



Stimulus Fenomena Lokal sebagai Asesmen Diagnostik HOTS IPA Terpadu

Katarina Herwanti^{1)*}, Saptono Nugrohad²⁾, Stanislaus Christo Petra Nugraha³⁾,
Dail Umamil Asri⁴⁾

¹⁾PGSD, Universitas Negeri Semarang

²⁾SMA Negeri 3 Salatiga

³⁾University of Dundee Scotland

⁴⁾FIB, Universitas Diponegoro

Correspondence: katarinaherwanti@mail.unnes.ac.id

ABSTRAK

Penelitian ini bertujuan menganalisis fungsi stimulus fenomena lokal sebagai asesmen diagnostik Higher Order Thinking Skills (HOTS) dalam pembelajaran IPA terpadu. Penelitian menggunakan pendekatan mixed methods dengan desain explanatory sequential. Subjek penelitian melibatkan 592 peserta didik pada pembelajaran IPA. Instrumen penelitian berupa asesmen HOTS berbasis stimulus fenomena lokal yang menuntut penalaran inferensial, pemahaman proses, dan klasifikasi konsep. Data kuantitatif dianalisis secara deskriptif untuk menggambarkan capaian HOTS, sedangkan data kualitatif dianalisis melalui pola pilihan jawaban dan distraktor untuk mengidentifikasi miskonsepsi konseptual. Hasil penelitian menunjukkan bahwa sebagian besar peserta didik mampu menarik inferensi kausal dan memahami proses fenomena alam yang disajikan. Namun demikian, ditemukan miskonsepsi konseptual yang signifikan pada aspek klasifikasi konsep, meskipun pemahaman proses relatif baik. Temuan ini mengindikasikan adanya ketidaksinkronan antara penalaran ilmiah dan pelabelan konsep. Penelitian ini menyimpulkan bahwa stimulus fenomena lokal efektif digunakan sebagai asesmen diagnostik HOTS IPA terpadu untuk mengungkap capaian penalaran sekaligus miskonsepsi konseptual peserta didik dalam pembelajaran IPA.

Kata Kunci: HOTS IPA; Fenomena Lokal; Asesmen Diagnostik; Miskonsepsi Konsep; Pembelajaran IPA

This is an open access article under the CC - BY license.



PENDAHULUAN

Pembelajaran Ilmu Pengetahuan Alam (IPA) pada abad ke-21 tidak lagi berfokus semata pada penguasaan konsep, tetapi pada pengembangan kemampuan berpikir tingkat tinggi (Higher Order Thinking Skills / HOTS) yang memungkinkan peserta didik menganalisis fenomena, mengevaluasi bukti, serta menarik inferensi berbasis penalaran ilmiah. Dalam konteks literasi sains kontemporer, HOTS diposisikan sebagai komponen inti yang menghubungkan pemahaman konsep dengan kemampuan menggunakan pengetahuan tersebut dalam konteks nyata (Lahagu dkk., 2024; Ploj Virtič, 2022). Oleh karena itu, asesmen IPA perlu dirancang tidak hanya untuk mengukur ketepatan jawaban, tetapi juga kualitas proses berpikir ilmiah yang mendasarinya.

Namun demikian, berbagai kajian mutakhir menunjukkan bahwa praktik asesmen sains di sekolah masih cenderung berorientasi pada pengukuran capaian faktual dan prosedural, sehingga kurang optimal dalam mengungkap proses penalaran ilmiah dan struktur konseptual peserta didik (Elshall & Badir, 2025). Studi terkini menegaskan bahwa desain asesmen yang tidak secara eksplisit mengintegrasikan tuntutan berpikir tingkat tinggi (higher-order thinking skills/HOTS) berisiko hanya merekam performa permukaan tanpa menangkap kualitas argumentasi dan pengambilan keputusan konseptual (Lian, 2023; OECD, 2023; Xiao dkk., 2025). Kondisi ini memungkinkan peserta didik memperoleh skor yang relatif baik, tetapi masih menyimpan miskonsepsi mendasar yang tidak terdeteksi melalui asesmen konvensional berbasis jawaban tunggal.

