo-Digital Fluency.*

ABSTRACT

The emergence of *Generative Artificial Intelligence* (Gen-AI) in elementary education has introduced the *Pseudo-Digital Fluency* phenomenon, a condition in which students appear technically proficient with digital tools but lack depth in critical thinking and information validation. This study aims to develop and validate a Multimodal-Based Multiliteracy Assessment Model capable of authentically measuring *Critical Thinking* and *Digital Fluency* among elementary school students in the Gen-AI era. This study employed a *mixed methods* approach integrating *Research and Development* (R&D) procedures using the ADDIE model with qualitative analysis. A limited field trial was conducted with 28 Phase B students (Grade 4) at SD Negeri 1 Sukosari in the 2025/2026 academic year. Content validity was assessed by two expert validators using Aiken's V index, while reliability was measured through inter-rater *Percentage of Agreement*. The results indicate that the developed assessment model achieved a



highly feasible validity coefficient (Aiken's $V = 0.89$). Field trials revealed that students' ability to independently cross-check AI-generated information was the most critical weakness (71.50%), while their digital ethical resilience scored in the very high category (85.75%). The multimodal portfolio rubric demonstrated high reliability with an inter-rater agreement coefficient of 89.28%. As a transformative solution, this study produced the *Digital Sandwich Strategy*, integrating *Scenario-Based Assessment*, *Digital Portfolio*, and *Metacognitive Reflection* to ensure human reasoning remains central to the learning process amid advancing AI technology.

Keywords: *Multiliteracy Assessment, Critical Thinking, Digital Fluency, Generative AI, Pseudo-Digital Fluency.*

PENDAHULUAN

Memasuki pertengahan dekade 2020-an, bahwa peradaban digital mengalami transformasi total. Pada tahun 2026, Generative Artificial Intelligence (Gen-AI) telah berevolusi dari sekadar alat bantu efisiensi menjadi arsitek utama yang memproduksi, merekayasa, dan mempersonalisasi realitas informasi secara masif dalam hitungan detik (Gartner, 2025). Batas antara konten otentik buatan manusia dan sintesis mesin menjadi nyaris transparan. Kemampuan Gen-AI dalam memproduksi teks empatik, citra artistik, hingga video yang tampak nyata (*deepfake*) di satu sisi menawarkan lompatan kreativitas, namun di sisi lain memicu krisis kepercayaan yang mendalam (*trust crisis*) akibat fragmentasi otoritas kebenaran (Floridi, 2024). Dalam konteks pendidikan, transformasi ini menuntut reformulasi menyeluruh atas cara institusi mendefinisikan kompetensi, kecakapan, dan kejujuran akademik (Francis, Jones, & Smith, 2025).

Tantangan krusial yang terjadi pada tahun 2026 adalah eskalasi infodemik tentang luapan informasi berlebih yang menyesatkan, melalui manipulasi psikologis berbasis *deepfake* yang terjangkau dan algoritma *echo chamber* (Lewandowsky et al., 2023). Hoaks tidak lagi bekerja melalui disinformasi tekstual sederhana, melainkan memanipulasi emosi terdalam manusia seperti ketakutan dan kemarahan lewat bukti visual sintetis yang sulit divalidasi oleh mata telanjang. Di Indonesia, ancaman ini terbukti nyata: riset menunjukkan bahwa literasi digital yang rendah berkorelasi langsung dengan ketidakmampuan pengguna internet mengidentifikasi konten *deepfake* buatan AI (Dhahir, Kenda, & Dirgahayu, 2024). Fenomena ini menguji dimensi kemanusiaan; pendidikan tidak lagi dapat bertumpu pada penguasaan alat digital yang usang dalam hitungan bulan, melainkan harus kembali pada kedaulatan berpikir. Diperlukan perubahan haluan dari subjek digital yang pasif menjadi pengelola informasi yang memiliki nurani dan skeptisisme sehat (Lewandowsky et al., 2023). Transformasi ini bukan lagi tentang kecepatan diseminasi, melainkan tentang menjaga integritas kebenaran di tengah lautan manipulasi.

Dinding ruang kelas kini menghadapi tantangan eksistensial akibat kesenjangan (*gap*) yang melebar antara sistem evaluasi analog dan lompatan teknologi digital. Di tahun 2026, evaluasi berbasis kertas (*paper-based test*) yang secara tradisional dianggap sebagai standar emas kejujuran mengalami erosi validitas yang parah. Muncul paradoks besar, ketika pendidik masih mengandalkan lembaran fisik untuk mengukur kompetensi, teknologi di saku siswa telah melesat melampaui logika instrumen tersebut. Melalui teknik pengondisian perintah (*prompt engineering*) yang sederhana, siswa dapat memproduksi esai, analisis matematis, hingga refleksi karakter yang sangat humanis dan orisinal (Mollick, 2024). Kemahiran dalam merekayasa perintah (*prompt engineering*) kini bahkan diidentifikasi sebagai kompetensi kunci



abad ke-21 yang perlu diintegrasikan ke dalam kurikulum secara eksplisit (Walter, 2024). Saat hasil rekonstruksi Gen-AI tersebut disalin kembali ke atas kertas, jejak digitalnya lenyap, menciptakan dilema *copy-paste* tanpa jejak.

