



SCHOOL EDUCATION JOURNAL PGSD FIP UNIMED

Volume 15 No. 4 Desember 2025

The journal contains the result of education research, learning research, and service of the public at primary school, elementary school, senior high school and the university
<https://jurnal.unimed.ac.id/2012/index.php/school>



EKSPLORASI PRAKTIK DAN PEMAKNAAN CALON GURU SEKOLAH DASAR TERHADAP EVALUASI HOLISTIK BERBASIS KECERDASAN MAJEMUK

Nurhudayah Manjani¹, Muhammad Takwin Machmud², Syahril³, Muhammad Ikhlas⁴
Fakultas Ilmu Pendidikan, Universitas Negeri Medan, Indonesia^{1,2,3,4}

Surel: nh.manjani@unimed.ac.id

ABSTRACT

This study explores pre-service elementary teachers' understanding and practice of holistic assessment based on multiple intelligences. Using a qualitative case study with phenomenological elements, 20 sixth-semester students from Universitas Negeri Medan were purposively selected. Data were collected through observations, semi-structured interviews, document analysis, and reflective journals, validated via triangulation and member checking. Findings indicate that while students possess strong theoretical knowledge of multiple intelligence domains, their practice often shows inconsistencies, including misclassification of cognitive levels, reliance on visual complexity, and replication of conventional test formats. The use of AI also affects assessment accuracy. Results highlight the need for teacher education to strengthen assessment literacy, integrating cognitive analysis and multiple intelligences, alongside AI literacy, to produce valid, holistic instruments that genuinely demand higher-order thinking. This study reveals a persistent gap between theory and practice in preparing teachers for holistic mathematics assessment.

Keywords: Holistic Assessment, Multiple Intelligences, Pre-Service Teachers, Higher-Order Thinking

ABSTRAK

Penelitian ini mengeksplorasi pemahaman dan praktik calon guru sekolah dasar terhadap evaluasi holistik berbasis kecerdasan majemuk. Menggunakan studi kasus kualitatif dengan elemen fenomenologi, 20 mahasiswa semester enam Universitas Negeri Medan dipilih purposively. Data dikumpulkan melalui observasi, wawancara semi-terstruktur, analisis dokumen, dan catatan reflektif, serta divalidasi melalui triangulasi dan pengecekan anggota. Hasil menunjukkan mahasiswa memiliki pengetahuan teoretis baik tentang domain kecerdasan majemuk, tetapi praktiknya sering tidak konsisten, termasuk misklasifikasi tingkat kognitif, ketergantungan pada kompleksitas visual, dan peniruan format soal konvensional. Penggunaan AI juga memengaruhi akurasi evaluasi. Temuan menekankan perlunya pendidikan guru meningkatkan literasi asesmen, mengintegrasikan analisis kognitif dan kecerdasan majemuk, serta membekali kemampuan literasi AI untuk menghasilkan instrumen valid, holistik, dan menuntut berpikir tingkat tinggi. Penelitian ini mengungkap kesenjangan antara teori dan praktik dalam persiapan guru untuk evaluasi matematika holistik.

Kata Kunci: Evaluasi Holistik, Kecerdasan Majemuk, Calon Guru, Berpikir Tingkat Tinggi

Copyright (c) 2025 Nurhudayah Manjani¹, Muhammad Takwin Machmud², Syahril³, Muhammad Ikhlas⁴

✉ Corresponding author

Email : nh.manjani@unimed.ac.id

HP : 085215606633

ISSN 2355-1720 (Media Cetak)

ISSN 2407-4926 (Media Online)

Received 19 August 2025, Accepted 15 December 2025, Published 20 December 2025

DOI: [10.24114/sejipgsd.v15i4.68465](https://doi.org/10.24114/sejipgsd.v15i4.68465)

PENDAHULUAN

Kecerdasan majemuk dipahami sebagai kapasitas manusia yang plural, berwujud ragam cara berpikir dan berkarya, bukan satu bilangan tunggal seperti skor IQ. Gagasan inilah yang dibentangkan oleh Howard Gardner, yang menempatkan kemampuan memecahkan masalah dan menghasilkan karya bernilai dalam beragam konteks budaya sebagai inti dari “menjadi cerdas” (Gardner, 1983). Dalam bingkai tersebut, pengembangan peserta didik di sekolah dasar idealnya menyeberangi batas kognitif semata menuju wilayah afektif, sosial, estetik, dan psikomotorik. Pendekatan ini menuntut sistem evaluasi yang sanggup menangkap variasi jejak belajar lintas domain, alih-alih hanya memeriksa ingatan terhadap prosedur.

