

Unjuk Kerja (*Performance Assessment*) sebagai Bentuk Penilaian Autentik dalam Meningkatkan Keterampilan Berpikir Kritis Peserta Didik Sekolah Dasar: Studi Literatur

Erlysta Eka Purnama ✉, Universitas PGRI Madiun

Dhika Mahendra, Universitas PGRI Madiun

Sindy Puji Lestari, Universitas PGRI Madiun

✉ erlystaicha@gmail.com

Abstract: *Critical thinking is a 21st-century competency that elementary students must master, yet empirical studies in Indonesia still report low-to-moderate achievement at this level, partly due to the dominance of written-test-based assessment that fails to authentically capture students' reasoning processes. This article examines the role of performance assessment as a form of authentic assessment in developing elementary students' critical thinking skills, while mapping the research gaps that remain open. The study employs a literature review method with a narrative-thematic approach toward 20 national and international journal sources published between 2015 and 2025, screened through a two-stage selection process that yielded 11 core studies for synthesis in a comparison matrix. The synthesis shows that performance assessment is consistently reported to contribute to students' interpretation, analysis, and inference indicators; however, of the six elementary-level studies found, none explicitly integrated performance assessment with critical thinking measurement within a single instrument. The article concludes that performance assessment holds strong potential for fostering critical thinking in elementary schools, although the empirical evidence remains fragmented, and therefore proposes a conceptual framework linking critical thinking indicators to performance assessment rubrics as an agenda for further development.*

Keywords: *performance assessment, authentic assessment, critical thinking skills.*

Abstrak: Keterampilan berpikir kritis merupakan kompetensi abad ke-21 yang wajib dikuasai peserta didik sekolah dasar, namun berbagai kajian empiris di Indonesia masih menunjukkan capaian yang rendah hingga sedang pada jenjang ini, salah satunya karena dominasi instrumen penilaian berbasis tes tertulis yang kurang menangkap proses bernalar secara autentik. Artikel ini bertujuan mengkaji peran unjuk kerja (*performance assessment*) sebagai bentuk penilaian autentik dalam mengembangkan keterampilan berpikir kritis peserta didik sekolah dasar, sekaligus memetakan kesenjangan penelitian yang masih terbuka. Penelitian menggunakan metode studi literatur dengan pendekatan naratif-tematik terhadap 20 sumber jurnal nasional dan internasional terbitan 2015–2025, disaring melalui dua tahap seleksi hingga menghasilkan 11 studi inti untuk disintesis pada matriks perbandingan. Hasil sintesis menunjukkan unjuk kerja secara konsisten dilaporkan berkontribusi terhadap indikator interpretasi, analisis, dan inferensi peserta didik, tetapi dari enam studi berjenjang SD yang ditemukan, tidak satu pun secara eksplisit mengintegrasikan unjuk kerja dengan pengukuran berpikir kritis dalam satu instrumen. Artikel ini menyimpulkan bahwa unjuk kerja berpotensi besar menumbuhkan berpikir kritis di SD, namun bukti empirisnya masih terfragmentasi sehingga kerangka konseptual pengaitan indikator berpikir kritis dengan rubrik unjuk kerja diajukan sebagai agenda pengembangan selanjutnya.

Kata kunci: Unjuk kerja, Penilaian autentik, Keterampilan berpikir kritis.



Copyright ©2026 Prosiding Konferensi Ilmiah Dasar

Published by Universitas PGRI Madiun. This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

PENDAHULUAN

Keterampilan berpikir kritis merupakan salah satu kompetensi inti abad ke-21 yang menjadi tumpuan kurikulum di hampir seluruh negara, termasuk Indonesia melalui kerangka Kurikulum Merdeka (Kemendikbudristek, 2022). Kompetensi ini tidak sekadar menuntut peserta didik menghafal informasi, tetapi mengolah, menafsirkan, dan mengevaluasi informasi untuk mengambil keputusan yang beralasan (Facione, 1990). OECD (2018) menempatkan *critical thinking* sebagai salah satu kompetensi transformatif yang menentukan kesiapan generasi muda menghadapi kompleksitas abad ke-21. Meski demikian, sejumlah studi lapangan di Indonesia justru melaporkan capaian keterampilan berpikir kritis peserta didik sekolah dasar (SD) masih berada pada kategori rendah hingga sedang. Rahmawati, Pujiastuti, dan Cahyaningtyas (2023) menemukan variasi capaian berpikir kritis siswa kelas empat SD yang timpang antarsekolah, sementara Juliyanika dan Batubara (2022) melalui pemetaan tren artikel pada jurnal-jurnal pendidikan dasar nasional mencatat bahwa penelitian tentang berpikir kritis di SD masih didominasi uji coba model pembelajaran dan relatif jarang menyentuh sisi instrumen penilaiannya.