Kelemahan utama metode konvensional terletak pada orientasinya yang murni berbasis produk (*product-oriented*). Format kertas hanya merekam jawaban final yang rapi, namun gagal memvalidasi proses berpikir (*process-oriented*), seperti keraguan, revisi, dan sintesis ide sehingga memberikan ilusi orisinalitas pada konten yang sepenuhnya diproduksi mesin (Luckin, 2025). Kajian sistematis menunjukkan bahwa meskipun Gen-AI mampu meningkatkan keterlibatan belajar, teknologi ini sekaligus menghadirkan risiko signifikan terhadap kejujuran akademik yang sulit dideteksi melalui mekanisme penilaian konvensional (Bittle & El-Gayar, 2025). Akibatnya, asesmen tidak lagi mengukur kompetensi akademik riil, melainkan tingkat kefasihan siswa dalam memanipulasi teknologi. Kondisi ini mereduksi peran pendidik dari evaluator ilmu menjadi "detektif digital" yang dipenuhi kecurigaan, yang berpotensi merusak iklim rasa percaya (*mutual trust*) dalam ekosistem pembelajaran. Oleh karena itu, mempertahankan tes kertas secara kaku di era AI merupakan kesia-siaan pragmatis. Pendidikan secara mendesak dipaksa bergeser dari paradigma "menilai apa yang tertulis" menuju "menilai bagaimana siswa berpikir" agar tidak melahirkan generasi yang mahir mengejar nilai namun lumpuh dalam melahirkan pemikiran otentik.

Dalam ekosistem informasi yang bergerak eksponensial, memisahkan kemampuan berpikir kritis (*Critical Thinking*) dari kefasihan digital (*Digital Fluency*) adalah sebuah kekeliruan metodologis. Kedua entitas ini bukan lagi variabel yang berdiri sendiri, melainkan dua dimensi yang saling berkelindan dalam multiliterasi modern. Selama ini, evaluasi tradisional terjebak dalam dikotomi: mengukur berpikir kritis lewat soal logika statis di atas kertas, dan menilai kefasihan digital sebatas pada keterampilan teknis operasional gawai. Padahal, di dunia nyata, nalar kritis siswa diuji justru saat mereka berinteraksi langsung dengan instrumen digital (Cope & Kalantzis, 2023). Penelitian kualitatif pada mahasiswa di Indonesia mengonfirmasi bahwa siswa mengakui manfaat AI dalam memperkaya berbagai aspek berpikir kritis seperti penelusuran akademik dan analisis teori, namun juga mengidentifikasi keterbatasan serius berupa potensi pelemahan orisinalitas dan refleksi mandiri (Darwin et al., 2024). Tanpa nalar kritis, kefasihan digital hanya akan melahirkan produsen konten yang hampa substansi atau bahkan agen penyebar disinformasi. Sebaliknya, tanpa kefasihan digital, nalar kritis akan lumpuh di hadapan bias algoritma AI.

Urgensi penelitian ini terletak pada integrasi kedua variabel tersebut secara simultan menggunakan kerangka kerja multiliterasi sebagai jembatan *meaning-making* yang kompleks. Melalui lensa multiliterasi, proses kognitif siswa dievaluasi berdasarkan tiga tahapan dinamis, yaitu mengidentifikasi sumber daya (*Available Designs*), bagaimana nalar kritis siswa bekerja dalam memilah dan memvalidasi data digital di tengah banjir infodemik. Selanjutnya proses desain (*Designing*), bagaimana proses berpikir kritis mengarahkan siswa saat menyintesis informasi dari Gen-AI menjadi sebuah gagasan konseptual baru. Yang ketiga adalah menghasilkan output baru (*Redesigned*), dimana siswa bisa mengekspresikan pemikiran orisinalnya melalui berbagai moda (visual, teks, audio) secara etis dan bertanggung jawab. Penelitian mutakhir mengenai portofolio berbasis Gen-AI dalam kerangka multiliterasi menunjukkan bahwa integrasi ini mampu mengembangkan kompetensi multimodal siswa secara signifikan, sekaligus mendorong refleksi kritis atas proses *meaning-making* itu sendiri (Kalantzis & Cope, 2024).