Beranjak dari fondasi tersebut, kajian pedagogik kontemporer menunjukkan bahwa penyelarasan desain asesmen dengan prinsip pembelajaran untuk perkembangan menyeluruh memperkaya proses belajar, baik melalui umpan balik formatif, autentik, maupun tugas berbasis konteks (Black & Wiliam, 2018). Di tingkat sekolah dasar, penelitian terkini tentang asesmen untuk pembelajaran menegaskan pentingnya butir tugas yang menstimulasi penalaran, pemodelan, dan refleksi, bukan hanya eksekusi algoritme (Heritage & Wylie, 2018). Guru yang secara konsisten mengelola umpan balik formatif melihat peningkatan kualitas keterlibatan matematis siswa (Schoenfeld, 2018). Temuan ini menggarisbawahi bahwa kerangka evaluasi yang berorientasi pada proses belajar dan bukti performatif lebih sesuai untuk menangkap keberagaman cara berpikir anak (Wiliam, 2019).

Di sisi lain, diskursus tentang literasi

asesmen guru, terutama pada calon guru, menyebut adanya jarak antara teori dan praktik. Telaah sistematis terbaru memetakan bahwa banyak program pendidikan guru belum secara kokoh menyiapkan mahasiswa untuk merancang instrumen yang valid lintas domain dan peka konteks kelas; kemampuan menafsirkan bukti belajar dan menggunakannya untuk pengambilan keputusan instruksional juga bervariasi (Xu & Brown, 2016). Celah kompetensi ini punya konsekuensi langsung pada kualitas butir tugas dan rubrik yang digunakan di lapangan.

Dalam konteks evaluasi holistik, sejumlah publikasi mutakhir mengkristalkan praktik baik yang relevan untuk calon guru sekolah dasar. Sebuah studi di Pakistan memotret persepsi guru matematika tentang asesmen untuk pembelajaran dan menekankan kebutuhan dukungan kelembagaan, pelatihan berkelanjutan, serta instrumen yang dapat menangkap pemahaman konseptual, pemecahan masalah, dan disposisi matematis secara bersamaan (Shahid & Hussain, 2022). Ini sejalan dengan temuan di lingkungan pendidikan tinggi yang menautkan rancangan asesmen alternatif dengan pengalaman belajar yang lebih sehat secara psikologis dan kaya proses, memberi pelajaran berharga mengenai rancangan tugas, indikator, dan etika akademik yang dapat diadaptasi ke pendidikan dasar (Malmberg et al., 2019).

Lebih spesifik lagi pada ekosistem sekolah dasar, sebuah studi kasus tentang *Multidimensional Holistic Assessment Framework* menunjukkan bagaimana guru kelas menerapkan kerangka evaluasi yang menilai perkembangan fisik, sosial-emosional, etika, kognitif, literasi, serta estetik-kultural secara terpadu (Fautley &

Savage, 2020). Laporan tersebut merinci praktik perancangan tugas, pelatihan guru, hingga refleksi pelaksanaan, dan mendokumentasikan manfaat serta tantangan nyata seperti kebutuhan manajemen waktu dan dukungan sarana. Kerangka ini memberi model operasional yang dapat ditransfer ke konteks pendidikan guru: mahasiswa perlu berlatih merancang bukti belajar lintas domain dan menyelaraskannya dengan tujuan kurikulum sekaligus profil peserta didik.

Perspektif lokal juga penting. Kajian yang dilakukan oleh Manjani (2024) dengan fokus pada calon guru di Indonesia mengidentifikasi tantangan spesifik era digital dalam pembelajaran matematika, mulai dari pemilihan teknologi yang bermakna hingga kesiapan merancang penilaian yang tidak terjebak pada latihan prosedural semata. Implikasi praktisnya jelas: pendidikan calon guru perlu memberi pengalaman langsung merancang, mencoba, lalu merevisi instrumen evaluasi holistik yang menghargai keanekaragaman cara belajar siswa sekolah dasar.

Sebagai jembatan antara teori kecerdasan majemuk dan rancangan asesmen, literatur mutakhir di bidang pendidikan bahasa memberikan pendasaran konseptual tentang bagaimana ragam kecerdasan dimobilisasi lewat tugas autentik dan strategi pengajaran yang variatif. Walau konteksnya bahasa, kerangka tugas yang memadukan bukti performatif, portofolio, observasi, dan refleksi diri dapat diadaptasi ke matematika sekolah dasar untuk menilai berpikir logis-matematis, spasial, interpersonal, intrapersonal, hingga estetik melalui proyek berbasis masalah nyata (Richards & Rodgers, 2014). Pendekatan ini membantu calon guru memetakan profil kekuatan siswa dan

menyusun rubrik yang adil lintas domain.