Dalam konteks geosains, tantangan tersebut semakin kompleks karena banyak konsep melibatkan proses dinamis, spasial, dan temporal yang tidak dapat diamati secara langsung. Penelitian terbaru menunjukkan bahwa peserta didik sering mengalami kesulitan membedakan fenomena yang tampak serupa secara visual namun berbeda secara mekanisme prosesnya, sehingga memunculkan alternatif konsepsi yang persisten, khususnya pada

topik tektonik lempeng dan sistem kebumian (Conrad & Libarkin, 2022; Pallant dkk., 2024). Hal ini menegaskan pentingnya asesmen yang tidak hanya menilai ketepatan jawaban akhir, tetapi juga mampu mengidentifikasi pola penalaran dan dasar konseptual yang melandasi respons peserta didik.

Pemanfaatan fenomena lokal sebagai stimulus pembelajaran dan asesmen menawarkan pendekatan yang lebih kontekstual dan bermakna. Kajian terkini menunjukkan bahwa konteks lokal atau place-based science learning dapat meningkatkan keterhubungan konseptual, relevansi personal, serta mendorong penalaran inferensial dan evaluatif dalam pembelajaran sains (Holmes dkk., 2021). Dalam kerangka asesmen, stimulus berbasis fenomena autentik dapat difungsikan sebagai alat diagnostik untuk mengungkap pemahaman proses dan miskonsepsi melalui analisis respons, termasuk efektivitas distraktor yang dirancang secara konseptual (Haladyna & Rodriguez, 2021; Rezigalla dkk., 2024a). Dengan demikian, asesmen tidak lagi semata-mata berfungsi sebagai alat pengukuran hasil, melainkan sebagai instrumen diagnosis struktur berpikir ilmiah.

Meskipun penelitian mengenai HOTS dan literasi sains terus berkembang, pemanfaatan fenomena lokal dalam asesmen masih lebih sering ditempatkan sebagai konteks pembelajaran daripada sebagai perangkat diagnostik yang secara sistematis memetakan struktur penalaran dan konsistensi konseptual peserta didik. Sebagian besar studi berfokus pada peningkatan capaian skor atau level literasi sains (Fortus dkk., 2022; Ploj Virtič, 2022), sementara hubungan antara kemampuan penalaran inferensial dan ketepatan klasifikasi konsep jarang dianalisis secara terintegrasi. Padahal, ketidaksinkronan antara keduanya berpotensi menyembunyikan miskonsepsi yang tidak terdeteksi melalui asesmen konvensional.

Berdasarkan celah tersebut, penelitian ini menempatkan stimulus fenomena lokal sebagai instrumen asesmen diagnostik HOTS IPA terpadu untuk menganalisis keterkaitan antara penalaran inferensial, klasifikasi konsep, dan pemahaman proses geologi. Berbeda dari penelitian sebelumnya yang menekankan capaian performa, studi ini menyoroti konsistensi struktur pengetahuan peserta didik sebagai indikator kualitas berpikir ilmiah. Dengan demikian, kontribusi penelitian ini terletak pada integrasi fungsi evaluatif dan diagnostik dalam satu desain asesmen kontekstual yang mampu mengungkap capaian sekaligus potensi miskonsepsi secara lebih komprehensif.

METODE

Penelitian ini menggunakan pendekatan mixed methods dengan desain explanatory sequential, yaitu pengumpulan dan analisis data kuantitatif pada tahap awal yang diikuti dengan analisis kualitatif untuk menjelaskan dan memperdalam temuan kuantitatif (Venkatesh dkk., 2024). Desain ini dipilih karena penelitian tidak hanya bertujuan menggambarkan capaian HOTS peserta didik, tetapi juga mengungkap proses berpikir dan miskonsepsi konseptual yang tercermin dari pola respons asesmen.

Subjek penelitian melibatkan 592 peserta didik yang mengikuti pembelajaran IPA terpadu. Pemilihan subjek dilakukan secara purposive, dengan pertimbangan bahwa peserta didik telah memperoleh materi dasar terkait fenomena alam dan dinamika bumi dalam kurikulum IPA. Penelitian dilaksanakan pada konteks pembelajaran reguler sehingga data yang diperoleh merepresentasikan kondisi autentik kelas.