Kesenjangan ini penting dicermati sebab penilaian bukan sekadar alat pelaporan hasil belajar, melainkan juga penggerak (*driver*) proses belajar itu sendiri. Ketika instrumen penilaian yang digunakan guru masih didominasi tes pilihan ganda atau isian singkat, proses pembelajaran cenderung terpusat pada pencapaian jawaban tunggal, bukan pada proses bernalar (Sudrajat, 2017). Wiggins (1998) telah lama mengingatkan bahwa penilaian konvensional semacam ini gagal menangkap unjuk kemampuan peserta didik dalam situasi nyata (*real-world performance*), sehingga diperlukan penilaian autentik yang menuntut peserta didik mendemonstrasikan pengetahuan dan keterampilannya secara langsung. Salah satu wujud penilaian autentik yang paling banyak direkomendasikan pada jenjang dasar dan menengah adalah unjuk kerja (*performance assessment*), yakni penilaian yang mengamati dan menilai proses maupun produk kinerja peserta didik berdasarkan kriteria rubrik yang telah ditetapkan (Kusnandar, 2016).

Secara konseptual, unjuk kerja memiliki keselarasan kuat dengan tuntutan pengukuran berpikir kritis. Taksonomi Facione (1990) menguraikan berpikir kritis kedalam enam indikator inti antara lain interpretasi, analisis, evaluasi, inferensi, eksplanasi, dan regulasi diri yang seluruhnya lebih mudah teramati, melalui unjuk kerja peserta didik saat memecahkan masalah, berargumentasi, atau mempresentasikan hasil kerja, dibandingkan melalui jawaban tertutup pada tes objektif. Fox, Freeman, Hughes, dan Murphy (2017) menegaskan bahwa penilaian autentik memberi ruang bagi peserta didik menunjukkan penalaran tingkat tinggi sekaligus memberi umpan balik formatif yang lebih kaya bagi guru. Sejumlah penelitian empiris di Indonesia turut mendukung klaim ini: Putri, Rosidin, dan Distrik (2020) melaporkan bahwa *performance assessment* yang dipadukan model *Project Based Learning* berkontribusi terhadap peningkatan keterampilan berpikir kritis dan kreatif siswa; sementara pada jenjang SMP, Chahnia dkk. (2024) serta Rahayu dan Ubabuddin (2023) melaporkan bahwa *performance assessment* efektif menilai keterampilan praktik sekaligus pemahaman konseptual peserta didik.

Urgensi kajian ini juga berkaitan erat dengan kebijakan asesmen dalam Kurikulum Merdeka, yang secara eksplisit mendorong guru menerapkan asesmen formatif dan autentik sebagai bagian dari pembelajaran berdiferensiasi (Kemendikbudristek, 2022). Namun, penerapan di lapangan masih menyisakan persoalan: guru SD kerap mengakui kesulitan menyusun rubrik unjuk kerja yang valid sekaligus efisien digunakan dalam kelas dengan jumlah siswa besar (Achmad, 2022; Nugroho dkk., 2021). Tanpa peta konseptual yang jelas tentang bagaimana indikator berpikir kritis dapat diterjemahkan menjadi kriteria unjuk kerja yang operasional, potensi unjuk kerja sebagai instrumen penilaian autentik berisiko tereduksi menjadi sekadar penilaian keterampilan motorik atau presentasi, tanpa benar-benar menyentuh dimensi bernalar peserta didik.

Berbeda dari kajian literatur sebelumnya yang umumnya berfokus pada satu mata pelajaran atau satu jenjang pendidikan, artikel ini secara khusus melakukan sintesis lintas mata pelajaran pada jenjang SD sekaligus membandingkannya dengan bukti dari jenjang SMP dan SMA untuk memetakan pola dan kesenjangan penelitian secara lebih utuh. Kontribusi artikel ini terletak pada tiga hal: (1) memetakan sejauh mana penelitian tentang unjuk kerja dan berpikir kritis pada jenjang SD sungguh-sungguh terintegrasi, bukan sekadar berdampingan; (2) menyusun kerangka konseptual keterkaitan indikator berpikir kritis Facione dengan komponen rubrik unjuk kerja yang dapat menjadi rujukan pengembangan instrumen; dan (3) merumuskan agenda penelitian lanjutan yang lebih terarah bagi peneliti pendidikan dasar.