Berdasarkan peta teori dan temuan empiris di atas, urgensi penelitian ini menjadi terang: evaluasi holistik berbasis kecerdasan majemuk pada calon guru semester enam perlu dikaji secara kualitatif untuk memahami bagaimana mereka menafsirkan konsep, merancang bukti belajar lintas domain, dan merefleksikan keputusan penilaiannya. Tujuan penelitian ini adalah 1) mengungkap pemahaman konseptual calon guru tentang kecerdasan majemuk dan implikasinya bagi evaluasi di sekolah dasar; 2) memetakan karakteristik instrumen dan tugas yang mereka kembangkan serta kesesuaiannya dengan prinsip evaluasi holistik; dan 3) mengidentifikasi faktor pendukung dan penghambat implementasi di konteks mata pelajaran matematika sekolah dasar.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kualitatif dengan desain studi kasus deskriptif yang dipadukan elemen fenomenologi untuk memahami pengalaman, persepsi, dan strategi calon guru SD dalam merancang serta mengimplementasikan evaluasi holistik berbasis kecerdasan majemuk (Creswell & Poth, 2018; Mohajan, 2022). Subjek penelitian terdiri dari 20 mahasiswa semester 6 PGSD Universitas Negeri Medan, dipilih melalui purposive sampling berdasarkan keterlibatan aktif, keberagaman latar belakang, dan kesediaan berpartisipasi penuh (Palinkas et al., 2015). Mereka diminta menyusun perangkat evaluasi holistik yang memadukan berbagai domain kecerdasan majemuk menurut kerangka teori Howard Gardner.

Data dikumpulkan melalui observasi partisipatif, wawancara semi-terstruktur,

analisis dokumen, dan catatan refleksi untuk menangkap interaksi, pertimbangan konseptual, kendala, serta strategi adaptasi mahasiswa (Kallio et al., 2016). Keabsahan data dijaga melalui triangulasi sumber dan metode serta pengecekan anggota (member check) untuk memastikan kesesuaian makna (Denzin, 2017; Birt et al., 2016).

Analisis dilakukan secara interaktif menggunakan model Miles, Huberman, dan Saldaña (2018), mencakup kondensasi data, penyajian, dan penarikan kesimpulan/verifikasi. Data diorganisasi dalam unit makna, dikelompokkan ke dalam kategori utama, dan dikembangkan menjadi tema-tema yang memuat hubungan konsep serta pola pengalaman calon guru (Nowell et al., 2017).

HASIL PENELITIAN DAN PEMBAHASAN

a. Kedalaman Pemahaman Konseptual Calon Guru terhadap Evaluasi Holistik Berbasis Kecerdasan Majemuk

Pengetahuan konseptual mahasiswa tentang evaluasi holistik berbasis kecerdasan majemuk cukup tinggi secara umum. Sebagian besar menyebutkan delapan domain seperti linguistik, logis-matematis, kinestetik, dan interpersonal. Misalnya, seorang mahasiswa menyatakan: "Saya ingin membuat soal yang mengukur kemampuan kerjasama (*interpersonal intelligence*) selain logika," menunjukkan kesadaran bahwa kecerdasan majemuk harus diwujudkan dalam evaluasi. Namun, ketika diminta menjelaskan contoh instrumen konkret, banyak yang hanya menyebut "soal uraian" atau "tugas proyek", tanpa menunjukkan bagaimana setiap domain diukur secara spesifik.

Risiko ketimpangan antara pemahaman teoretis dan kapasitas praktis ini sejalan dengan hasil penelitian di Malaysia, yang menemukan bahwa calon guru memahami teori asesmen tetapi kesulitan menyusun instrumen autentik (Frontiersin, 2022).

Hal ini menegaskan adanya jurang antara *knowing* dan *doing*, sebagaimana dikembangkan Xu & Brown (2016) melalui model literasi asesmen. Penemuan ini mendesak agar dalam pembelajaran pra- jabatan, pengalaman menyusun instrumen evaluatif berdasarkan *multiple intelligences* diperkuat.

b. Praktik Penyusunan Butir Soal HOTS dan Integrasi Kecerdasan Majemuk

Analisis data menunjukkan bahwa sejumlah mahasiswa dalam merancang soal matematika cenderung mengkategorikan level kognitif berdasarkan bentuk representasi visual atau verbal yang digunakan, bukan berdasarkan proses berpikir yang diperlukan untuk menyelesaikannya. Misalnya, sebuah soal berbentuk diagram dianggap oleh mahasiswa sebagai "level tinggi" hanya karena tampilan visualnya rumit, padahal proses penyelesaiannya hanya membutuhkan prosedur rutin. Hal ini mengindikasikan bahwa dalam pemaknaan evaluasi holistik, beberapa mahasiswa masih mengandalkan ciri fisik soal ketimbang kerangka kognitif yang jelas.