Instrumen penelitian berupa asesmen HOTS IPA terpadu yang dikembangkan berbasis stimulus fenomena lokal. Asesmen dirancang untuk mengukur kemampuan penalaran inferensial, pemahaman proses, dan klasifikasi konsep pada level analisis dan evaluasi sesuai taksonomi Bloom revisi (Hingstman dkk., 2023; Lieberei dkk., 2025; Rezigalla dkk., 2024b). Distraktor disusun secara konseptual untuk merepresentasikan miskonsepsi umum sehingga respons peserta didik dapat digunakan sebagai data diagnostik proses berpikir (Haladyna & Rodriguez, 2021).

Pengumpulan data dilakukan menggunakan Google Form yang dibagikan pada saat pembelajaran reguler. Peserta didik mengerjakan asesmen secara mandiri tanpa umpan balik langsung agar respons mencerminkan pemahaman aktual.

Analisis data kuantitatif dilakukan menggunakan statistik deskriptif (rerata, median, rentang, dan distribusi skor) untuk menggambarkan capaian HOTS peserta didik. Analisis kualitatif dilakukan melalui analisis pola pilihan jawaban dan distraktor guna mengidentifikasi miskonsepsi konseptual dan ketidaksinkronan antara penalaran dan klasifikasi konsep. Integrasi hasil kuantitatif dan kualitatif dilakukan

pada tahap interpretasi sesuai karakteristik desain explanatory sequential (Rezigalla dkk., 2024a; Sajjad dkk., 2020).

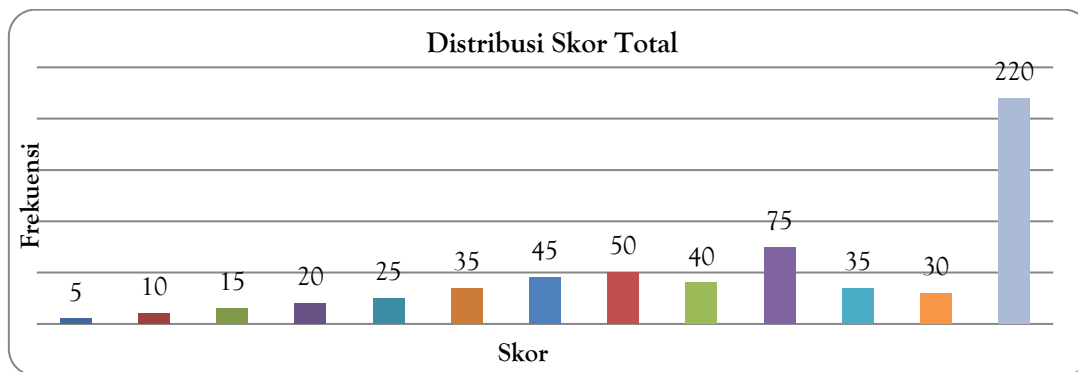
HASIL DAN PEMBAHASAN

Hasil penelitian ini menyajikan analisis terintegrasi antara capaian kuantitatif peserta didik dan interpretasi kualitatif respons asesmen Higher Order Thinking Skills (HOTS) berbasis stimulus fenomena lokal. Penyajian hasil difokuskan pada distribusi skor, capaian indikator penalaran, serta pola miskonsepsi konseptual yang terungkap melalui analisis respons dan distraktor

Capaian HOTS Peserta Didik pada Asesmen IPA Terpadu

Secara keseluruhan, hasil asesmen menunjukkan bahwa capaian peserta didik berada pada kategori menengah-tinggi. Dari skor maksimum instrumen sebesar 16 poin, skor aktual peserta didik berada pada rentang 0–12 dengan nilai rata-rata 9,08 dan median 9. Informasi statistik ringkas ini menunjukkan bahwa sebagian besar peserta didik mampu menjawab lebih dari separuh tuntutan asesmen.

Distribusi skor peserta didik disajikan pada Gambar 1. Grafik tersebut menunjukkan distribusi yang relatif simetris dengan konsentrasi tertinggi pada rentang skor 8–10. Tidak tampak penumpukan ekstrem pada skor sangat rendah maupun sangat tinggi, sehingga dapat disimpulkan bahwa instrumen tidak terlalu mudah dan tidak terlalu sulit.

