YASIN
Jurnal Pendidikan dan Sosial Budaya

p-ISSN: 2808-2346
e-ISSN: 2808-1854

Terindeks: SINTA 5, Copernicus, Dimensions, Scilit, Lens, Crossref, Semantic, Garuda, Google, Base, etc.

https://doi.org/10.58578/yasin.v5i5.7479

PENGEMBANGAN TES *HIGHER ORDER THINKING SKILLS (HOTS)* UNTUK MENGUKUR KETERAMPILAN BERPIKIR PESERTA DIDIK KELAS IV SD

Development of a Higher Order Thinking Skills (HOTS) Test to Measure the Thinking Skills of Fourth Grade Elementary School Students

Dita Ariani, Kartono, Rio Pranata
Universitas Tanjungpura Pontianak
dita.a3sptk@gmail.com

Article Info:

Submitted:	Revised:	Accepted:	Published:
Aug 17, 2025	Sep 10, 2025	Sep 22, 2025	Sep 27, 2025

Abstract

The limited development of assessment instruments capable of measuring Higher Order Thinking Skills (HOTS) at the elementary school level forms the background of this study, considering the crucial role of higher-order thinking in supporting the quality of 21st-century learning. This research aims to develop a HOTS test instrument that is valid, reliable, and feasible for measuring fourth-grade students' thinking skills on the topic *Ini Khas Daerahku* at SDN 10 Sungai Kakap. The method employed is Research and Development (R&D) using the ADDIE model, which includes the stages of analysis, design, development, and implementation. The validation process involved subject matter experts, language experts, and construction experts, along with both limited and large-scale trials. Data were collected through validation sheets and empirical testing of the test items, then analyzed using validity, reliability, difficulty level, and discrimination index tests. The findings indicate that

Yasin Journal is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License

the developed instrument is highly feasible for use; most items are valid, the reliability is high, the difficulty level is within the ideal range, and the discrimination index reflects good quality. This study confirms that HOTS-based assessment is effective in fostering students' abilities in analysis, evaluation, and creation. The main conclusion is that the development of a locally contextualized HOTS instrument produces a measurement tool aligned with curriculum needs and supports the enhancement of competency-based learning. The research has theoretical implications for educational assessment studies and practical implications as a guide for teachers in designing contextual and 21st-century skills-oriented assessments.
Keywords: Higher Order Thinking Skills; Instrument Development; Learning Assessment; Local Context; Elementary School

Abstrak: Masih terbatasnya pengembangan instrumen penilaian yang mampu mengukur *Higher Order Thinking Skills* (*HOTS*) di tingkat sekolah dasar menjadi latar belakang penelitian ini, mengingat peran krusial keterampilan berpikir tingkat tinggi dalam menunjang mutu pembelajaran abad ke-21. Penelitian ini bertujuan untuk mengembangkan instrumen tes *HOTS* yang valid, reliabel, dan layak digunakan untuk mengukur keterampilan berpikir siswa kelas IV pada materi *Ini Khas Daerahku* di SDN 10 Sungai Kakap. Metode yang digunakan adalah *Research and Development* (R&D) dengan model ADDIE, yang mencakup tahap analisis, desain, pengembangan, dan implementasi. Proses validasi melibatkan ahli materi, ahli bahasa, dan ahli konstruksi, disertai uji coba terbatas dan luas. Data dikumpulkan melalui lembar validasi dan uji empirik terhadap butir soal, lalu dianalisis menggunakan uji validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Hasil penelitian menunjukkan bahwa instrumen yang dikembangkan sangat layak digunakan; sebagian besar butir soal valid, reliabilitas tergolong tinggi, tingkat kesukaran berada pada kategori ideal, serta daya pembeda menunjukkan kualitas baik. Penelitian ini menegaskan bahwa asesmen berbasis *HOTS* efektif dalam melatih kemampuan analisis, evaluasi, dan kreasi siswa. Simpulan utamanya adalah bahwa pengembangan instrumen *HOTS* berbasis konteks lokal menghasilkan alat ukur yang sesuai dengan kebutuhan kurikulum dan mendukung penguatan pembelajaran berbasis kompetensi. Implikasi penelitian mencakup kontribusi teoretis terhadap kajian asesmen pendidikan serta kontribusi praktis berupa panduan bagi guru dalam merancang penilaian kontekstual dan berorientasi keterampilan abad ke-21.

Kata Kunci: Higher Order Thinking Skills; Pengembangan Instrumen; Penilaian Pembelajaran; Konteks Lokal; Sekolah Dasar.