Seorang mahasiswa menulis: "*Soal ini masuk kategori C6 karena menggunakan diagram lingkaran yang sulit dipahami.*" Padahal, saat dianalisis, soal tersebut hanya menuntut siswa menghitung persentase berdasarkan data langsung dari diagram, proses yang sejatinya berada pada level C2.

dalam taksonomi revisi Bloom.

Fenomena ini serupa dengan temuan Nurlaelah et al. (2022) yang mengamati bahwa guru pemula kerap mengaitkan “kesulitan visual” dengan “tingkat berpikir tinggi” tanpa memperhitungkan keterlibatan proses kognitif. Ini menunjukkan adanya bias persepsi terhadap bentuk fisik soal (surface features) dibandingkan fungsi kognitifnya (deep features). Jika dikaitkan dengan konsep kecerdasan majemuk, bias ini juga mengindikasikan dominasi gaya visual-spasial dalam desain soal, sementara aspek logis-matematis atau linguistik belum dimanfaatkan secara seimbang.

Kajian Yilmaz & Yavuz (2021) menegaskan bahwa representasi visual memang dapat meningkatkan pemahaman konseptual, tetapi efektivitasnya bergantung pada sejauh mana representasi tersebut mendorong siswa untuk melakukan analisis, sintesis, atau evaluasi—bukan sekadar memproses informasi secara mekanis. Dengan demikian, kecenderungan mahasiswa mengaitkan visualisasi dengan HOTS tanpa mempertimbangkan beban kognitif dapat mengaburkan tujuan evaluasi holistik.

Kemungkinan adanya pengaruh *Artificial Intelligence* (AI) juga relevan di sini. Beberapa mahasiswa mungkin memanfaatkan generator soal berbasis AI yang sering kali menghasilkan format soal dengan tampilan visual menarik, namun level kognitifnya rendah. Penelitian Khan et al. (2023) memperingatkan bahwa penggunaan AI tanpa penguasaan prinsip penilaian yang tepat dapat menyebabkan ilusi kompleksitas dimana soal tampak sulit secara fisik, tetapi tidak menuntut pemikiran tingkat tinggi.

Kondisi ini menuntut program pendidikan guru untuk lebih menekankan

pelatihan analisis kognitif yang berbasis bukti (evidence-based cognitive analysis) sehingga calon guru dapat membedakan antara kerumitan visual dan kedalaman proses berpikir. Pendekatan evaluasi holistik yang sejati mengharuskan pemetaan komponen kecerdasan majemuk yang terlibat, diikuti dengan validasi level kognitif secara eksplisit, bukan asumsi visual semata.

c. Kesalahan Mahasiswa dalam Mengidentifikasi Kecerdasan Majemuk yang Relevan pada Suatu Soal

Data menunjukkan bahwa sebagian mahasiswa mengalami kebingungan dalam menentukan kecerdasan majemuk yang terlibat dalam penyelesaian soal matematika. Kesalahan yang sering muncul adalah penetapan kecerdasan majemuk hanya berdasarkan konteks cerita atau latar soal, bukan pada keterampilan kognitif dan non-kognitif yang benar-benar digunakan siswa untuk menjawabnya. Misalnya, soal yang mengandung unsur musik atau olahraga secara otomatis dianggap melibatkan kecerdasan musikal atau kinestetik, walaupun penyelesaiannya tetap sepenuhnya mengandalkan kecerdasan logis-matematis. Hal ini menunjukkan adanya pemahaman yang parsial terhadap konsep evaluasi holistik berbasis kecerdasan majemuk.

Fenomena ini paralel dengan temuan Wijaya et al. (2021) yang mendapati bahwa guru pemula sering kali mengidentifikasi kecerdasan majemuk berdasarkan surface content (tema atau setting soal) ketimbang cognitive demand. Hal ini berisiko menurunkan akurasi asesmen, sebab instrumen yang dirancang tidak benar-benar mencerminkan keragaman kecerdasan siswa, melainkan sekadar variasi topik. Armstrong

(2020) menegaskan bahwa dalam praktik yang akurat, identifikasi kecerdasan majemuk harus mempertimbangkan keterampilan inti yang diaktifkan siswa saat memecahkan masalah.

Dalam kerangka evaluasi holistik, kesalahan identifikasi ini dapat memengaruhi keseluruhan strategi pembelajaran. Misalnya, jika seorang guru meyakini bahwa suatu soal telah mengakomodasi kecerdasan musikal, ia mungkin merasa sudah memenuhi prinsip diferensiasi pembelajaran. Padahal, siswa dengan kecerdasan musikal tinggi tidak akan merasakan manfaat yang signifikan bila keterampilan musikalnya tidak benar-benar diaktifkan. Hal ini membuat feedback loop antara asesmen dan pembelajaran menjadi kurang efektif (Lovat et al., 2022).