METODE

Penelitian ini menggunakan metode studi literatur (*literature review*) dengan pendekatan naratif-tematik yang diperkaya prosedur penelusuran sistematis, mengadaptasi tahapan yang dikembangkan Triandini, Jayanatha, Indrawan, Putra, dan Iswara (2019), yaitu: (1) perumusan pertanyaan penelitian, (2) penyusunan strategi pencarian, (3) penerapan kriteria seleksi, (4) evaluasi dan analisis data, serta (5) interpretasi dan pelaporan temuan. Pertanyaan penelitian yang dirumuskan adalah: “Bagaimana unjuk kerja (*performance assessment*) berperan sebagai bentuk penilaian autentik dalam mengembangkan keterampilan berpikir kritis peserta didik sekolah dasar berdasarkan bukti pada literatur yang tersedia, dan pada aspek mana kesenjangan penelitian masih terbuka?”

Penelusuran literatur dilakukan pada basis data *Google Scholar*, Garuda (Garba Rujukan Digital), dan *ScienceDirect* menggunakan kombinasi kata kunci “unjuk kerja”, “*performance assessment*”, “penilaian autentik”, “*authentic assessment*”, “berpikir kritis”, “*critical thinking*”, dan “sekolah dasar”/“*elementary school*”. Penelusuran dibatasi pada artikel terbitan 2015–2025, dengan pengecualian untuk sejumlah rujukan konseptual berusia lebih lama (Facione, 1990; Wiggins, 1998) yang hingga kini masih menjadi rujukan standar di bidang asesmen pendidikan. Seleksi dilakukan dua tahap: tahap pertama menyaring berdasarkan kesesuaian judul dan abstrak sehingga menyisakan sekitar 32 artikel kandidat; tahap kedua menyaring berdasarkan teks penuh dengan kriteria inklusi (a) membahas unjuk kerja/penilaian autentik dan/atau keterampilan berpikir kritis, (b) berlatar jenjang pendidikan dasar, menengah, atau tinggi sebagai pembanding, dan (c) diterbitkan pada jurnal terindeks atau prosiding bereputasi. Proses ini menghasilkan 20 sumber sebagai basis sitasi, dengan 11 di antaranya yang paling relevan disintesis secara rinci pada matriks perbandingan.

Instrumen pengumpulan data berupa lembar pemetaan literatur (*literature mapping sheet*) yang memuat kolom peneliti/tahun, jenjang pendidikan dan mata pelajaran, metode penelitian, fokus kajian (unjuk kerja saja, berpikir kritis, atau terintegrasi), dan temuan utama. Analisis data dilakukan secara tematik dengan mengelompokkan temuan ke dalam tiga kategori: (1) jenjang pendidikan tempat penelitian dilakukan, (2) fokus kajian, dan (3) pola temuan terkait dampak unjuk kerja terhadap indikator berpikir kritis. Hasil pengelompokan selanjutnya divisualisasikan dalam diagram batang untuk memperlihatkan distribusi dan kesenjangan penelitian secara lebih terukur, sebelum dibahas secara kualitatif pada bagian pembahasan.

HASIL PENELITIAN

Berdasarkan prosedur seleksi yang dijelaskan pada bagian metode, 11 dari 20 sumber yang terkumpul memuat informasi jenjang pendidikan, metode, dan fokus kajian yang cukup lengkap untuk disintesis dalam matriks perbandingan. Tabel dibawah menyajikan hasil pemetaan tersebut.

TABEL 1. *Matriks Sintesis Studi tentang Unjuk Kerja dan Keterampilan Berpikir Kritis*

Peneliti	Jenjang & Mapel	Fokus Kajian	Metode	Temuan Utama
Putri, Rosidin, & Distrik (2020)	SMA – Fisika.	Unjuk kerja & berpikir kritis (terintegrasi)	Eksperimen (model PjBl)	Performance assessment berbasis PjBl dilaporkan berkontribusi pada peningkatan keterampilan berpikir kritis dan kreatif siswa
Nugroho, Yudha, Sundari, & Praja (2021)	SD – PJOK.	Unjuk kerja	Deskriptif-analitis	Instrumen asesmen unjuk kerja tergolong praktis, namun guru masih kesulitan menyusun kriteria penilaian yang rinci
Marselina, Lasmawan, & Dantes (2021)	SD kelas V – IPS.	Berpikir kritis	R&D pengembangan instrumen	Instrumen pengukuran berpikir kritis terbukti valid dan reliabel, namun masih berbasis tes tertulis, bukan unjuk kerja
Hapsari & Gularso (2024)	SD kelas IV – umum	Berpikir kritis	R&D (uji kepraktisan)	Instrumen penilaian berbasis berpikir kritis dinilai praktis oleh guru, tetapi belum dirancang dalam format unjuk kerja
Setyawan & Kristanti (2021)	SD – IPA	Berpikir kritis	Deskriptif kualitatif	Discovery learning meningkatkan keterampilan berpikir kritis; penilaian yang dipakai masih observasi umum, bukan rubrik unjuk kerja terstruktur
Rahayu Ubabuddin (2023)	& SMP – PAI	Unjuk kerja & berpikir kritis (parsial)	Deskriptif kualitatif	Penilaian unjuk kerja dan praktik meningkatkan pemahaman serta keterampilan performatif siswa dalam ibadah praktis
Erawati (2023)	SMA / promosi kesehatan	Unjuk kerja	R&D pengembangan instrumen	Instrumen unjuk kerja untuk promosi kesehatan dinyatakan valid dan layak digunakan
Chahnia (2024)	dkk. SMP (MTs) – PAI	Unjuk kerja & berpikir kritis	Kualitatif deskriptif (field research)	Performance assessment efektif menilai hafalan meningkatkan keterlibatan belajar siswa