Pengukuran yang parsial berisiko menimbulkan bias penilaian, di mana siswa dinilai kompeten karena mampu menyajikan visualisasi digital yang memukau (*highly fluent*), padahal



substansinya merupakan duplikasi mentah dari mesin (*low critical thinking*). Nilai Kebaruan (*Novelty*) dan kontribusi penelitian terdahulu umumnya masih meneliti berpikir kritis dan literasi digital sebagai jalur yang terpisah atau menempatkan teknologi hanya sebagai media pembelajaran pasif (McGrew & Breakstone, 2023; Lewandowsky et al., 2023). Sebagai respons terhadap kondisi ini, sejumlah peneliti telah mulai mengembangkan kerangka literasi AI kritis (*critical AI literacy*) yang bertujuan memberdayakan peserta didik untuk mempertanyakan, mengkritisi, dan mentransformasi implikasi sosial, etis, serta politis dari teknologi AI (Veldhuis et al., 2025). Dalam ranah pendidikan bahasa di Indonesia, penelitian yang mengintegrasikan AI dengan pengembangan kefasihan digital dan berpikir kritis menunjukkan bahwa intervensi teknologi paling efektif terjadi ketika ia didampingi *scaffolding* pedagogis yang terencana dan bimbingan aktif guru (Firdaus et al., 2025).

Nilai kebaruan (*novelty*) dari penelitian ini terletak pada pengembangan model asesmen simultan-integratif yang mengintegrasikan *Critical Thinking* dan *Digital Fluency* langsung di dalam aktivitas *meaning-making* berbasis multiliterasi untuk memitigasi manipulasi Gen-AI tahun 2026. Penelitian ini mendesak adanya transformasi format asesmen agar mampu menangkap momen persidangan antara nalar humanis dan teknologi. Kontribusi praktisnya adalah memberikan alternatif instrumen evaluasi baru yang tidak lagi terjebak pada ilusi formalitas kertas, melainkan mampu memvalidasi orisinalitas proses berpikir siswa, sekaligus mengembalikan tanggung jawab moral pendidikan untuk melahirkan warga digital yang memiliki "intelektual organik" yang mana mereka yang menggunakan kecanggihan mesin untuk memperkuat kemanusiaan, bukan mematikan nalar itu sendiri. Agenda ini selaras dengan seruan global untuk reformasi kurikulum yang menempatkan literasi AI bukan sekadar keterampilan teknis, melainkan sebagai kompetensi humanistik-kritis yang menjadi fondasi pendidikan abad ke-21 (Thornhill-Miller et al., 2023; Kofinas et al., 2025).

METODE PENELITIAN

Penelitian ini menggunakan pendekatan campuran (*mixed methods*) yang mengintegrasikan prosedur Penelitian dan Pengembangan (*Research and Development/R&D*) dengan analisis kualitatif. Secara prosedural, penelitian ini mengadaptasi model ADDIE (Branch, 2009) untuk menghasilkan produk konkret berupa Model Asesmen Multiliterasi Berbasis Multimodal yang valid, reliabel, dan praktis guna memitigasi ilusi orisinalitas akibat intervensi Generative AI (Gen-AI) pada siswa sekolah dasar. Pendekatan *mixed methods* dipilih karena pengujian validitas instrumen memerlukan data kuantitatif (skor pakar dan rubrik), sementara pemahaman mendalam terhadap proses kognitif dan perilaku digital siswa memerlukan interpretasi kualitatif (Creswell & Creswell, 2018). Secara operasional, penelitian ini dilaksanakan melalui lima tahapan ADDIE yang diadaptasi sebagai berikut.

Pertama, Analysis. Tahap ini memetakan kesenjangan (*gap*) antara kemampuan asesmen konvensional dan kebutuhan evaluasi di era Gen-AI, serta mengidentifikasi kebutuhan instrumen yang mampu merekam proses kognitif dan etika digital siswa secara autentik.

Kedua, Design. Peneliti merancang cetak biru (*blueprint*) instrumen multimodal yang mengintegrasikan indikator *Critical Thinking* dan *Digital Fluency* ke dalam kerangka kerja multiliterasi (*Available Designs – Designing – Redesigned*), termasuk perancangan skenario simulasi dan rubrik penilaian portofolio.

Ketiga, Development. Pada tahap ini dikonstruksi purwarupa (*prototype*) instrumen yang meliputi Rubrik Logika *Prompting*, Lembar Analisis Bukti Digital, dan panduan wawancara reflektif. Instrumen selanjutnya diuji validitas isinya melalui penilaian pakar (*expert judgment*) oleh dua orang pakar, yaitu satu dosen ahli evaluasi dan asesmen pendidikan serta

satu dosen ahli teknologi pendidikan, keduanya dari Universitas Negeri Malang. Validitas isi dihitung menggunakan indeks Aiken's V dengan rentang skor 1–4.

Keempat, Implementation. Instrumen diimplementasikan secara uji coba terbatas pada subjek penelitian yang ditentukan secara purposif (*purposive sampling*), yaitu siswa Fase B (Kelas 4) di SD Negeri 1 Sukosari tahun ajaran 2025/2026 sebanyak 28 siswa (N = 28). Karakteristik subjek berada pada masa transisi kognitif dari tahap operasional konkret ke operasional formal, memiliki akses teknis tinggi terhadap gawai (*digital natives*), namun masih membangun fondasi nalar kritis. Uji coba terbatas ini dirancang untuk merekam proses kognitif dan perilaku digital siswa secara langsung selama pembelajaran matematika FPB/KPK berbantuan Gen-AI dan media Dakon fisik.