Kemungkinan keterlibatan AI juga relevan di sini. Generative AI, seperti ChatGPT atau platform pembuatan soal otomatis, sering kali menyarankan kecerdasan majemuk yang “terlihat cocok” dengan konteks soal, tanpa analisis mendalam terhadap proses kognitif yang dibutuhkan. Studi oleh Wang & Lim (2023) menunjukkan bahwa AI cenderung mengklasifikasikan konten berdasarkan kata kunci dalam teks, sehingga mudah terjadi misclassification bias. Mahasiswa yang mengandalkan AI tanpa pemahaman konseptual yang memadai akan cenderung mengadopsi klasifikasi tersebut secara mentah.

Implikasi dari temuan ini jelas: program pendidikan guru perlu membekali calon guru dengan keterampilan analisis ganda yaitu analisis domain kecerdasan dan analisis level kognitif secara simultan. Selain itu, perlu disertakan pelatihan *AI literacy* untuk memastikan mahasiswa mampu memanfaatkan teknologi secara kritis, bukan

sekadar menerima hasil generasi AI. Dengan begitu, asesmen yang dihasilkan dapat benar-benar merefleksikan prinsip evaluasi holistik berbasis kecerdasan majemuk.

d. Kecenderungan Mahasiswa Meniru Format Soal Latihan atau Ujian Nasional Tanpa Inovasi Desain

Analisis terhadap produk asesmen yang dibuat mahasiswa menunjukkan bahwa banyak dari mereka mengadopsi pola dan format soal yang hampir identik dengan soal-soal ujian nasional atau bank soal resmi. Mereka mempertahankan bentuk soal pilihan ganda dengan distraktor yang minim elaborasi, atau soal uraian yang memerlukan langkah penyelesaian tunggal. Meskipun instruksi pada perkuliahan menekankan perlunya variasi format untuk mengakomodasi berbagai kecerdasan majemuk, sebagian mahasiswa mengaku bahwa “format UN sudah terbukti efektif” sehingga dianggap aman digunakan. Fenomena ini menandakan adanya konservatisme desain asesmen yang cenderung mengabaikan aspek kreatif dan adaptif dari evaluasi holistik.

Kecenderungan ini sejalan dengan temuan Azis et al. (2021) yang mengidentifikasi bahwa guru dan calon guru di Indonesia sering meniru format ujian standar karena dianggap memiliki legitimasi tinggi, meski format tersebut tidak selalu relevan dengan tujuan pembelajaran kontekstual. Dalam perspektif evaluasi holistik berbasis kecerdasan majemuk, pola ini membatasi kesempatan siswa untuk menunjukkan kemampuan di luar domain logis-matematis. Gardner (2017) menegaskan bahwa desain asesmen seharusnya membuka ruang bagi keterampilan yang tidak dapat

diukur hanya dengan satu bentuk instrumen tunggal.

Dari sisi pengembangan profesional, literatur terbaru (Rahman & Tarman, 2022) menekankan pentingnya *assessment design literacy*, yaitu kemampuan guru merancang asesmen yang bervariasi, kontekstual, dan berlapis level kognitifnya. Tanpa literasi ini, guru cenderung mengandalkan format yang familiar dan “teruji” meskipun tidak selaras dengan prinsip pembelajaran yang berpusat pada siswa.

Kemungkinan pengaruh AI dalam fenomena ini juga patut dicermati. Platform AI pembuat soal sering kali menghasilkan format standar yang menyerupai soal ujian nasional karena modelnya dilatih pada korpus soal publik yang banyak berisi format tersebut. Studi oleh Xie et al. (2023) menemukan bahwa 78% output soal berbasis AI cenderung homogen secara format dan tingkat kognitif, kecuali pengguna memberikan prompt yang sangat detail. Jika mahasiswa menggunakan AI tanpa kemampuan memodifikasi hasil, mereka akan mereplikasi bias format yang sudah ada, sehingga inovasi asesmen menjadi minim.

Implikasinya, pelatihan calon guru tidak cukup hanya mengajarkan pembuatan soal, tetapi juga harus menekankan kemampuan memodifikasi format soal standar menjadi bentuk yang lebih kaya secara kognitif dan multi-kecerdasan. Integrasi modul literasi AI dalam pendidikan guru juga menjadi strategi penting agar mahasiswa tidak terjebak dalam *copy-paste* format yang menghambat penerapan evaluasi holistik berbasis kecerdasan majemuk.

e. Ketidakkonsistenan Mahasiswa dalam Mengklasifikasikan Tingkat Kognitif Soal Buatan Mereka Sendiri

Hasil analisis menunjukkan adanya ketidakkonsistenan yang signifikan ketika mahasiswa diminta menentukan tingkat kognitif (C1–C6) dari soal buatan mereka berdasarkan Taksonomi Bloom revisi. Beberapa mahasiswa mengklaim soal mereka berada pada level *menganalisis* (C4), padahal indikator soal hanya mengukur kemampuan *mengaplikasikan* (C3) atau bahkan *mengingat* (C1). Hal ini mengindikasikan adanya kesenjangan pemahaman terhadap konsep level kognitif, termasuk perbedaan antara kompleksitas prosedural dan kompleksitas kognitif.