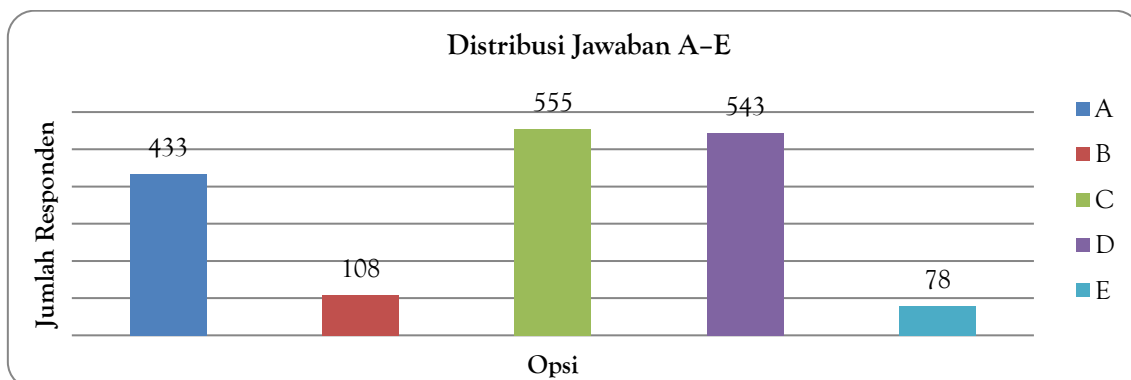


Gambar 1. Distribusi Skor Total HOTS Peserta Didik

Pola distribusi ini mengindikasikan bahwa instrumen mampu menjangkau spektrum kemampuan peserta didik secara memadai. Dalam konteks asesmen HOTS, capaian skor yang relatif tinggi tidak dapat ditafsirkan secara sederhana sebagai rendahnya tuntutan kognitif, karena kualitas asesmen HOTS ditentukan oleh jenis proses berpikir yang diukur, bukan oleh proporsi jawaban benar semata (Sajjad dkk., 2020).

Penalaran Inferensial terhadap Fenomena Alam

Capaian peserta didik pada indikator penalaran inferensial ditunjukkan pada Gambar 2, yang menyajikan distribusi pilihan jawaban pada butir yang mengukur kemampuan peserta didik dalam menarik inferensi penyebab terjadinya fenomena alam berdasarkan stimulus yang diberikan.



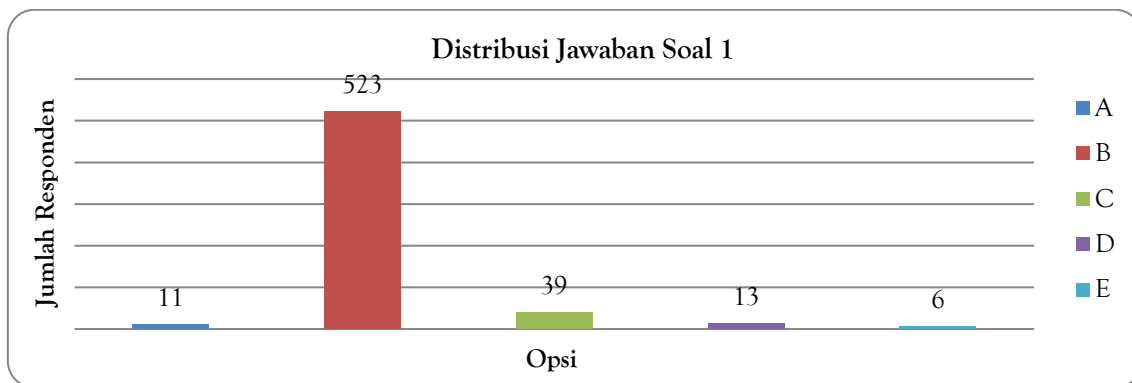
Gambar 2. Respon Peserta Didik pada Indikator Penalaran Inferensial Proses Fenomena

Grafik tersebut menunjukkan bahwa mayoritas peserta didik memilih jawaban yang merepresentasikan pemahaman bahwa fenomena yang disajikan dipicu oleh mekanisme proses tertentu, bukan oleh asumsi umum atau penjelasan berbasis ciri permukaan. Persentase jawaban benar yang sangat dominan menunjukkan bahwa peserta didik mampu mengintegrasikan informasi eksplisit dan implisit dalam stimulus untuk membangun hubungan sebab-akibat secara ilmiah.

Temuan ini memperkuat pandangan bahwa stimulus fenomena lokal efektif dalam memfasilitasi penalaran inferensial, yaitu kemampuan inti dalam HOTS IPA yang menuntut analisis dan evaluasi informasi, bukan sekadar reproduksi pengetahuan (Rahayu & Rosawati, 2023).