TABEL 1 LANJUTAN. *Matriks Sintesis Studi tentang Unjuk Kerja dan Keterampilan Berpikir Kritis*

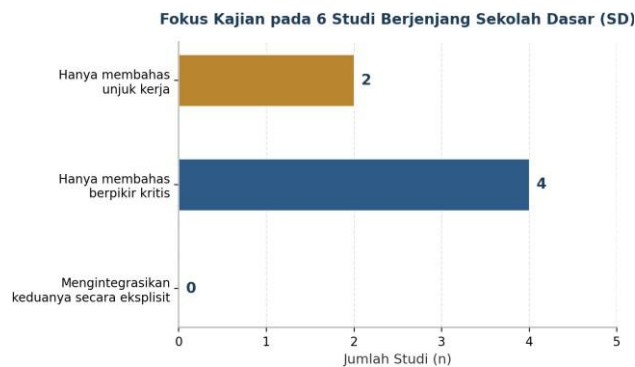
Peneliti	Jenjang & Mapel	Fokus Kajian	Metode	Temuan Utama
Achmad (2022)	SD – PAI	Unjuk kerja	Kualitatif studi kasus	Guru menerapkan penilaian autentik pada aspek kognitif, afektif, dan psikomotorik, tetapi mengalami kendala pemahaman prosedur penilaian
Fox, Freeman, Hughes, & Murphy (2017)	Perguruan tinggi (multidisiplin, internasional)	Perguruan tinggi (multidisiplin, internasional)	Tinjauan literatur (review)	Authentic assessment memberi ruang bernalar tingkat tinggi, namun implementasinya menghadapi tantangan beban penilaian dan reliabilitas antar-penilai
Rahmawati, Pujiastuti, & Cahyaningtyas (2023)	SD kelas IV	Berpikir kritis	Deskriptif kuantitatif	Capaian keterampilan berpikir kritis bervariasi antarsekolah, cenderung rendah-sedang, tanpa dikaitkan dengan bentuk unjuk kerja tertentu

Pemetaan pada table diatas memperlihatkan tiga pola menonjol. Pertama, dari sebelas studi, enam di antaranya berlatar jenjang SD, dua berlatar SMP, dan dua berlatar SMA, dan satu berlatar pendidikan tinggi internasional. Kedua, mata pelajaran yang paling banyak menerapkan unjuk kerja adalah mata pelajaran yang bermuatan praktik atau psikomotorik seperti PJOK dan Pendidikan Agama Islam, sementara mata pelajaran yang berorientasi literasi-numerasi (IPS, Matematika) lebih banyak diukur lewat instrumen tes berpikir kritis konvensional. Ketiga, dari sebelas studi tersebut hanya tiga (Putri dkk., 2020; Rahayu & Ubabuddin, 2023; Chahnia dkk., 2024) yang secara eksplisit memadukan unjuk kerja dengan pengukuran berpikir kritis dalam satu kajian yang sama, dan ketiganya berlatar jenjang SMP/SMA, bukan SD.

**GAMBAR 1.** *Distribusi Jenjang Pendidikan pada 11 Studi yang Direview*

Gambar 1 menegaskan secara visual bahwa jenjang SD justru menjadi jenjang dengan jumlah studi terbanyak yaitu enam studi dibandingkan SMP, SMA, maupun

perguruan tinggi. Temuan ini pada awalnya tampak bertentangan dengan asumsi umum bahwa penelitian tentang unjuk kerja di SD masih minim, namun ketika keenam studi tersebut ditelaah lebih jauh berdasarkan fokus kajiannya, pola yang muncul justru berbeda dari yang diharapkan.