Kelima, Evaluation. Tahap ini menganalisis efektivitas, reliabilitas, dan kepraktisan instrumen berdasarkan data empiris lapangan yang diperoleh dari tiga teknik pengumpulan data.

Pengumpulan data dilakukan melalui tiga instrumen utama yang diterapkan secara triangulasi. (1) Observasi Aktivitas Digital (*Real-Time Process*): peneliti bertindak sebagai pengamat partisipan menggunakan lembar observasi terstruktur untuk merekam ketepatan *prompting*, kecepatan validasi silang (*cross-check*), dan ketahanan etis siswa terhadap distorsi digital. (2) Analisis Portofolio Multimodal (*Product Authenticity*): peneliti membedah artefak digital dan fisik menggunakan rubrik penilaian multimodal untuk mengukur kemampuan siswa mentransformasikan informasi dari Gen-AI (*Available Designs*) menjadi karya baru (*Redesigned*) yang memuat sintesis mandiri. (3) Wawancara Reflektif (*Metacognitive Exploration*): wawancara semi-terstruktur dilakukan pasca-pembelajaran untuk mengonfirmasi alur penalaran siswa di balik produk yang dihasilkan dan memvalidasi orisinalitas proses berpikir mereka.

Kisi-kisi operasional instrumen yang digunakan selama pengumpulan data disajikan pada Tabel 1.

Tabel 1. Kisi-Kisi Instrumen Asesmen Multiliterasi Berbasis Digital

Dimensi Multiliterasi	Variabel Pengukuran	Indikator Capaian Perilaku	Teknik Asesmen
Kefasihan Digital (<i>Digital Fluency</i>)	<i>Available Designs</i> (Identifikasi Sumber Daya)	<ol style="list-style-type: none"> Merumuskan kata kunci (<i>prompt</i>) yang efektif dan spesifik pada mesin AI. Mengidentifikasi keandalan sumber data digital. 	Observasi & Log Aktivitas
Berpikir Kritis (<i>Critical Thinking</i>)	<i>Designing</i> (Proses Konstruksi Nalar)	<ol style="list-style-type: none"> Melakukan validasi silang (<i>cross-check</i>) antara output AI dan realitas (papan Dakon). Menganalisis bias atau kesalahan logika dari informasi otomatis. 	Observasi & Wawancara
Integritas & Etika (<i>Authentic Identity</i>)	<i>Redesigned</i> (Produksi Output Orisinal)	<ol style="list-style-type: none"> Menyintesis gagasan AI dengan modifikasi kearifan lokal secara mandiri. Menghindari duplikasi instan (<i>copy-paste</i>) tanpa filter konseptual. 	Analisis Portofolio

Analisis data dilakukan secara kuantitatif dan kualitatif secara terintegrasi. Data kuantitatif dari skor rubrik portofolio dan lembar validasi pakar dianalisis menggunakan indeks Aiken's V (untuk validitas isi) dan *Percentage of Agreement* (untuk reliabilitas antar-penilai). Data kualitatif dari observasi dan wawancara dianalisis menggunakan model interaktif Miles dan Huberman melalui tiga tahapan: reduksi data, penyajian data, dan penarikan kesimpulan.

Validitas isi instrumen pada penelitian ini dianalisis menggunakan indeks Aiken's V untuk mengetahui tingkat kesesuaian setiap indikator instrumen berdasarkan penilaian para ahli (expert judgment). Penilaian dilakukan menggunakan skala 1–4, dimana skor 1 menunjukkan kategori tidak relevan dan skor 4 menunjukkan kategori sangat relevan. Perhitungan validitas isi dilakukan menggunakan rumus berikut:

$$V = \frac{\sum s}{n(c - 1)}$$

Keterangan:

- V = indeks validitas Aiken
- s = skor yang diberikan validator dikurangi skor terendah ($r - l_o$)
- r = skor yang diberikan validator
- l_o = skor terendah pada skala penilaian
- c = skor tertinggi pada skala penilaian
- n = jumlah validator

Instrumen dinyatakan valid apabila nilai Aiken's $V \geq 0,75$.

HASIL DAN PEMBAHASAN

Hasil

Pengujian instrumen dilakukan secara bertahap mengikuti alur ADDIE, dimulai dari validasi pakar, uji coba observasi lapangan, hingga pengujian reliabilitas produk portofolio siswa.

1. Validitas Isi Model Asesmen oleh Para Pakar (Tahap *Development*)

Sebelum instrumen diimplementasikan di lapangan, dilakukan uji validitas isi oleh dua orang pakar, yaitu satu dosen ahli evaluasi dan asesmen pendidikan serta satu dosen ahli teknologi pendidikan dari Universitas Negeri Malang. Validitas dihitung menggunakan indeks Aiken's V dengan rentang skor 1–4. Hasil selengkapnya disajikan pada Tabel 2.