Pemahaman Hubungan Proses Alam dan Pemanfaatan oleh Manusia

Selain penalaran inferensial, asesmen juga mengukur kemampuan peserta didik dalam mengaitkan proses fenomena alam dengan konteks pemanfaatannya dalam kehidupan manusia. Hasil respons peserta didik pada indikator ini ditampilkan pada Gambar 3.



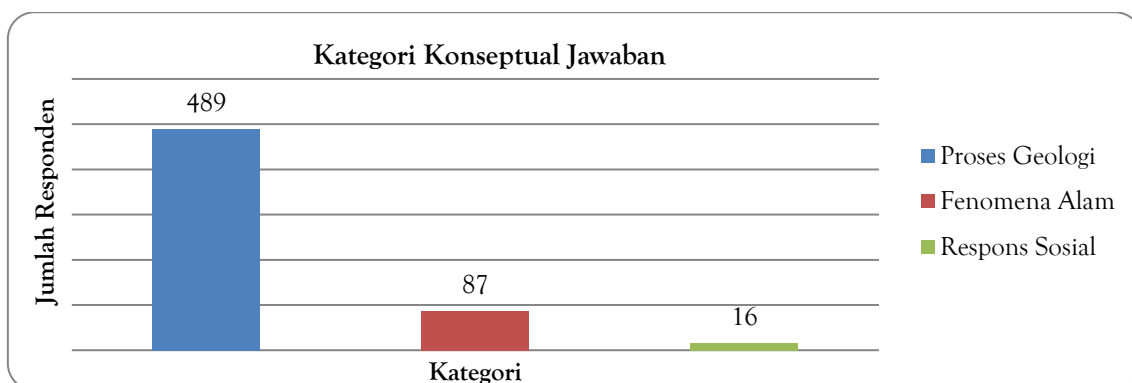
Gambar 3. Respons Peserta Didik pada Indikator Hubungan Proses Alam dan Pemanfaatan Manusia

Grafik menunjukkan bahwa sebagian besar peserta didik mampu menghubungkan karakteristik proses alam dengan bentuk pemanfaatan yang sesuai. Tingginya proporsi jawaban tepat mengindikasikan bahwa peserta didik tidak hanya memahami proses alam secara terpisah, tetapi juga mampu mengaitkannya dengan konteks kehidupan nyata.

Hasil ini menunjukkan bahwa asesmen berbasis fenomena lokal berpotensi memperkuat literasi sains kontekstual, karena peserta didik didorong untuk melihat keterkaitan antara konsep IPA dan realitas sosial di sekitarnya (Holmes dkk., 2021).

Miskonsepsi Konseptual pada Klasifikasi Fenomena

Berbeda dengan capaian tinggi pada pemahaman proses, hasil yang kontras muncul pada indikator klasifikasi konsep. Respons peserta didik pada indikator ini disajikan pada Gambar 4.



Gambar 4. Perbandingan Ketepatan Klasifikasi Konsep Fenomena

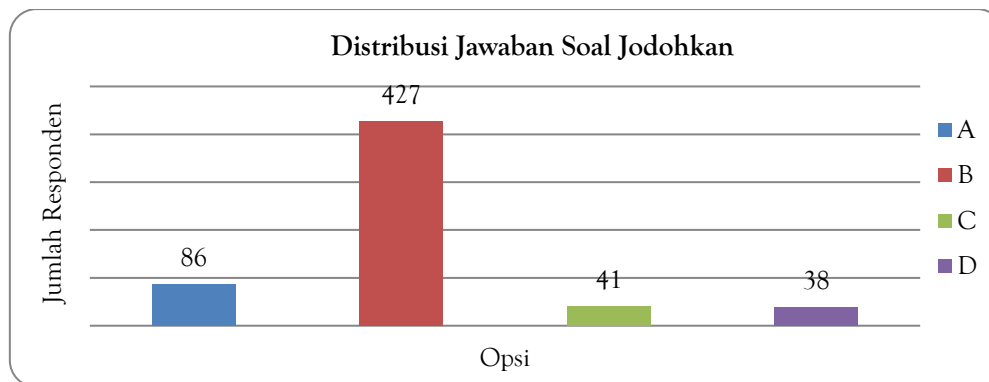
Grafik tersebut menunjukkan bahwa meskipun peserta didik mampu menjelaskan proses terjadinya fenomena secara ilmiah, sebagian besar masih melakukan kesalahan dalam mengklasifikasikan fenomena

tersebut. Peserta didik cenderung mengelompokkan fenomena berdasarkan kemiripan visual atau istilah populer, bukan berdasarkan mekanisme proses yang telah mereka pahami.