GAMBAR 2. Distribusi Jenjang Pendidikan pada 11 Studi yang Direview

Gambar 2 memperlihatkan bahwa dari enam studi berjenjang SD, dua studi (Nugroho dkk., 2021; Achmad, 2022) hanya membahas unjuk kerja tanpa mengukur berpikir kritis secara eksplisit, empat studi (Marselina dkk., 2021; Hapsari & Gularso, 2024; Setyawan & Kristanti, 2021; Rahmawati dkk., 2023) hanya membahas berpikir kritis tanpa menempatkannya dalam kerangka unjuk kerja, dan tidak satu pun dari keenam studi tersebut mengintegrasikan kedua konstruk dalam satu instrumen yang sama. Dengan kata lain, meskipun secara kuantitas penelitian tentang SD cukup banyak ditemukan, secara kualitas keterkaitan tema, penelitian yang benar-benar menjawab judul “unjuk kerja untuk meningkatkan berpikir kritis siswa SD” justru belum ditemukan dalam korpus yang dikaji. Bukti dukungan terhadap tesis judul ini sejauh ini masih bersifat tidak langsung (*indirect evidence*): diinferensikan dari kombinasi teori (Facione, 1990; Wiggins, 1998), bukti dari jenjang di atasnya (Putri dkk., 2020; Chahniah dkk., 2024), dan bukti parsial mengenai efektivitas unjuk kerja maupun efektivitas pengukuran berpikir kritis yang berdiri sendiri di SD.

Pola ini juga sejalan dengan temuan Juliyantika dan Batubara (2022) yang memetakan tren artikel berpikir kritis pada jurnal-jurnal pendidikan dasar nasional, mayoritas artikel yang mereka telaah berangkat dari uji coba model atau strategi pembelajaran, sedangkan artikel yang berangkat dari pengembangan instrumen penilaian termasuk unjuk kerja jumlahnya jauh lebih sedikit. Dengan kata lain, kesenjangan yang ditemukan pada Tabel 1 dan Gambar 2 bukan sekadar kebetulan dari sebelas studi yang dikaji dalam artikel ini, melainkan mencerminkan kecenderungan yang lebih luas pada peta penelitian pendidikan dasar di Indonesia secara umum

PEMBAHASAN

Temuan di atas perlu dibaca bukan sebagai bantahan terhadap potensi unjuk kerja dalam menumbuhkan berpikir kritis, melainkan sebagai peta kesenjangan yang menjelaskan mengapa klaim tersebut meski masuk akal secara teoretis belum banyak diuji langsung pada peserta didik SD. Secara konseptual, keselarasan antara unjuk kerja dan berpikir kritis sesungguhnya kuat. Wiggins (1998) menekankan bahwa penilaian autentik menuntut peserta didik “melakukan” (*doing*) pengetahuannya, bukan sekadar mengenalinya di antara pilihan jawaban, sementara enam indikator berpikir kritis Facione (1990) interpretasi, analisis, evaluasi, inferensi, eksplanasi, dan regulasi diri ada dasarnya adalah proses yang paling mudah teramati ketika peserta didik mengerjakan tugas nyata bukan tugas tertutup. Kekuatan konseptual inilah yang tampaknya mendorong peneliti

pada jenjang SMP dan SMA (Putri dkk., 2020; Chahnia dkk., 2024) untuk secara sengaja memadukan keduanya. Namun demikian, temuan pada Gambar 2 menunjukkan bahwa pengaitan serupa belum banyak dilakukan pada jenjang SD, sehingga literatur yang ada belum cukup untuk memastikan bahwa hubungan tersebut berlaku sama kuatnya pada peserta didik yang secara perkembangan kognitif berada pada tahap operasional konkret.

Justru pada titik inilah unjuk kerja diperkirakan memiliki relevansi yang lebih tinggi dibandingkan tes tertulis untuk peserta didik SD. Anak usia sekolah dasar cenderung lebih mudah menunjukkan penalarannya melalui aktivitas konkret mendemonstrasikan langkah percobaan, menjelaskan alasan di balik sebuah keputusan dalam permainan kelompok, atau mempresentasikan hasil pengamatan dibandingkan mengartikulasikannya dalam bentuk tulisan abstrak. Argumen ini konsisten dengan temuan Setyawan dan Kristanti (2021) serta Marselina dkk. (2021) yang melaporkan peningkatan indikator berpikir kritis melalui aktivitas belajar berbasis pengalaman langsung, meskipun keduanya belum menggunakan istilah maupun instrumen “unjuk kerja” secara formal. Dengan demikian, dapat dikatakan bahwa praktik pembelajaran di lapangan sesungguhnya telah bergerak ke arah yang selaras dengan gagasan unjuk kerja, tetapi sisi instrumentasi penilaiannya belum menyusul.