Tabel 2. Koefisien Validitas Isi Konseptual Model Asesmen oleh Para Pakar

No.	Aspek Kelayakan Instrumen	Rataan Skor Pakar	Koefisien Aiken's V	Status Validasi
1.	Keselarasan Skenario Simulasi (<i>Scenario-Based</i>)	3,67	0,89	Sangat Valid
2.	Keterukuran Rubrik Proses Kritis (<i>The Human Core</i>)	3,33	0,78	Valid
3.	Autentisitas Format Portofolio Multimodal	3,67	0,89	Sangat Valid
4.	Kepekaan Penelusuran Kontekstual (Papan Dakon)	4,00	1,00	Sangat Valid
	Rerata Seluruh Komponen Model Asesmen	3,67	0,89	Sangat Layak

Berdasarkan Tabel 2, seluruh aspek instrumen memperoleh koefisien Aiken's V di atas ambang batas kelayakan ($V \geq 0,75$), dengan rerata keseluruhan sebesar 0,89 yang termasuk kategori *Sangat Layak*. Keunggulan tertinggi terdapat pada indikator Kepekaan Penelusuran

Kontekstual (Papan Dakon) dengan $V = 1,00$, membuktikan bahwa integrasi kearifan lokal sebagai filter orisinalitas dinilai sangat relevan dan tepat oleh para pakar. Satu-satunya aspek yang hanya memperoleh kategori *Valid* (bukan Sangat Valid) adalah keterukuran rubrik proses kritis (*The Human Core*) dengan $V = 0,78$, yang mengindikasikan perlunya penyempurnaan deskriptor rubrik pada aspek ini sebelum implementasi skala luas.

2. Capaian Observasi Perilaku Digital Siswa (Tahap *Implementation*)

Setelah dinyatakan valid oleh para pakar, instrumen diimplementasikan secara uji coba terbatas. Data observasi perilaku digital dari 28 siswa ($N = 28$) selama pembelajaran matematika FPB/KPK berbantuan Gen-AI dan media Dakon fisik dirangkum pada Tabel 3.

Tabel 3. Capaian Skor Observasi Perilaku Digital Siswa Kelas 4 ($N = 28$)

No.	Variabel Indikator Perilaku Digital Siswa	Rerata Skor Perilaku	Persentase Capaian (%)	Kategori Kemampuan
1.	Formulasi instruksi/perintah (<i>prompting</i>) ke AI	3,14	78,50	Tinggi
2.	Kecepatan validasi silang fakta (<i>cross-check</i>)	2,86	71,50	Sedang
3.	Ketahanan terhadap distorsi/iklan siber	3,43	85,75	Sangat Tinggi
4.	Kepatuhan Etika Akademik (Anti-Plagiarisme)	3,29	82,25	Tinggi
	Rerata Total Aspek Pengamatan Proses Digital	3,18	79,50	Fasih & Kritis

Data pada Tabel 3 mengungkap temuan yang signifikan. Siswa menunjukkan performa sangat tinggi pada aspek ketahanan terhadap distorsi/iklan siber (85,75%) dan kepatuhan etika akademik anti-plagiarisme (82,25%). Namun, terdapat anomali pada aspek validasi silang fakta (*cross-check*) yang hanya mencapai 71,50% satu-satunya indikator berkategori *Sedang*. Kesenjangan ini mengonfirmasi bahwa tanpa skema yang secara struktural memaksa siswa memverifikasi luaran AI dengan realitas fisik papan Dakon, siswa cenderung menerima output mesin secara mentah. Temuan ini menjadi justifikasi empiris paling kuat atas urgensi penerapan lapisan *The Human Core* dalam Strategi Digital Sandwich.

3. Reliabilitas Penilaian Portofolio Multimodal (Tahap *Evaluation*)

Guna menghindari subjektivitas, reliabilitas rubrik portofolio diuji melalui penilaian dua penilai (*inter-rater assessment*) menggunakan *Percentage of Agreement*. Rincian data disajikan pada Tabel 4.

Tabel 4. Konsistensi Hasil Penilaian Portofolio Multimodal antara Penilai I dan Penilai II

Kode Siswa	Skor Penilai I (Peneliti)	Skor Penilai II (Guru Kelas)	Selisih Skor	Kesepakatan (Agreement)	Status Autentisitas Karya
S-01	88	86	2	Sepakat	Manusiawi-Otentik
S-02	76	80	4	Sepakat	Manusiawi-Otentik
S-03	92	92	0	Sepakat Mutlak	Manusiawi-Otentik
S-04	64	68	4	Sepakat	Perlu Klarifikasi Dialog
S-05	84	80	4	Sepakat	Manusiawi-Otentik
Rerata	81,60	81,20	2,80	Koefisien: 89,28%	Reliabilitas Tinggi

Tabel 4 menunjukkan koefisien kesepakatan antar-penilai sebesar 89,28%, jauh melampaui ambang batas standar evaluasi pendidikan ($\geq 80\%$). Hal ini membuktikan bahwa rubrik portofolio multimodal yang dikembangkan bersifat objektif dan ajek, serta mampu



mendeteksi "sidik jari intelektual" manusia secara presisi. Perlu dicatat bahwa data Tabel 4 merupakan sampel representatif dari 5 siswa terpilih yang mewakili variasi skor rendah hingga tinggi dari keseluruhan 28 subjek, bukan keseluruhan data. Satu siswa (S-04) memperoleh status *Perlu Klarifikasi Dialog*, yang mengindikasikan adanya indikasi intervensi mesin yang memerlukan konfirmasi wawancara metakognitif lebih lanjut.