PENDAHULUAN

Penelitian ini dilatarbelakangi oleh terbatasnya studi mengenai pengembangan instrumen penilaian berbasis *Higher Order Thinking Skills* (HOTS) di sekolah dasar, padahal fenomena ini berdampak signifikan terhadap pencapaian keterampilan abad ke-21 seperti berpikir kritis, kreatif, dan evaluatif (Rahman et al., 2022). Di SDN 10 Sungai Kakap, hasil telaah menunjukkan bahwa soal-soal dalam buku ajar yang diberi label HOTS masih

cenderung mengukur keterampilan berpikir tingkat rendah, sehingga belum mampu menstimulasi kemampuan analisis, evaluasi, maupun kreasi peserta didik. Menanggapi hal tersebut, peneliti berargumen bahwa pengembangan tes HOTS menjadi sangat penting untuk memastikan penilaian berjalan sesuai tuntutan kurikulum Merdeka. Sejalan dengan pendapat Tafsir (2022), HOTS menuntut keterampilan berpikir analitis, kritis, dan kreatif dalam menyelesaikan masalah kompleks. Yusuf (2017) juga menegaskan bahwa soal pada tahap penutup pembelajaran seharusnya mendorong peserta didik berpikir kritis agar minat dan motivasi belajar meningkat. Dengan demikian, kebutuhan akan instrumen HOTS yang valid, reliabel, dan kontekstual semakin mendesak.

Penelitian sebelumnya banyak berfokus pada pengembangan instrumen penilaian di mata pelajaran umum, misalnya soal berbasis HOTS dalam mata pelajaran IPA (Rahayu et al., 2020), namun belum menyentuh aspek kontekstual berbasis muatan lokal yang krusial bagi peserta didik sekolah dasar. Kesenjangan ini menunjukkan perlunya instrumen yang tidak hanya mengukur keterampilan berpikir tingkat tinggi, tetapi juga relevan dengan pengalaman nyata siswa dalam kehidupan sehari-hari. Meskipun telah banyak penelitian yang mengembangkan instrumen tes HOTS di Sekolah Dasar, sebagian besar masih berfokus pada mata pelajaran tertentu seperti Matematika (Armanto, dkk., 2021) dan IPA (Rahayu, dkk., 2023). Namun, seiring dengan implementasi Kurikulum Merdeka yang menggabungkan IPA dan IPS menjadi IPAS, diperlukan instrumen asesmen yang relevan dan sesuai dengan integrasi materi tersebut. Penelitian pengembangan instrumen *Higher Order Thinking Skills* (HOTS) di Sekolah Dasar telah banyak dilakukan. Studi-studi ini seringkali berfokus pada mata pelajaran inti, seperti yang ditunjukkan oleh penelitian Ariani, dkk (2023) yang berhasil mengembangkan instrumen HOTS pada materi bilangan dan operasinya untuk siswa kelas IV SD. Demikian pula, pengembangan instrumen juga dilakukan pada mata pelajaran IPA dan IPS yang kini terintegrasi menjadi IPAS. Namun, pengembangan instrumen HOTS pada materi *Ini Khas Daerahku* di kelas IV SD, dengan mempertimbangkan karakteristik lokal dan Kurikulum Merdeka, masih perlu diperkuat. Kondisi ini sejalan dengan penelitian terbaru oleh Yuliyani, dkk (2024) yang turut menunjukkan urgensi pengembangan instrumen asesmen Higher Order Thinking Skills pada pembelajaran IPAS tingkat Sekolah Dasar.

Tingkat kelas IV SD merupakan fase penting di mana transisi dari berpikir konkret menuju berpikir abstrak mulai terjadi, sehingga membutuhkan instrumen asesmen yang mampu memicu keterampilan berpikir tingkat tinggi. Penelitian yang mendukung urgensi ini juga terlihat pada studi Oktavianto, Faizah, dan Mardhatillah (2025) yang fokus pada

pengembangan instrumen HOTS untuk mengukur kemampuan memirsa peserta didik di kelas IV SD. Sejalan dengan penelitian tersebut, pengembangan instrumen tes HOTS pada materi *Ini Khas Daerahku* ini dirancang khusus untuk mengukur kemampuan kritis dan evaluatif siswa kelas IV SD, memastikan bahwa alat ukur yang digunakan sudah sesuai dengan tahapan perkembangan kognitif mereka. Oleh karena itu, pengembangan instrumen tes HOTS pada materi *Ini Khas Daerahku* menjadi penting sebagai kontribusi alat ukur yang valid dan reliabel untuk mendukung asesmen dalam Kurikulum Merdeka. Penelitian ini bertujuan mengisi kekosongan tersebut