Kesenjangan instrumentasi ini juga tampak dari sisi praktis. Nugroho dkk. (2021) dan Achmad (2022) dua studi SD yang secara eksplisit membahas unjuk kerja sama-sama melaporkan bahwa guru mengalami kesulitan menyusun kriteria atau rubrik penilaian yang rinci sekaligus efisien digunakan pada kelas dengan jumlah siswa besar. Kendala serupa turut disinggung Fox, Freeman, Hughes, dan Murphy (2017) dalam konteks pendidikan tinggi: penilaian autentik memang memperkaya bukti penalaran peserta didik, tetapi menimbulkan beban penilaian yang lebih besar bagi pengajar serta risiko rendahnya reliabilitas antarpemilai (*inter-rater reliability*) jika rubrik tidak dirancang secara cermat. Bagi konteks SD, tantangan ini berpotensi lebih besar karena guru kelas umumnya mengampu banyak mata pelajaran sekaligus dan belum tentu memperoleh pelatihan khusus mengenai penyusunan rubrik berpikir kritis.

Untuk menjembatani kesenjangan konseptual-instrumen tersebut, artikel ini mengajukan kerangka penskoran unjuk kerja yang secara eksplisit mengintegrasikan indikator berpikir kritis Facione (1990) sebagai agenda pengembangan lanjutan. Skor unjuk kerja secara umum dapat dirumuskan sebagai berikut

$$NUK = (\sum Si / Smaks) \times 100.....(1)$$

NUK adalah nilai unjuk kerja, Si adalah skor capaian indikator ke-i yang diberikan penilai berdasarkan rubrik, dan Smaks adalah skor maksimal seluruh indikator yang dinilai. Formula ini lazim digunakan dalam pengembangan instrumen unjuk kerja (Kusnandar, 2016), tetapi indikator Si di dalamnya umumnya masih berupa aspek keterampilan motorik atau prosedural, belum menyentuh dimensi bernalar. Untuk mengakomodasi dimensi berpikir kritis, formula tersebut dapat diperluas menjadi skor unjuk kerja berbobot berpikir kritis sebagai berikut.

$$NUK-BK = \sum (Wj \times Sj) / \sum Wj.....(2)$$

Wj adalah bobot indikator berpikir kritis ke-j (interpretasi, analisis, evaluasi, inferensi, eksplanasi, dan regulasi diri) yang disesuaikan dengan tingkat perkembangan kognitif peserta didik SD misalnya memberi bobot lebih besar pada interpretasi dan inferensi konkret, serta bobot lebih kecil pada regulasi diri yang menuntut metakognisi lebih matang dan Sj adalah skor capaian indikator tersebut yang teramati selama unjuk kerja berlangsung. Persamaan ini bersifat konseptual dan belum diuji secara empiris pada populasi SD diajukan sebagai kerangka awal yang dapat divalidasi melalui penelitian pengembangan (R&D) maupun pendekatan psikometrik seperti pemodelan *Rasch*

(Sumintono & Widhiarso, 2013), yang telah terbukti relevan untuk menguji validitas instrumen asesmen pendidikan di Indonesia.

Kontribusi konseptual ini juga relevan dengan arah kebijakan Kurikulum Merdeka yang mendorong asesmen formatif dan berdiferensiasi (Kemendikbudristek, 2022). Tren penelitian yang dipetakan Juliyantika dan Batubara (2022) menunjukkan bahwa literatur pendidikan dasar nasional masih condong pada uji coba model pembelajaran untuk menaikkan skor berpikir kritis, sementara pengembangan instrumen penilaiannya tertinggal. Bila dibiarkan, kesenjangan ini berisiko melanggengkan situasi di mana guru menerapkan model pembelajaran yang mendorong bernalar tingkat tinggi, tetapi tetap menilainya dengan instrumen tes tertulis konvensional yang justru tidak selaras dengan proses belajarnya sendiri sebuah inkonsistensi antara praktik pembelajaran dan praktik penilaian yang oleh Wiggins (1998) disebut sebagai kegagalan validitas konsekuensial (*consequential validity*) asesmen.

Dibandingkan dengan kajian literatur sejenis yang cenderung berfokus pada satu mata pelajaran (misalnya hanya PJOK sebagaimana Nugroho dkk., 2021, atau hanya IPA sebagaimana Setyawan & Kristanti, 2021), sintesis lintas mata pelajaran dan lintas jenjang pada artikel ini memungkinkan teridentifikasinya pola yang tidak tampak bila hanya satu bidang studi yang dikaji, yaitu bahwa persoalan utamanya bukan pada ketiadaan unjuk kerja maupun ketiadaan pengukuran berpikir kritis di SD, melainkan pada belum bertemunya keduanya dalam satu instrumen. Perbedaan sudut pandang inilah yang menjadi nilai tambah (*added value*) artikel ini dibandingkan tinjauan literatur yang sudah ada, sekaligus menjadi dasar argumentasi bahwa agenda penelitian ke depan semestinya diarahkan pada pengembangan dan validasi instrumen, bukan sekadar replikasi uji coba model pembelajaran baru.