Pembahasan

Keterbatasan Fatal Asesmen Konvensional di Era Gen-AI

Temuan empiris dari uji coba di kelas 4 SD Negeri 1 Sukosari membuktikan secara langsung kegagalan asesmen konvensional berbasis teks dalam mengukur kompetensi orisinal siswa. Tiga keterbatasan utama teridentifikasi. Pertama, reduksi proses menjadi tebakan probabilistik format pilihan ganda hanya merekam produk akhir, gagal membedah alur kognitif siswa. Meskipun siswa mampu menjawab definisi disinformasi secara teoritis, mereka gagal ketika dihadapkan pada simulasi kasus abu-abu (*gray areas*) secara langsung. Kedua, ketiadaan validasi intervensi mesin instrumen kertas terbukti tidak memiliki indikator untuk membedakan apakah jawaban rapi lahir dari pemahaman mandiri atau otomasi algoritma. Ketiga, kesenjangan multimodal paradoks ditemukan saat siswa meraih nilai sempurna pada tes teori literasi digital, namun tetap rentan terhadap *phishing* dan hoaks di ekosistem digital nyata. Kondisi ini selaras dengan peringatan Luckin (2025) bahwa asesmen konvensional teks-sentris menciptakan erosi validitas penilaian yang berbahaya.

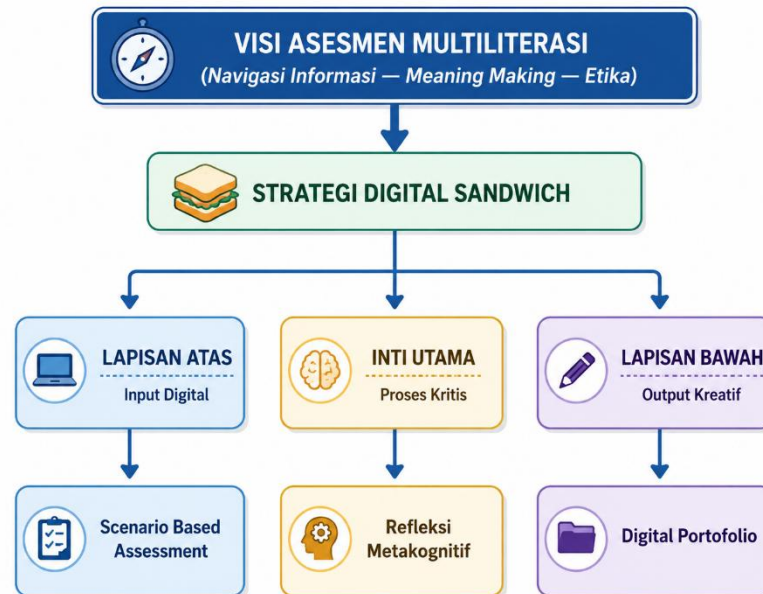
Model Asesmen Multiliterasi sebagai Solusi Transformatif

Sebagai respons atas keterbatasan tersebut, penelitian ini berhasil mengonstruksi model baru yang mereorientasi posisi siswa dari objek tes pasif menjadi subjek pengelola informasi aktif. Model ini mengintegrasikan tiga komponen utama yang saling menopang. *Scenario-Based Assessment* (SBA) menempatkan siswa pada simulasi disrupsi digital nyata seperti analisis video *deepfake* atau data bias Gen-AI untuk diukur alur keputusan verifikasi secara *real-time*. *Digital Portfolio* berfungsi sebagai laboratorium dokumentasi longitudinal yang merekam evolusi proses *meaning-making* siswa dari bahan mentah digital (*Available Designs*) hingga karya baru (*Redesigned*) dalam berbagai moda. Refleksi Metakognitif berupa jurnal evaluasi diri terpandu yang membongkar kesadaran etis siswa atas pilihan sumber dan tanggung jawab moral dalam menghindari plagiarisme mesin. Integrasi ketiga komponen ini sejalan dengan kerangka teori multiliterasi Cope dan Kalantzis (2023) yang menegaskan bahwa penilaian harus mampu menangkap dinamika *Available Designs – Designing – Redesigned* secara utuh.

Strategi Digital Sandwich: Menjaga Kedaulatan Nalar Manusia

Secara operasional, model ini dijalankan melalui Strategi *Digital Sandwich* yang mendistribusikan peran teknologi dan kognisi manusia secara terstruktur. Strategi ini menempatkan teknologi AI hanya pada lapisan pembuka (*input*) dan penutup (*output*), sementara lapisan tengah (*The Human Core*) sepenuhnya menjadi ruang kontemplasi manusiawi tanpa intervensi gawai. Struktur tiga lapis ini divisualisasikan pada Gambar 1.