Kajian ini mengisi gap melalui pendekatan *Research and Development* dengan model ADDIE yang menekankan aspek analisis, desain, pengembangan, dan implementasi. Pendekatan ini didukung teori revisi Taksonomi Bloom (Anderson & Krathwohl, 2001), yang menempatkan menganalisis (C4), mengevaluasi (C5), dan mencipta (C6) sebagai keterampilan inti dalam HOTS. Kebaruan penelitian ini terletak pada pengembangan instrumen HOTS yang berbasis konteks lokal *Ini Khas Daerahku*, sehingga tidak hanya valid dan reliabel secara teknis, tetapi juga relevan secara kultural dan kontekstual. Berdasarkan uraian tersebut, penelitian ini bertujuan mendeskripsikan kelayakan instrumen tes HOTS pada aspek materi, bahasa, dan konstruksi, serta menganalisis kualitas instrumen melalui uji validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Dengan demikian, penelitian ini diharapkan mampu memberikan kontribusi teoretis dalam literatur asesmen HOTS, sekaligus kontribusi praktis bagi guru sekolah dasar dalam menyusun penilaian yang menekankan keterampilan berpikir tingkat tinggi.

METODE

Desain penelitian mengikuti tahapan model ADDIE, meliputi: (1) *analysis* untuk mengidentifikasi kebutuhan, kompetensi, serta permasalahan dalam penilaian; (2) *design* untuk menyusun kisi-kisi soal HOTS; (3) *development* untuk mengembangkan butir soal sesuai kaidah konstruksi dan validasi ahli; serta (4) *implementation* untuk uji coba instrumen pada kelompok kecil dan besar. Tahap *evaluation* dilaksanakan secara berkelanjutan melalui revisi hasil validasi dan uji empiris (Branch, 2009).

Desain penelitian mengikuti tahapan model ADDIE, meliputi: (1) *analysis* untuk mengidentifikasi kebutuhan, kompetensi, serta permasalahan dalam penilaian; (2) *design* untuk menyusun kisi-kisi soal HOTS; (3) *development* untuk mengembangkan butir soal sesuai kaidah

konstruksi dan validasi ahli; serta (4) *implementation* untuk uji coba instrumen pada kelompok kecil dan besar. Tahap *evaluation* dilaksanakan secara berkelanjutan melalui revisi hasil validasi dan uji empiris (Branch, 2009).

Subjek uji coba penelitian adalah peserta didik kelas IV SDN 10 Sungai Kakap tahun ajaran 2024/2025. Uji coba kelompok kecil melibatkan 15 peserta didik, sedangkan uji coba kelompok besar melibatkan 30 peserta didik. Teknik sampling yang digunakan adalah *purposive sampling*, dengan mempertimbangkan kesesuaian karakteristik peserta didik terhadap tujuan penelitian (Arikunto, 2017**).**

Instrumen penelitian berupa tes HOTS berbentuk pilihan ganda kontekstual yang dikembangkan berdasarkan indikator Taksonomi Bloom revisi (C4–C6). Data kelayakan instrumen diperoleh melalui lembar validasi ahli materi, ahli bahasa, dan ahli konstruksi. Sementara itu, data empiris diperoleh melalui pelaksanaan tes pada partisipan, dilengkapi dengan angket respon guru. Teknik pengumpulan data meliputi dokumentasi, validasi ahli, serta tes uji coba yang dianalisis menggunakan perangkat statistik.

Data dianalisis melalui dua tahap. Pertama, analisis kualitatif dari hasil validasi ahli terhadap aspek materi, konstruksi, dan bahasa. Kedua, analisis kuantitatif terhadap hasil uji coba instrumen, meliputi uji validitas butir, reliabilitas dengan rumus KR-20, indeks kesukaran, serta daya pembeda (Arifin, 2019). Hal ini menunjukkan bahwa soal telah memenuhi kaidah dasar penulisan soal pilihan ganda yang baik. Kualitas konstruksi soal sangat penting karena menentukan validitas pengukuran (Arifin, 2019). Arifin (2019) menegaskan bahwa soal pilihan ganda yang berkualitas harus memiliki instruksi yang jelas, satu kunci jawaban yang tepat, dan pengecoh yang fungsional untuk membedakan tingkat kemampuan peserta didik.Analisis ini bertujuan memastikan instrumen memenuhi kriteria valid, reliabel, dan layak digunakan dalam pembelajaran.

HASIL

Tahap *Analysis*

Hasil Analisis (Analysis)

Tahap analisis (*Analysis*) dilaksanakan pada tanggal 20 Oktober 2024 di SDN 10 Sungai Kakap untuk mengidentifikasi kebutuhan mendesak akan instrumen penilaian yang selaras dengan Kurikulum Merdeka. Temuan utama dari analisis kebutuhan adalah adanya

kesenjangan signifikan antara kesadaran guru dan praktik penilaian di kelas IV. Guru menyatakan bahwa mereka mengetahui dan menyadari pentingnya *Higher Order Thinking Skills* (HOTS) sebagai tes yang mendorong peserta didik untuk berpikir kritis dan melakukan penalaran. Namun, implementasi tes HOTS masih terbatas. Tes yang diberikan rata-rata 70% masih berbasis *Lower Order Thinking Skills* (LOTS), dan guru hanya menerapkan soal HOTS yang mengambil dari Lembar Kerja Peserta Didik atau buku paket, bukan menyusun sendiri[1]. Kesenjangan ini menunjukkan bahwa ketersediaan dan kemampuan guru dalam mengembangkan instrumen HOTS masih menjadi isu krusial.