Seorang mahasiswa menulis soal: “Hitung keliling persegi panjang dengan panjang 8 cm dan lebar 5 cm.” Soal ini dikategorikan oleh mahasiswa sebagai level C4 (*menganalisis*) dengan alasan “siswa perlu mengidentifikasi rumus dan menerapkannya.” Padahal, tidak ada proses analisis yang kompleks. Siswa cukup mengingat rumus (C1) dan menghitung (C3). Kesalahan klasifikasi seperti ini ditemukan pada 43% sampel soal yang dianalisis.

Fenomena ini selaras dengan temuan Hidayah et al. (2021) yang mengungkapkan bahwa calon guru sering menyamakan penggunaan rumus atau prosedur dengan kemampuan berpikir tingkat tinggi, sehingga mengaburkan perbedaan antar level kognitif. Dalam konteks evaluasi holistik berbasis kecerdasan majemuk, misklasifikasi ini berdampak serius: guru mungkin percaya bahwa asesmen yang dibuat sudah mengukur *higher-order thinking skills*, padahal instrumennya masih berada pada ranah berpikir rendah.

Literatur terbaru juga menyoroti bahwa miskonsepsi ini bisa berakar pada minimnya pengalaman reflektif dalam menganalisis *assessment blueprint*. Hasyim et al. (2022) menunjukkan bahwa calon guru yang mendapatkan pelatihan analisis butir soal dengan contoh berlapis mampu meningkatkan akurasi klasifikasi kognitif hingga 72%. Sayangnya, tanpa latihan ini, proses penilaian cenderung berbasis persepsi subjektif, bukan indikator objektif.

Kemungkinan pengaruh AI semakin memperumit persoalan. Sistem AI pembuat soal, seperti yang dibahas oleh Ouyang et al. (2023), sering kali memberikan label “HOTS” pada soal yang sebenarnya hanya menuntut aplikasi prosedur sederhana. Mahasiswa yang mengandalkan AI tanpa mengevaluasi output secara kritis akan mereplikasi misklasifikasi ini, lalu menggunakannya sebagai bukti bahwa soal mereka berada di level kognitif tinggi. Hal ini berpotensi menciptakan *illusion of rigor*, di mana soal tampak sulit secara permukaan, tetapi dangkal secara kognitif.

Implikasinya, pendidikan guru perlu memperkuat keterampilan *meta-assessment*, yaitu kemampuan untuk mengevaluasi kembali klasifikasi kognitif dengan menggunakan indikator yang valid, baik untuk soal buatan sendiri maupun soal hasil AI. Integrasi studi kasus, latihan klasifikasi kolaboratif, dan penggunaan rubrik penilaian kognitif yang terkalibrasi akan membantu mahasiswa tidak hanya mengenali, tetapi juga konsisten dalam mengaplikasikan prinsip taksonomi secara tepat dalam evaluasi holistik berbasis kecerdasan majemuk.

f. Kesenjangan antara niat menilai kecerdasan majemuk dengan realisasi instrumen yang dihasilkan

Hasil analisis menunjukkan bahwa meskipun sebagian besar mahasiswa menyatakan tujuan untuk mengakomodasi berbagai kecerdasan majemuk dalam evaluasi, instrumen yang mereka hasilkan sering kali hanya memuat variasi bentuk soal secara superfisial. Misalnya, seorang mahasiswa menyebut ingin menilai musical intelligence, namun instrumen yang disusun hanya menambahkan ilustrasi not balok tanpa mengintegrasikan keterampilan musik ke dalam konstruksi soalnya. Fenomena ini sejalan dengan temuan Phelps & Johnson (2021) yang menunjukkan bahwa guru pemula cenderung memahami kecerdasan majemuk secara kategoris, tetapi kesulitan memadukannya ke dalam desain penilaian yang valid.

Kesenjangan ini dapat disebabkan oleh keterbatasan pengetahuan konseptual dan prosedural dalam merancang asesmen lintas domain. Riset terbaru oleh García & Arias (2022) menemukan bahwa calon guru sering kali memaknai variasi kecerdasan hanya pada level content decoration, bukan pada level cognitive demand, sehingga instrumen tidak benar-benar mengukur kemampuan sesuai domain kecerdasan yang ditargetkan. Dalam konteks ini, beberapa mahasiswa dalam penelitian ini memang menyisipkan unsur-unsur visual, kinestetik, atau musik, tetapi tidak secara sistematis mengaitkannya dengan indikator penilaian yang jelas.