Temuan ini menunjukkan adanya ketidaksinkronan antara pemahaman proses dan pelabelan konsep. Kondisi ini sejalan dengan temuan penelitian pendidikan sains yang menyatakan bahwa miskonsepsi konseptual sering bersifat persisten dan berakar pada skema pengetahuan intuitif yang telah lama terbentuk (Conrad & Libarkin, 2022).

Analisis Distraktor sebagai Diagnostik Miskonsepsi

Fungsi diagnostik asesmen semakin terlihat melalui analisis distraktor. Pola pilihan jawaban peserta didik pada distraktor disajikan pada Gambar 5.



Gambar 5. Pola Pilihan Distraktor dan Indikasi Miskonsepsi Konseptual

Grafik menunjukkan bahwa distraktor yang paling banyak dipilih merupakan opsi yang secara konseptual tampak masuk akal karena selaras dengan intuisi atau pengalaman sehari-hari, tetapi tidak sesuai dengan mekanisme ilmiah yang dijelaskan dalam stimulus. Sebaliknya, distraktor yang tidak memiliki keterkaitan logis dengan konteks fenomena hanya dipilih oleh sebagian kecil peserta didik. Pola ini mengindikasikan bahwa kesalahan yang muncul bukan hasil tebakan acak, melainkan refleksi dari skema konseptual tertentu yang telah lebih dahulu terbentuk.

Konsistensi pilihan pada distraktor tertentu memperlihatkan bahwa peserta didik cenderung menggunakan kategori intuitif dalam menginterpretasikan fenomena, meskipun sebelumnya mampu menjelaskan proses kausal secara benar. Hal ini menunjukkan adanya pemisahan antara pemahaman mekanisme dan struktur klasifikasi dalam representasi pengetahuan mereka. Temuan tersebut memperkuat pandangan bahwa distraktor yang dirancang berbasis miskonsepsi umum dapat berfungsi sebagai alat diagnostik yang efektif untuk mengidentifikasi pola penalaran dan kesalahan konseptual secara lebih mendalam (Rezigalla dkk., 2024a).

Dengan demikian, analisis distraktor dalam penelitian ini tidak hanya berperan sebagai pelengkap butir pilihan ganda, tetapi sebagai sumber informasi penting untuk memetakan struktur berpikir peserta didik. Informasi ini relevan bagi guru dalam merancang intervensi pembelajaran yang menargetkan miskonsepsi yang bersifat sistematis.

Sintesis Hasil dan Makna Pedagogis

Temuan penelitian ini menunjukkan bahwa asesmen HOTS berbasis stimulus fenomena lokal tidak hanya merefleksikan tingkat capaian kognitif, tetapi juga mengungkap koherensi struktur konseptual peserta didik. Integrasi antara distribusi skor, capaian penalaran inferensial, ketepatan klasifikasi konsep, dan pola distraktor memperlihatkan bahwa kemampuan menjelaskan proses ilmiah tidak selalu sejalan dengan ketepatan pelabelan konseptual. Dengan kata lain, penalaran kausal dan organisasi konsep merupakan dua dimensi yang dapat berkembang secara tidak paralel.

Implikasi pedagogisnya terletak pada perlunya asesmen yang mampu mengidentifikasi ketidaksinkronan tersebut sejak awal. Informasi dari pola respons dan distraktor memberikan dasar yang lebih diagnostik dibanding sekadar skor akhir, sehingga guru dapat merancang intervensi yang menargetkan miskonsepsi konseptual secara spesifik. Pendekatan ini sejalan dengan pandangan bahwa asesmen merupakan bagian integral

dari proses pembelajaran dan berfungsi sebagai alat umpan balik untuk memperbaiki struktur pengetahuan peserta didik (Abate dkk., 2020; Zohar & Dori, 2003).