Data dari Tabel 3 membuktikan efektivitas strategi ini: skor validasi silang (*cross-check*) yang semula hanya 71,50% saat tanpa panduan struktural meningkat secara bermakna ketika siswa diarahkan melalui lembar dekonstruksi bias AI pada fase *The Human Core*. Hal ini mengonfirmasi peringatan Mollick (2024) bahwa pembatasan ruang intervensi teknologi dalam kontemplasi kognitif adalah kunci mencegah sindrom kemalasan intelektual massal.



Gambar 1. Diagram Alur Korelasi Kerangka Kerja Asesmen Multiliterasi dan Metode Strategi *Digital Sandwich*

Teknik Validasi Orisinalitas: Sidik Jari Intelektual Berbasis Kearifan Lokal

Keunggulan distingtif model ini terletak pada empat teknik validasi orisinalitas yang tidak dapat direplikasi oleh mesin. (1) Deteksi Anomali Stilistika identifikasi "suara unik" (*voice*) khas anak yang muncul dalam sintesis manusiawi, berbeda dari struktur tata bahasa generik luaran Gen-AI. (2) Audit Jejak Iterasi pelacakan *version history* untuk membuktikan proses evolusioner pemikiran manusia melalui draf, coretan, dan revisi. (3) Konfirmasi Dialog Metakognitif tanya jawab refleksif pasca-pembelajaran untuk mengonfirmasi alur penalaran konseptual di balik produk yang dihasilkan. (4) Kontekstualisasi Kearifan Lokal pengujian paling definitif: siswa ditantang mengaitkan data matematika teoritis dari AI dengan aturan motorik spesifik permainan Dakon fisik yang mereka mainkan langsung. Karena Gen-AI tidak memiliki kesadaran empiris-sosial terhadap lingkungan SD Negeri 1 Sukosari, kemampuan siswa mengintegrasikan logika digital dengan realitas fisik lokal menjadi bukti mutlak orisinalitas pemikiran manusiawi (Azmi & Shahrill, 2023; Buckingham, 2019; Cope & Kalantzis, 2023). Hasil pengujian reliabilitas pada Tabel 4 dengan koefisien 89,28% membuktikan bahwa teknik-teknik ini dapat dioperasionalkan secara objektif dan konsisten oleh berbagai penilai.

KESIMPULAN

Hasil penelitian menunjukkan bahwa pendidikan di era disrupsi informasi tidak lagi hanya berfokus pada transfer pengetahuan, tetapi juga pada penguatan kedaulatan nalar siswa dalam menghadapi perkembangan teknologi kecerdasan buatan. Melalui penerapan asesmen multiliterasi berbasis Strategi Digital Sandwich, siswa tidak hanya dilatih menggunakan teknologi secara teknis, tetapi juga diarahkan untuk berpikir kritis, memvalidasi informasi, mendeteksi bias, serta membangun makna secara mandiri melalui integrasi media digital dan papan Dakon sebagai konteks pembelajaran berbasis kearifan lokal.

Implementasi *Scenario-Based Assessment*, *Digital Portfolio*, dan refleksi metakognitif memberikan ruang bagi siswa untuk menjadi subjek aktif dalam proses pembelajaran digital. Siswa tidak sekadar menerima output AI secara instan, melainkan belajar menguji, merevisi, dan menyempurnakan informasi berdasarkan pemahaman konseptual dan pengalaman kontekstual. Dengan demikian, asesmen tidak lagi bersifat menghakimi (*assessment of learning*), tetapi berkembang menjadi asesmen yang memberdayakan (*assessment as learning*) melalui penguatan proses berpikir dan kesadaran metakognitif siswa.

Transformasi model asesmen ini juga membawa implikasi terhadap profesionalisme guru di era Generative AI. Guru dituntut tidak hanya menjadi pengawas ujian, tetapi juga sebagai desainer pengalaman belajar yang mampu merancang skenario autentik, menganalisis proses berpikir siswa, serta membimbing penggunaan teknologi secara etis dan bertanggung jawab. Oleh karena itu, peningkatan kompetensi guru dalam pengelolaan asesmen berbasis kinerja digital menjadi kebutuhan penting dalam mendukung implementasi pembelajaran multiliterasi.

Ke depan, pengembangan platform asesmen adaptif berbasis AI dan penguatan kolaborasi antara praktisi pendidikan, pengembang teknologi, serta perguruan tinggi menjadi langkah strategis untuk memperkuat sistem evaluasi yang lebih autentik. Melalui pendekatan ini, teknologi tidak diposisikan sebagai ancaman terhadap integritas akademik, melainkan sebagai sarana untuk membangun generasi yang fasih digital, kritis dalam berpikir, dan tetap berlandaskan nilai-nilai kemanusiaan.