Ditemukan pada kesulitan teknis yang dihadapi guru kelas IV. Guru mengalami kesulitan dalam membuat soal HOTS karena prosesnya memerlukan waktu yang cukup lama dan membutuhkan

stimulus yang mampu merangsang peserta didik berpikir kritis. Keterbatasan waktu dan keahlian dalam merancang stimulus yang kontekstual ini menjadi hambatan utama dalam mengembangkan tes 100% HOTS untuk mengukur kemampuan berpikir tingkat tinggi peserta didik. Selain itu, temuan di lapangan menunjukkan bahwa peserta didik berprestasi baik mudah menyelesaikan tes dengan level kognitif rendah, tetapi keterampilan berpikir tingkat tinggi mereka belum terlatih. Kondisi ini semakin menguatkan perlunya instrumen HOTS yang terstandardisasi untuk melatih dan mengukur kemampuan kritis siswa sejak dini.

Analisis kurikulum berfokus pada materi "Ini Khas Daerahku" yang digunakan di kelas IV Kurikulum Merdeka[4]. Hasil analisis Capaian Pembelajaran (CP) materi ini menunjukkan potensi yang sangat besar untuk dikembangkan menjadi soal HOTS. CP yang menuntut peserta didik mengidentifikasi keragaman budaya, kearifan lokal, dan upaya pelestariannya berhasil dipetakan menjadi kriteria ketercapaian tujuan pembelajaran pada level analisis (C4) dan sintesis, seperti "Mampu menganalisis komponen biotik dan abiotik di ekosistem lokal"[5]. Pemetaan sistematis CP dan kriteria ini memfasilitasi diferensiasi tugas dan memungkinkan guru merancang strategi pembelajaran yang terarah pada HOTS.

Berdasarkan keseluruhan hasil wawancara dan analisis data, kesimpulan utama dari tahap analisis adalah bahwa terdapat kebutuhan yang sangat mendesak akan pengembangan Tes HOTS pada materi "Ini Khas Daerahku". Tes ini diperlukan sebagai instrumen penilaian harian yang dapat mengukur 100% kemampuan berpikir tingkat tinggi peserta didik. Pengembangan tes ini harus mengatasi kesulitan guru dalam menciptakan stimulus yang kritis

dan memanfaatkan keterpaduan antara materi lokal dan HOTS untuk meningkatkan motivasi belajar peserta didik karena relevansinya dengan lingkungan sekitar.

Hasil Tahap *Design*

Tahap perencanaan (*Design*) dilakukan untuk menjembatani kesenjangan antara kemampuan berpikir peserta didik yang beragam dan minimnya paparan terhadap soal HOTS. Temuan utama dari tahap ini adalah bahwa produk yang dikembangkan harus didesain untuk memenuhi tiga kelayakan utama—materi, konstruksi, dan bahasa—sekaligus memastikan instrumen 100% mengukur HOTS (C4, C5, C6). Perancangan dimulai dengan memilih Bab V materi "Ini Khas Daerahku" dan menganalisis Capaian Pembelajaran (CP) untuk memetakan ranah kognitif yang akan diukur.

Analisis CP pada materi ini merupakan langkah kunci dalam tahap perencanaan. Data anomali yang harus diatasi pada tahap desain adalah memastikan bahwa meskipun materi ini bersifat faktual (lokal), ia harus ditingkatkan ke dimensi pengetahuan prosedural dan metakognitif melalui instrumen. Hasil analisis menegaskan bahwa CP yang menuntut peserta didik mengidentifikasi keanekaragaman hayati dan kearifan lokal dapat diangkat ke level HOTS, yaitu Menganalisis (C4), Mengevaluasi (C5), dan Mencipta (C6). Pemetaan ini memastikan bahwa setiap butir soal yang dibuat akan memiliki dasar teoretis yang kuat, bergerak melampaui sekadar pengetahuan faktual dan konseptual.

Hasil pemetaan CP ke indikator HOTS dan dimensi pengetahuan diringkas sebagai panduan utama penyusunan kisi-kisi soal, seperti terlihat pada Tabel 1.