Dengan demikian, kontribusi penelitian ini bukan hanya pada penggunaan fenomena lokal sebagai konteks, tetapi pada penempatannya sebagai instrumen diagnostik yang mengungkap konsistensi antara penalaran dan klasifikasi konsep. Asesmen yang dirancang secara kontekstual dan analitik memungkinkan pendidik memperoleh gambaran yang lebih komprehensif mengenai kualitas berpikir ilmiah peserta didik dibandingkan asesmen berbasis jawaban tunggal semata.

Kontribusi dan Kebaruan Penelitian

Kebaruan penelitian ini terletak pada integrasi eksplisit antara analisis penalaran inferensial dan klasifikasi konsep dalam satu desain asesmen berbasis fenomena lokal. Berbeda dari studi sebelumnya yang umumnya menilai HOTS melalui capaian skor atau level literasi sains, penelitian ini mengidentifikasi ketidaksinkronan antara kemampuan menjelaskan mekanisme proses dan ketepatan pelabelan konseptual sebagai indikator kualitas struktur pengetahuan. Pendekatan ini menunjukkan bahwa evaluasi HOTS tidak cukup dilakukan melalui pengukuran performa inferensial semata, tetapi perlu dilengkapi dengan analisis konsistensi konseptual melalui pola distraktor. Dengan demikian, penelitian ini memperluas fungsi asesmen kontekstual dari sekadar alat evaluatif menjadi instrumen diagnostik yang mampu memetakan koherensi representasi ilmiah peserta didik.

SIMPULAN

Penelitian ini menunjukkan bahwa asesmen HOTS berbasis stimulus fenomena lokal tidak hanya mampu mengukur capaian penalaran inferensial dan pemahaman proses ilmiah, tetapi juga mengungkap ketidaksinkronan antara kemampuan menjelaskan mekanisme fenomena dan ketepatan klasifikasi konseptual. Temuan tersebut menegaskan bahwa performa inferensial yang relatif tinggi tidak selalu mencerminkan koherensi struktur pengetahuan secara menyeluruh. Kebaruan penelitian ini terletak pada integrasi analisis penalaran dan klasifikasi konsep dalam satu desain asesmen kontekstual, sehingga fungsi asesmen berkembang dari sekadar evaluatif menjadi diagnostik. Melalui analisis pola respons dan distraktor, asesmen mampu memetakan miskonsepsi yang tidak tampak melalui skor agregat semata. Implikasinya, pengembangan asesmen IPA terpadu perlu mempertimbangkan konsistensi antara proses berpikir dan organisasi konsep sebagai indikator kualitas pemahaman ilmiah. Dengan demikian, stimulus fenomena lokal dapat dimanfaatkan secara lebih strategis untuk memperoleh informasi yang lebih komprehensif mengenai struktur pengetahuan peserta didik.

Daftar Pustaka

- Abate, T., Michael, K., & Angell, C. (2020). Assessment of Scientific Reasoning: Development and Validation of Scientific Reasoning Assessment Tool. *Eurasia Journal of Mathematics, Science and Technology Education*, 16(12), em1927. <https://doi.org/10.29333/ejmste/9353>
- Conrad, D., & Libarkin, J. C. (2022). Using Conceptual Metaphor Theory within the Model of Educational Reconstruction to identify students' alternative conceptions and improve instruction: A plate tectonics example. *Journal of Geoscience Education*, 70(2), 262-277. <https://doi.org/10.1080/10899995.2021.1983941>
- Elshall, A. S., & Badir, A. (2025). Balancing AI-assisted learning and traditional assessment: The FACT assessment in environmental data science education. *Frontiers in Education*, 10. <https://doi.org/10.3389/educ.2025.1596462>
- Fortus, D., Lin, J., Neumann, K., & Sadler, T. D. (2022). The role of affect in science literacy for all. *International Journal of Science Education*, 44(4), 535-555. <https://doi.org/10.1080/09500693.2022.2036384>
- Haladyna, T. M., & Rodriguez, M. C. (2021). Using Full-information Item Analysis to Improve Item Quality. *Educational Assessment*, 26(3), 198-211. <https://doi.org/10.1080/10627197.2021.1946390>
- Hingstman, M., Warrens, M. J., Doolaard, S., & Bosker, R. J. (2023). The effects of Success for All in the Netherlands on the reading achievement of first-grade students at risk of reading problems. *Studies in Educational Evaluation*, 77, 101257. <https://doi.org/10.1016/j.stueduc.2023.101257>