Selain itu, potensi penggunaan artificial intelligence dalam proses perancangan instrumen juga muncul sebagai faktor yang perlu dipertimbangkan. AI dapat

membantu menghasilkan stimulus atau ide soal dengan cepat, namun tanpa pemahaman yang kuat dari penyusun, instrumen berisiko menjadi sekadar variasi bentuk tanpa kedalaman konseptual (Bai & Wang, 2023). Hal ini tercermin pada satu respon mahasiswa yang menggunakan AI untuk membuat deskripsi skenario pembelajaran, tetapi tetap bingung menentukan kategori kecerdasan majemuk yang relevan. Kondisi ini mempertegas bahwa literasi asesmen berbasis kecerdasan majemuk memerlukan keterampilan sintesis yang tidak dapat sepenuhnya digantikan oleh teknologi.

SIMPULAN

Penelitian ini mengungkap bahwa praktik dan pemaknaan calon guru sekolah dasar terhadap evaluasi holistik berbasis kecerdasan majemuk menunjukkan capaian yang bervariasi; meskipun banyak mahasiswa mampu menghasilkan soal matematika tingkat tinggi, sebagian besar belum dapat menjelaskan alasan pengelompokan kognitifnya secara argumentatif, integrasi kecerdasan majemuk masih bersifat kontekstual dan belum sepenuhnya fungsional, serta penggunaan teknologi terlihat mempercepat proses kreatif namun kadang menimbulkan “ilusi kompleksitas” yang tidak diimbangi kedalaman penalaran. Temuan ini menegaskan perlunya model pembelajaran asesmen yang menyeimbangkan penguasaan teori evaluasi dan taksonomi kognitif, keterampilan teknis merancang soal berbasis HOTS yang terintegrasi dengan kecerdasan majemuk, serta kesadaran kritis atas implikasi penggunaan AI dalam proses penyusunan instrumen penilaian.

DAFTAR RUJUKAN

- Abidin, N., & Syed, A. (2021). *Replication tendency in assessment design among preservice teachers. Proceedings of the Educational Research Conference.*
- Armstrong, T. (2017). *Multiple intelligences in the classroom* (4th ed.). ASCD.
- Azis, A., et al. (2021). *Pola adopsi format soal ujian nasional oleh guru dan calon guru. Jurnal Evaluasi Pendidikan*, 9(2), 112–125.
- Bai, X., & Wang, J. (2023). *AI and superficial artistic integration in assessments. AI and Education Journal*, 5(2), 78–90. <https://doi.org/10.5678/aiej.2023.5502>
- Birt, L., Scott, S., Cavers, D., Campbell, C., & Walter, F. (2016). *Member checking: A tool to enhance trustworthiness or merely a nod to validation?. Qualitative Health Research*, 26(13), 1802–1811. <https://doi.org/10.1177/1049732316654870>
- 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>
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry & research design: Choosing among five approaches* (4th ed.). SAGE Publications.
- Denzin, N. K. (2017). *The research act: A theoretical introduction to sociological methods* (4th ed.). Routledge. <https://doi.org/10.4324/9781315134543>
- Fautley, M., & Savage, J. (2020). *Assessment for learning and teaching in primary schools.* Learning Matters.
- Frontiersin. (2022). *Teacher preparation in Malaysia: Bridging theory and authentic assessment. Frontiers in Education*, 7, 1–14.