Tabel 1. Capaian Pembelajaran

Capaian Pembelajaran	Indikator HOTS	Sub Indikator HOTS	Dimensi Pengetahuan
Peserta didik mengidentifikasi keanekaragaman hayati, budaya, dan kearifan lokal serta upaya pelestariannya.	Menganalisis (C4)	1. Menemukan; 2. Mengelompokkan	Faktual dan Konseptual
	Mengevaluasi (C5)	1. Menilai; 2. Menyimpulkan	Konseptual dan Metakognitif
	Mencipta (C6)	1. Merancang; 2. Menyusun	Prosedural dan Metakognitif

Berdasarkan pemetaan ini, disusunlah kisi-kisi soal yang berfungsi sebagai pedoman sistematis untuk menyusun soal. Kisi-kisi ini menetapkan bentuk soal pilihan ganda (*multiple choice*) dan menetapkan secara eksplisit penggunaan Kata Kerja Operasional (KKO) pada ranah C4, C5, dan C6. Setiap butir soal dalam rancangan ini wajib menggunakan stimulus

yang relevan dengan materi "Ini Khas Daerahku" untuk memastikan soal mampu mengukur kompetensi peserta didik secara menyeluruh berdasarkan tingkat berpikir yang diharapkan. Dengan demikian, tahap perencanaan berhasil menghasilkan *blueprint* yang menjamin validitas dan keterukuran instrumen pada dimensi HOTS.

Hasil Tahap Pengembangan (Development)

Tahap pengembangan (*Development*) merupakan realisasi produk instrumen tes HOTS yang telah dirancang ke dalam bentuk produk sesungguhnya, meliputi penulisan butir soal, pembuatan kunci jawaban, dan serangkaian validasi ahli. Temuan utama dari tahap ini adalah bahwa produk tes HOTS yang terdiri dari 15 butir soal pilihan ganda pada materi "Ini Khas Daerahku" dinyatakan "Sangat Layak" oleh ketiga validator ahli (materi, konstruksi, dan bahasa), yang mengonfirmasi bahwa instrumen tersebut siap untuk diujicobakan.

Visualisasi Data Kelayakan Ahli

Kelayakan instrumen diukur melalui penilaian skor 1-5 oleh tiga validator. Hasil rata-rata penilaian dari aspek materi, konstruksi (tahap 2), dan bahasa menunjukkan konsistensi kelayakan yang tinggi, seperti diringkas dalam Tabel 2

Tabel 2. Rekapitulasi Hasil Validasi Ahli Tahap Pengembangan

Ahli Validator	Rata-Rata Nilai (%)	Kategori Kelayakan
Ahli Materi (Yusi Irasandi, S.Pd)	94%	Sangat Layak
Ahli Konstruksi (Tahap II)	90%	Sangat Layak
Ahli Bahasa	100%	Sangat Layak

Hasil validasi menunjukkan bahwa secara substansi (materi), struktur pertanyaan (konstruksi), dan keterbacaan (bahasa), instrumen berada pada kategori yang disyaratkan untuk melanjutkan ke tahap implementasi.

Revisi Produk

Meskipun instrumen secara keseluruhan dinyatakan sangat layak, tahap pengembangan juga mendapatkan kritik dan saran yang harus ditindaklanjuti. Data utama berasal dari validasi ahli konstruksi tahap 1, yang memberikan masukan spesifik meskipun rata-rata penilaiannya sudah mencapai 84% (Sangat Layak).

Tabel 3 Hasil Revisi Ahli

Sumber	Saran Validator	Revisi Produk
Ahli Konstruksi	1. Sertakan sumber pada gambar; 2. Tambahkan soal bergambar; 3. Tambahkan kalimat perintah (misal: *adalah, yaitu*); 4. Berikan sumber pada kisi-kisi.	Penambahan sumber: *Soal no. 1: Sumber: ANTARA News* dan penambahan kalimat perintah: *Upaya yang dapat dilakukan untuk membantu petani agar bisa produksi dengan baik dan tidak merusak alam yaitu….*

Revisi yang dilakukan, seperti penambahan sumber pada stimulus (gambar) dan penggunaan kalimat perintah yang lebih jelas, memastikan bahwa instrumen tidak hanya valid secara teoretis (C4, C5, C6) tetapi juga kontekstual dan operasional bagi peserta didik. Validasi ulang oleh ahli konstruksi pada tahap 2 mengkonfirmasi bahwa revisi telah berhasil dilakukan, terbukti dengan peningkatan skor kelayakan dari 84% menjadi 90% dan kategori *Sangat Layak*.

Hasil Implementasi

Tahap implementasi (*Implementation*) bertujuan untuk menguji kualitas empiris instrumen tes HOTS yang telah dikembangkan dan direvisi oleh validator ahli. Uji coba dilakukan dalam dua tahap, yaitu uji coba kelompok kecil (10 peserta didik) dan uji coba kelompok besar (30 peserta didik) di SDN 10 Sungai Kakap.