Implikasi praktis dari peta kesenjangan ini juga perlu disampaikan secara proporsional kepada guru di lapangan. Selama ini, pelatihan penilaian autentik bagi guru SD umumnya masih bersifat umum seperti mengenalkan definisi dan jenis-jenis penilaian autentik tanpa contoh rubrik konkret yang secara eksplisit menyasar indikator berpikir kritis. Nugroho dkk. (2021) menunjukkan bahwa instrumen unjuk kerja yang sudah ada di lapangan pun cenderung berhenti pada penilaian ketepatan gerak atau langkah kerja, bukan pada kualitas penalaran di baliknya. Bila kerangka pada Persamaan (2) dikembangkan lebih lanjut menjadi rubrik yang disertai deskriptor perilaku konkret untuk setiap indikator berpikir kritis misalnya deskriptor “mampu menjelaskan alasan pemilihan strategi” untuk indikator eksplanasi, atau “mampu membandingkan dua kemungkinan solusi sebelum memutuskan” untuk indikator evaluasi guru berpeluang lebih besar menilai berpikir kritis tanpa harus menambah instrumen baru di luar unjuk kerja yang sudah lazim mereka gunakan.

Perbandingan dengan konteks internasional turut memperkuat argumen ini. Fox dkk. (2017) mencatat bahwa keberhasilan penilaian autentik di berbagai negara umumnya ditopang oleh ketersediaan bank rubrik yang telah divalidasi dan dapat diadaptasi lintas mata pelajaran, bukan disusun ulang oleh setiap guru secara mandiri. Kondisi ini belum banyak ditemukan pada konteks SD di Indonesia, sehingga pengembangan bank rubrik unjuk kerja berbasis indikator berpikir kritis yang teruji secara psikometrik dan dapat diadaptasi lintas mata pelajaran dapat menjadi kontribusi terapan yang menjembatani kesenjangan antara idealisme kebijakan Kurikulum Merdeka dan kesiapan guru di lapangan.

SIMPULAN

Sintesis terhadap 11 studi inti menunjukkan bahwa unjuk kerja (*performance assessment*) memiliki landasan konseptual yang kuat sebagai bentuk penilaian autentik yang selaras dengan indikator berpikir kritis Facione (1990), dan bukti pada jenjang SMP-SMA memperlihatkan hasil yang konsisten mendukung peran tersebut. Akan tetapi, pada jenjang SD, dukungan empirisnya masih bersifat tidak langsung: enam studi SD yang

ditemukan terbelah antara yang membahas unjuk kerja saja atau berpikir kritis saja, tanpa satu pun yang memadukan keduanya secara eksplisit. Dengan demikian, klaim bahwa unjuk kerja “meningkatkan” keterampilan berpikir kritis peserta didik SD lebih tepat dipahami sebagai proposisi teoretis yang didukung secara tidak langsung oleh bukti dari jenjang lain, bukan sebagai simpulan yang telah teruji langsung pada populasi SD.

Artikel ini turut mengajukan kerangka penskoran unjuk kerja berbobot berpikir kritis (Persamaan 2) sebagai kontribusi konseptual yang dapat menjadi titik tolak pengembangan instrumen pada penelitian selanjutnya. Beberapa keterbatasan perlu diakui: korpus literatur yang dikaji berjumlah relatif kecil dan didominasi sumber berbahasa Indonesia, proses seleksi belum mengikuti protokol PRISMA secara penuh, dan pengelompokan fokus kajian pada Tabel 1 tetap mengandung unsur interpretasi peneliti yang berpotensi bias. Oleh karena itu, tiga agenda penelitian lanjutan direkomendasikan: pertama, pengembangan dan validasi psikometrik rubrik unjuk kerja yang secara eksplisit memuat indikator berpikir kritis untuk peserta didik SD lintas mata pelajaran; kedua, penelitian eksperimen atau kuasi-eksperimen yang menguji langsung efektivitas rubrik tersebut dibandingkan penilaian konvensional; dan ketiga, meta-analisis kuantitatif yang lebih luas dan mengikuti protokol sistematis penuh (mis. PRISMA) untuk menguatkan generalisasi temuan yang pada artikel ini masih bersifat pemetaan awal.