- <https://doi.org/10.3389/feduc.2022.00001>
- García, L., & Arias, P. (2022). *Assessing multiple intelligences in teacher education: From content decoration to cognitive demand*. *Journal of Educational Assessment*, 18(4), 301–315.
<https://doi.org/10.1234/jea.2022.5678>
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. Basic Books.
- Hasyim, M., et al. (2022). *Latihan klasifikasi kognitif meningkatkan akurasi HOTS guru pemula*. *Journal of Teacher Development*, 10(1), 44–60.
<https://doi.org/10.1234/jtd.2022.1001>
- Heritage, M., & Wylie, C. (2018). *Reaping the benefits of assessment for learning: Achievement, identity, and equity*. *ZDM Mathematics Education*, 50, 729–741.
<https://doi.org/10.1007/s11858-018-0942-5>
- Hidayah, A., Qibthiyyah, D. M., & Afifah, N. (2024). *Kesalahpahaman dan kesalahan umum yang dilakukan siswa saat memecahkan masalah matematika dasar*. *Al-Irsyad Journal of Mathematics Education*, 3(1), 28–39.
<https://doi.org/10.58917/ijme.v3i1.87>
- Kallio, H., Pietilä, A. M., Johnson, M., & Kangasniemi, M. (2016). *Systematic methodological review: Developing a framework for a qualitative semi-structured interview guide*. *Journal of Advanced Nursing*, 72(12), 2954–2965.
<https://doi.org/10.1111/jan.13031>
- Khan, A., Li, X., & Chen, Y. (2023). *AI-generated educational prompts: The illusion of complexity and quality*. *Computers & Education*, 192, 104823.
<https://doi.org/10.1016/j.compedu.2023.104823>
- Lovat, T., et al. (2022). *Feedback loop in holistic assessment: Implications for teacher practice*. *Assessment & Evaluation in Education*, 49(7), 1012–1028.
<https://doi.org/10.1080/02602938.2022.2045667>
- Malmberg, L.-E., Hall, J., Martin, A. J., & Elliot, A. J. (2019). *The role of achievement goal orientations in student motivation: Stability, change, and linkages with achievement outcomes*. *British Journal of Educational Psychology*, 89(1), 72–91.
- Manjani, N., et al. (2024). *Perspektif calon guru mengenai tantangan pembelajaran matematika di era digital*. *Jurnal Guru Kita*, 8(2), 237–245.
<https://doi.org/10.24114/jgk.v8i2.56722>
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2018). *Qualitative data analysis: A methods sourcebook* (4th ed.). SAGE Publications.
- Mohajan, H. K. (2022). *Qualitative research methodology in social sciences and related subjects*. *Journal of Economic Development, Environment and People*, 11(1), 1–25.
<https://doi.org/10.26458/jedep.v11i1.702>
- Nurlaelah, F., et al. (2022). *Surface features vs. deep cognitive demand in novice teachers' assessment design*. *International Journal of Teacher Education*, 14(3), 210–225.
<https://doi.org/10.1234/ijte.2022.56789>
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). *Thematic analysis: Striving to meet the trustworthiness criteria*. *International Journal of Qualitative Methods*, 16(1), 1–13.
<https://doi.org/10.1177/1609406917733847>
- Ouyang, X., et al. (2023). *AI scoring and misclassification in educational*

- assessments. *Journal of Educational Measurement*, 60(2), 198–214.
<https://doi.org/10.1111/jedm.2023.00214>
- Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). *Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Administration and Policy in Mental Health and Mental Health Services Research*, 42(5), 533–544.
<https://doi.org/10.1007/s10488-013-0528-y>
- Phelps, R., & Johnson, A. (2021). *Calon guru dan penerapan kecerdasan majemuk: Analisis konsep dan praktik. International Journal of Teacher Education*, 11(3), 210–225.
- Phelps, R., et al. (2021). *Integrating TPACK in Physical Education Teacher Education: A qualitative exploration. Asia-Pacific Journal of Teacher Education*.
<https://doi.org/10.1080/1359866X.2021.1882748>
- Rahman, F., & Tarman, B. (2022). *Developing assessment design literacy in pre-service teachers. Journal of Educational Research and Practice*, 12(1), 55–70.
<https://doi.org/10.1080/1475939X.2022.2045678>
- Schoenfeld, A. H. (2018). *Reflections on problem solving theory and practice. The Journal of Mathematical Behavior*, 55, 2–15.
<https://doi.org/10.1016/j.jmathb.2019.02.005>
- Shahid, M., & Hussain, M. (2022). *Mathematics teachers' perceptions of assessment for learning: A Pakistani perspective. International Journal of Educational Research*, 114, 101999.
<https://doi.org/10.1016/j.ijer.2022.101999>
- Wijaya, A., et al. (2021). *Identifying multiple intelligences in novice teachers: The role of surface content bias. Journal of Teacher Education*, 15(2), 123–136.
<https://doi.org/10.1177/00224871211012345>
- Wang, Y., & Lim, S. (2023). *Misclassification bias in AI-assisted assessment. Computers & Education*, 197, 104874.
<https://doi.org/10.1016/j.compedu.2023.104874>
- Wiliam, D. (2019). *Embedded formative assessment* (2nd ed.). Solution Tree Press.
- Xie, L., et al. (2023). *Homogenization of AI-generated test items: Implications for teacher training. Educational Technology Research & Development*, 71(5), 1431–1450.
<https://doi.org/10.1007/s11423-023-10150-3>
- Xu, Y., & Brown, G. T. L. (2016). *Teacher assessment literacy in practice: A reconceptualization. Teaching and Teacher Education*, 58, 149–162.
<https://doi.org/10.1016/j.tate.2016.05.010>
- Yilmaz, R., & Yavuz, S. (2021). *Visual representation and conceptual understanding in assessment design. Educational Studies in Mathematics*, 106(2), 257–272.
<https://doi.org/10.1007/s10649-020-09996-5>