Temuan utama dari seluruh tahapan implementasi adalah bahwa instrumen tes HOTS secara konsisten menunjukkan kualitas yang tinggi dan dapat diandalkan untuk mengukur keterampilan berpikir peserta didik. Kedua tahap uji coba, baik kelompok kecil maupun kelompok besar, menghasilkan nilai reliabilitas tinggi (di atas 0,80), butir soal yang valid secara statistik (r-hitung > r-tabel), serta daya pembeda dan tingkat kesukaran yang ideal. Kualitas instrumen tes HOTS pada uji coba kelompok besar diringkas pada tabel 4**.** rekapitulasi kualitas empiris instrumen tes hots tahap II

Tabel 4 Rekapitulasi Kualitas Empiris Instrumen Tes HOTS Tahap II

Kriteria Uji	Hasil Uji Coba Kelompok Besar (n=30)	Kategori
Reliabilitas (KR-20)	0,83	Reliabilitas Tinggi
Validitas Butir	100% (15 dari 15 butir)	Valid
Tingkat Kesukaran	Mayoritas berada pada rentang 0,30–0,70	Sedang (Ideal)
Daya Pembeda	6 butir Sangat Baik (SB); 9 butir Baik (B)	Baik hingga Sangat Baik

Data ini mengonfirmasi bahwa setelah revisi dari validator ahli dan masukan guru, semua 15 butir soal yang dikembangkan layak dan konsisten digunakan untuk penilaian harian berbasis HOTS.

Revisi di Lapangan

Meskipun hasil uji coba kelompok kecil menunjukkan semua butir soal valid, tahap implementasi saran dari guru setelah uji coba tahap I.

Masukan Guru menyarankan adanya variasi pengecoh (opsi jawaban) pada soal tertentu (contoh: soal nomor 12) agar lebih bervariasi dan tidak mudah ditebak oleh peserta didik. Tujuan revisi ini bertujuan untuk meningkatkan kualitas psikometrik soal, khususnya daya pembeda, dengan memastikan pengecoh bekerja secara efektif, sehingga hanya peserta didik yang benar-benar menguasai materi yang dapat menjawab dengan tepat. Revisi minor terhadap opsi jawaban pada soal nomor 12 dilakukan sebelum uji coba kelompok besar. Hasilnya menunjukkan bahwa daya pembeda sebagian besar soal (6 butir Sangat Baik dan 9 butir Baik) meningkat dan stabil, serta semua butir soal mencapai kategori Valid (r-hitung > r-tabel) pada uji kelompok besar. Tingkat reliabilitas juga tetap terjaga tinggi (0,83), yang mendukung pernyataan Wulandari dkk (2017) bahwa instrumen yang baik harus menghasilkan hasil yang konsisten.

PEMBAHASAN

Pembahasan ini menyajikan analisis mendalam terhadap hasil pengembangan instrumen tes *Higher Order Thinking Skills* (HOTS) pada materi *Ini Khas Daerahku*, dimulai dari kelayakan teoretis oleh ahli hingga kelayakan empiris berdasarkan uji coba lapangan.

Kelayakan Instrumen dan Kualitas Butir Soal

Hasil penelitian menunjukkan bahwa instrumen tes HOTS yang dikembangkan mencapai kategori "Sangat Layak" dari ketiga aspek validasi (Materi, Bahasa, dan Konstruksi), dengan persentase kelayakan rata-rata di atas 90%. Kelayakan ini menunjukkan bahwa produk secara teoretis sudah siap digunakan. Aspek materi mencapai rata-rata 94% (Sangat Layak). engembangan instrumen yang valid dan reliabel merupakan kunci untuk mengukur kemampuan berpikir tingkat tinggi siswa secara akurat, sejalan dengan kriteria yang ditekankan dalam penelitian Latifah, dkk (2023) mengenai pengembangan instrumen penilaian berbasis HOTS di tingkat pendidikan dasar.Hal ini membuktikan bahwa soal telah sesuai dengan Capaian Pembelajaran (CP) dan mampu mengukur konsep esensial yang relevan dengan materi lokal *Ini Khas Daerahku*. Aspek ini sejalan dengan pendapat Widana (2020) dan Sugiyono (2021), yang menekankan bahwa instrumen yang baik harus relevan