- Holmes, K., Mackenzie, E., Berger, N., & Walker, M. (2021). Linking K-12 STEM Pedagogy to Local Contexts: A Scoping Review of Benefits and Limitations. *Frontiers in Education*, 6. <https://doi.org/10.3389/educ.2021.693808>
- Lahagu, A., Harefa, A. R., Zega, N. A., & Gulo, H. (2024). Improving Higher-Order Thinking Skills-based Science Literacy Questions in Science Learning Using Reading to Learn Model. *BIOEDUKASI*, 422-428. <https://doi.org/10.19184/bioedu.v22i3.52026>
- Lian, L. H. (2023). Assessment of Higher-Order Thinking Skills: Is it Simply Determined by Verbs? *International Journal of Academic Research in Progressive Education and Development*, 12(2), Article 2. <https://doi.org/10.6007/IJARPED/v12-i2/16719>
- Lieberei, T., Großmann, L., Welter, V. D. E., Krüger, D., & Krell, M. (2025). Considering Multiple Sources of Validity Evidence Can Help to Address Challenges in the Development of Pedagogical Content Knowledge (PCK) Multiple-Choice Items. *Research in Science Education*, 55(5), 1349-1369. <https://doi.org/10.1007/s11165-024-10227-9>
- OECD. (2023, April 27). *Innovating Assessments to Measure and Support Complex Skills*. OECD Publishing. OECD. <https://doi.org/10.1787/e5f3e341-en>
- Pallant, A., Lore, C., Lee, H.-S., Seevers, S., & Lord, T. (2024). Integrating the plate tectonic and rock genesis systems for secondary school students. *Journal of Geoscience Education*, 73, 1-14. <https://doi.org/10.1080/10899995.2024.2407748>
- Ploj Virtič, M. (2022). Teaching science & technology: Components of scientific literacy and insight into the steps of research. *International Journal of Science Education*, 44(12), 1916-1931. <https://doi.org/10.1080/09500693.2022.2105414>
- Rahayu, S., & Rosawati, E. E. (2023). The development of higher-order thinking skills (HOTS) assessment instrument in chemistry using socioscientific issues context: A preliminary trial. *AIP Conference Proceedings*, 2572(1), 030009. <https://doi.org/10.1063/5.0118624>
- Rezigalla, A. A., Eleragi, A. M. E. S. A., Elhussein, A. B., Alfaifi, J., ALGhamdi, M. A., Al Ameer, A. Y., Yahia, A. I. O., Mohammed, O. A., & Adam, M. I. E. (2024a). Item analysis: The impact of distractor efficiency on the difficulty index and discrimination power of multiple-choice items. *BMC Medical Education*, 24(1), 445. <https://doi.org/10.1186/s12909-024-05433-y>
- Rezigalla, A. A., Eleragi, A. M. E. S. A., Elhussein, A. B., Alfaifi, J., ALGhamdi, M. A., Al Ameer, A. Y., Yahia, A. I. O., Mohammed, O. A., & Adam, M. I. E. (2024b). Item analysis: The impact of distractor efficiency on the difficulty index and discrimination power of multiple-choice items. *BMC Medical Education*, 24(1), 445. <https://doi.org/10.1186/s12909-024-05433-y>
- Sajjad, M., Iltaf, S., & Khan, R. A. (2020). Nonfunctional distractor analysis: An indicator for quality of Multiple choice questions. *Pakistan Journal of Medical Sciences*, 36(5), 982-986. <https://doi.org/10.12669/pjms.36.5.2439>
- Venkatesh, V., Brown, S., & Sullivan, Y. (2024). Conducting Mixed-Methods Research. Dalam *Conducting Mixed-Methods Research*. Virginia Tech Publishing. <https://doi.org/10.21061/conducting-mixed-methods-research>
- Xiao, X., Li, Y., He, X., Fang, J., Yan, Z., & Xie, C. (2025). An assessment framework of higher-order thinking skills based on fine-tuned large language models. *Expert Systems with Applications*, 272, 126531. <https://doi.org/10.1016/j.eswa.2025.126531>
- Zohar, A., & Dori, Y. J. (2003). Higher Order Thinking Skills and Low-Achieving Students: Are They Mutually Exclusive? *The Journal of the Learning Sciences*, 12(2), 145-181.