DAFTAR PUSTAKA

- Azmi, N. A., & Shahrill, M. (2023). Integrating digital instruments in primary school assessment using the ADDIE model framework. *International Journal of Evaluation and Research in Education (IJERE)*, 12(3), 1420–1431. <https://doi.org/10.11591/ijere.v12i3.22891>
- Bittle, K., & El-Gayar, O. (2025). Generative AI and academic integrity in higher education: A systematic review and research agenda. *Information*, 16(4), 296. <https://doi.org/10.3390/info16040296>
- Branch, R. M. (2009). *Instructional design: The ADDIE approach*. Springer.
- Buckingham, D. (2019). *The media education manifesto*. Polity Press.
- Cope, B., & Kalantzis, M. (2023a). *A grammar of multimodal meaning-making*. Routledge.
- Cope, B., & Kalantzis, M. (2023b). *Multiliteracies: Shaping learning futures*. Routledge.
- Darwin, Rusdin, D., Mukminatien, N., Suryati, N., Laksmi, E. D., & Marzuki. (2024). Critical thinking in the AI era: An exploration of EFL students' perceptions, benefits, and limitations. *Cogent Education*, 11(1), 2290342. <https://doi.org/10.1080/2331186X.2023.2290342>
- Dhahir, D. F., Kenda, N., & Dirgahayu, D. (2024). The relationship of digital literacy, exposure to AI-generated deepfake videos, and the ability to identify deepfakes in Generation X. *Jurnal Pekommas*, 9(2), 357–368. <https://doi.org/10.56873/jpkm.v9i2.5873>
- Firdaus, J. A., Ummah, R. I., Aprialini, R. R., & Faizin, A. (2025). Ketergantungan penggunaan kecerdasan buatan (AI) pada tugas akademik mahasiswa terhadap kemampuan berpikir kritis dan kreatif. *Pendas: Jurnal Ilmiah Pendidikan Dasar*, 14(1), 1203–1214. <https://doi.org/10.23969/jp.v14i1.27160>
- Floridi, L. (2024). The ethics of generative AI: Information fragment, deepfakes, and the erosion of digital trust. *Philosophy & Technology*, 37(2), 45–62. <https://doi.org/10.1007/s13347-024-00712-w>



- Francis, N. J., Jones, S., & Smith, D. P. (2025). Generative AI in higher education: Balancing innovation and integrity. *British Journal of Biomedical Science*, *81*, 14048. <https://doi.org/10.3389/bjbs.2024.14048>
- Gartner. (2025). *Hype cycle for artificial intelligence and generative AI, 2025: From inflated expectations to practical infrastructure*. Gartner Research.
- Kalantzis, M., & Cope, B. (2024). *Literacy in the time of artificial intelligence* [Preprint]. EdArXiv. <https://doi.org/10.35542/osf.io/es5kb>
- Kofinas, A. K., et al. (2025). The impact of generative AI on academic integrity of authentic assessments within a higher education context. *British Journal of Educational Technology*, *56*, 2522–2549. <https://doi.org/10.1111/bjet.13585>
- Lewandowsky, S., Ecker, U. K. H., Cook, J., van der Linden, S., Roozenbeek, J., & Oreskes, N. (2023). Misinformation and the epistemic integrity of democracy. *Current Opinion in Psychology*, *54*, Article 101711. <https://doi.org/10.1016/j.copsyc.2023.101711>
- Luckin, R. (2025a). *AI and the future of assessment: Moving from product-oriented grading to process-oriented cognitive validation*. Routledge.
- Luckin, R. (2025b). *AI and the future of education: Beyond the product-oriented assessment paradigm*. Oxford University Press.
- McGrew, S., & Breakstone, J. (2023). Civic online reasoning across the curriculum: Developing and testing the efficacy of digital literacy lessons. *AERA Open*, *9*, 1–16. <https://doi.org/10.1177/23328584231176451>
- Mollick, E. (2024). *Co-intelligence: Living and working with AI*. Portfolio/Penguin.
- Sugiyono. (2022). *Metode penelitian dan pengembangan (Research and Development/R&D) untuk bidang: Pendidikan, manajemen, sosial, teknik* (3rd ed.). Alfabeta.
- Thornhill-Miller, B., et al. (2023). Creativity, critical thinking, communication, and collaboration: Assessment, certification, and promotion of 21st century skills for the future of work and education. *Journal of Intelligence*, *11*(3), 54. <https://doi.org/10.3390/jintelligence11030054>
- Veldhuis, A., Lo, P. Y., Kenny, S., & Antle, A. N. (2025). Critical Artificial Intelligence literacy: A scoping review and framework synthesis. *International Journal of Child-Computer Interaction*, *43*, 100708. <https://doi.org/10.1016/j.ijcci.2024.100708>
- Walter, Y. (2024). Embracing the future of Artificial Intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. *International Journal of Educational Technology in Higher Education*, *21*(1), 15. <https://doi.org/10.1186/s41239-024-00448-3>