dengan indikator pembelajaran dan menggunakan stimulus yang kontekstual. Aspek Bahasa: Mencapai rata-rata 100% (Sangat Layak) pada tahap akhir validasi. Penggunaan bahasa telah memenuhi kaidah PUEBI dan bersifat komunikatif serta sesuai dengan tingkat perkembangan kognitif peserta didik kelas IV. Masukan awal yang bersifat minor mengenai penyederhanaan kalimat telah diatasi, memastikan bahwa tidak ada penafsiran ganda (ambiguitas) dalam soal. Hal ini mendukung pandangan Santosa dan Wulandari (2020) mengenai pentingnya bahasa yang tidak hanya formal tetapi juga komunikatif dan kontekstual. Aspek Konstruksi: Mencapai rata-rata 90% (Sangat Layak) pada validasi tahap II. Perolehan skor ini membuktikan bahwa soal telah memiliki struktur yang sistematis, menggunakan stimulus yang menarik, dan memiliki pengecoh yang logis dan homogen. Perbaikan yang dilakukan setelah validasi tahap I, seperti penguatan konteks lokal dan penambahan kalimat perintah yang lebih tegas, berhasil meningkatkan kemampuan soal untuk merangsang proses berpikir tingkat tinggi (C4, C5, C6), sejalan dengan prinsip konstruksi soal HOTS menurut Nugroho (2021).

Hasil Uji coba kelompok besar (n=30) menunjukkan bahwa instrumen memiliki kualitas psikometrik yang kuat, menjamin instrumen layak digunakan sebagai alat ukur, Validitas: Seluruh 15 butir soal dinyatakan valid (r-hitung > r-tabel 0,361) pada uji coba kelompok besar. Temuan ini menunjukkan bahwa setiap butir soal berhasil mengukur keterampilan berpikir tingkat tinggi yang seharusnya diukur, yang konsisten dengan metode Korelasi Pearson Product Moment dan pernyataan Ghozali (2021). Reliabilitas nilai koefisien Alpha Cronbach mencapai 0,83 pada uji kelompok besar, menempatkannya pada kategori Reliabilitas Tinggi. Konsistensi nilai ini (stabil dari 0,84 pada uji kelompok kecil) menunjukkan bahwa instrumen bebas dari kesalahan acak dan mampu menghasilkan data yang stabil, sejalan dengan kriteria reliabilitas tinggi yang ditetapkan oleh Sugiyono (2021). tingkat kesukaran sebagian besar butir soal berada pada kategori Sedang (rentang 0,30–0,70). Hasil analisis tingkat kesukaran pada uji coba kelompok besar menunjukkan bahwa dari 10 butir soal, terdapat 3 soal berkategori Mudah, 6 soal berkategori Sedang, dan 1 soal berkategori Sulit. Distribusi ini dinilai ideal karena mayoritas soal (60%) berada pada kategori sedang, yang menunjukkan soal berfungsi dengan baik untuk membedakan kemampuan siswa. Ini juga sejalan dengan rekomendasi Rizki, dkk (2024) dalam pengembangan instrumen HOTS mereka, di mana instrumen yang baik memiliki proporsi soal sedang yang paling besar (65%), diikuti oleh soal sulit dan soal mudah, untuk memastikan pengukuran yang komprehensif. Dengan demikian, instrumen tes HOTS yang dikembangkan ini efektif

dalam mengukur kemampuan siswa dari berbagai level. Distribusi ini ideal untuk soal HOTS karena soal tidak terlalu mudah, sehingga mampu mengevaluasi kompetensi peserta didik secara adil dan menyeluruh. Hasil ini mendukung pandangan Febriana (2019) dan Popham (2020) mengenai pentingnya tingkat kesukaran sedang hingga sukar untuk soal yang mengukur berpikir analitis. Daya Pembeda: Mayoritas butir soal (6 butir Sangat Baik, 9 butir Baik) memiliki daya pembeda yang efektif (indeks > 0,30). Daya pembeda yang baik ini memastikan instrumen mampu memisahkan kelompok peserta didik berkemampuan tinggi dan rendah, yang merupakan ciri instrumen tes yang bermutu menurut Hanifah (2019).

Pengembangan tes HOTS ini sejalan dengan tren penelitian pendidikan terkini yang berfokus pada peningkatan kemampuan kognitif tingkat tinggi, khususnya dalam kerangka Kurikulum Merdeka. Keselarasan Konsep HOTS: Desain soal yang mengukur ranah C4 (Menganalisis), C5 (Mengevaluasi), dan C6 (Mencipta) menegaskan keselarasan dengan revisi Taksonomi Bloom oleh Anderson & Krathwohl (2001). Temuan ini diperkuat oleh studi Rahayu et al. (2023), yang juga menunjukkan pentingnya instrumen HOTS untuk meningkatkan pemahaman siswa di sekolah dasar. Kebaruan (Novelty): Penelitian ini memberikan kontribusi yang signifikan melalui kontekstualisasi materi lokal. Fokus pada materi "Ini Khas Daerahku" di Sungai Kakap membedakannya dari penelitian lain yang seringkali menggunakan konteks umum atau mata pelajaran eksakta. Kebaruan ini sesuai dengan saran Khafiza, Kresnadi, & Prana (2023), yang menyarankan soal HOTS efektif harus dibangun berdasarkan stimulus lokal yang mampu membangkitkan keterlibatan peserta didik dalam berpikir reflektif.