DAFTAR PUSTAKA

1. Achmad, G. H. (2022). Penilaian autentik pada Kurikulum Merdeka Belajar dalam pembelajaran Pendidikan Agama Islam di sekolah dasar. *Edukatif: Jurnal Ilmu Pendidikan*, 4(3).
2. Chahnia, J., dkk. (2024). Penggunaan performance assessment sebagai instrumen penilaian pada mata pelajaran Al-Qur'an Hadis kelas VIII di MTs-S Thawalib Padang. *Jurnal Kewarganegaraan*, 8(2).
3. Erawati. (2023). Pengembangan instrumen penilaian unjuk kerja (performance assessment) dalam melaksanakan promosi kesehatan. *Edukatif: Jurnal Ilmu Pendidikan*, 5(1).
4. Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. The California Academic Press.
5. Fox, J., Freeman, S., Hughes, N., & Murphy, V. (2017). Keeping it real: A review of the benefits, challenges and steps towards implementing authentic assessment. *Journal of Teaching and Learning in Higher Education*, 9(3), 2801–2813.
6. Hapsari, Y. P., & Gularso, D. (2024). Pengembangan kepraktisan instrumen penilaian berbasis keterampilan berpikir kritis siswa kelas IV SD Negeri Wirosaban. *Jurnal Cendekia: Jurnal Pendidikan Matematika*, 8(1).
7. Juliyantika, T., & Batubara, H. H. (2022). Tren penelitian keterampilan berpikir kritis pada jurnal pendidikan dasar di Indonesia. *Jurnal Basicedu*, 6(3), 4731–4744. <https://doi.org/10.31004/basicedu.v6i3.2869>
8. Kemendikbudristek. (2022). *Panduan pembelajaran IPAS untuk SD/MI Kurikulum Merdeka*. Pusat Kurikulum dan Perbukuan.
9. Kusnandar. (2016). *Penilaian autentik (penilaian hasil belajar peserta didik berdasarkan Kurikulum 2013)*. Rajagrafindo Persada.
10. Marselina, K. T., Lasmawan, I. W., & Dantes, N. (2021). Pengembangan instrumen kemampuan berpikir kritis dan hasil belajar IPS pada siswa kelas V SD. *Jurnal Penelitian dan Evaluasi Pendidikan Indonesia*, 11(2), 105–114. <https://doi.org/10.23887/jpepi.v11i2.620>

11. Nugroho, W. A., Yudha, R. P., Sundari, S., & Praja, H. N. (2021). Analisis instrumen asesmen unjuk kerja pada pembelajaran PJOK di sekolah dasar Kota Cirebon. *Gelombang Olahraga: Jurnal Pendidikan Jasmani dan Olahraga*, 4(2), 126–141. <https://doi.org/10.31539/jpjo.v4i2.1795>
12. OECD. (2018). *The future of education and skills: Education 2030*. OECD Publishing.
13. Putri, N. S. Y., Rosidin, U., & Distrik, I. W. (2020). Pengaruh penerapan performance assessment dengan model PJBL terhadap keterampilan berpikir kritis dan kreatif siswa SMA. *Jurnal Pendidikan Fisika*, 8(1), 58–69.
14. Rahayu, P., & Ubabuddin, U. (2023). Penilaian unjuk kerja dan praktik dalam pembelajaran PAI. *JIP: Jurnal Ilmu Pendidikan*, 1(2), 304–313.
15. Rahmawati, H., Pujiastuti, P., & Cahyaningtyas, A. P. (2023). Kategorisasi kemampuan berpikir kritis siswa kelas empat sekolah dasar di SD se-Gugus II Kapanewon Playen, Gunung Kidul. *Jurnal Pendidikan dan Kebudayaan*, 8(1).
16. Setyawan, R. A., & Kristanti, H. S. (2021). Keterampilan berpikir kritis pada pembelajaran IPA melalui model pembelajaran discovery learning bagi siswa sekolah dasar. *Jurnal Basicedu*, 5(2), 1076–1082.
17. Sudrajat, A. (2017). Penilaian autentik dalam pembelajaran berbasis kompetensi. *Jurnal Pendidikan Dasar*, 9(1).
18. Sumintono, B., & Widhiarso, W. (2013). *Aplikasi pemodelan Rasch pada assessment pendidikan*. Trim Komunikata Publishing House.
19. Triandini, E., Jayanatha, S., Indrawan, A., Putra, G. W., & Iswara, B. (2019). Metode systematic literature review untuk identifikasi platform dan metode pengembangan sistem informasi di Indonesia. *Indonesian Journal of Information Systems*, 1(2), 63–77.
20. Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. Jossey-Bass.