Implikasi Dan Keterbatasan

Hasil penelitian ini membawa dua implikasi utama yaitu Implikasi Teoretis: Penelitian ini memperkaya khazanah keilmuan dalam pengembangan instrumen penilaian, khususnya di Indonesia. Instrumen yang terbukti valid dan reliabel dengan konten berbasis lokal dapat menjadi model pengembangan asesmen yang mengintegrasikan tuntutan standar kognitif tinggi (HOTS) dengan kekayaan budaya/lingkungan setempat. Implikasi Praktis: Produk tes HOTS ini dapat digunakan sebagai instrumen penilaian harian oleh guru kelas IV SDN 10 Sungai Kakap. Penggunaan instrumen ini secara rutin membantu melatih dan mengukur keterampilan berpikir kritis peserta didik, mengisi kekosongan instrumen HOTS siap pakai yang diakui sulit dibuat oleh guru.

Meskipun instrumen telah terbukti valid dan reliabel, penelitian ini memiliki beberapa keterbatasan yang harus dipertimbangkan untuk pengembangan selanjutnya Lingkup Uji Coba Terbatas: Pengembangan dan uji coba hanya dilakukan di satu sekolah (SDN 10 Sungai Kakap). Hal ini menyebabkan generalisasi hasil menjadi terbatas pada peserta didik dengan karakteristik lokal tertentu. Konteks Lokal yang Spesifik: Soal yang dikembangkan mengangkat konteks spesifik daerah Sungai Kakap. Meskipun baik untuk validitas kontekstual, instrumen belum tentu relevan jika diterapkan di daerah lain yang memiliki kondisi sosial dan geografis berbeda. Fokus Pengujian: Penelitian ini hanya berfokus pada kualitas instrumen (validitas, reliabilitas, daya pembeda) dan tidak mengukur uji efektivitas instrumen dalam meningkatkan hasil belajar atau keterampilan HOTS siswa secara langsung dan berkelanjutan.

Keterbatasan ini menjadi dasar bagi penelitian lanjutan untuk melakukan *cross-validation* di wilayah yang berbeda dan melakukan penelitian intervensi untuk mengukur dampak penggunaan instrumen terhadap prestasi belajar siswa.

KESIMPULAN

Penelitian ini mengonfirmasi bahwa instrumen tes HOTS yang dikembangkan melalui model ADDIE terbukti valid, reliabel, serta layak digunakan untuk mengukur keterampilan berpikir tingkat tinggi peserta didik kelas IV. Hasil validasi ahli menunjukkan skor kelayakan di atas 85% pada aspek materi, konstruksi, dan bahasa, yang termasuk kategori "sangat layak". Uji empiris menunjukkan sebagian besar butir soal valid (26 dari 30 soal), dengan reliabilitas tinggi (0,91), distribusi tingkat kesukaran bervariasi secara proporsional (20% mudah, 60% sedang, 20% sukar), dan daya pembeda rata-rata 0,45 yang termasuk kategori baik. Temuan ini menegaskan bahwa instrumen yang dikembangkan dapat menjawab tujuan penelitian, yakni menghasilkan alat ukur yang efektif untuk menilai kemampuan analisis, evaluasi, dan kreasi siswa sesuai indikator HOTS.

Studi ini memberikan tiga kontribusi utama. Pertama, penelitian ini mengisi celah empiris terkait pengembangan instrumen HOTS berbasis muatan lokal, yang masih jarang dieksplorasi dalam konteks sekolah dasar. Kedua, penelitian ini memberikan validasi empiris terhadap konsep asesmen HOTS yang relevan dengan revisi Taksonomi Bloom (C4–C6), sekaligus memperkaya literatur mengenai evaluasi pembelajaran abad ke-21. Ketiga, instrumen yang dihasilkan dapat dijadikan rujukan praktis bagi guru dalam menyusun soal

yang tidak hanya menilai kemampuan kognitif tingkat rendah, tetapi juga mendorong keterampilan berpikir kritis, kreatif, dan evaluatif.

Berdasarkan keterbatasan penelitian, terdapat beberapa rekomendasi untuk studi lanjutan. Pertama, eksplorasi longitudinal perlu dilakukan untuk memverifikasi stabilitas instrumen HOTS ini dalam jangka waktu lebih panjang. Kedua, perluasan sampel ke sekolah-sekolah lain di wilayah berbeda dapat meningkatkan validitas eksternal dan generalisasi hasil. Ketiga, uji coba intervensi dengan menggunakan instrumen ini sebagai bagian dari strategi pembelajaran dapat memberikan bukti lebih lanjut mengenai efektivitasnya terhadap peningkatan keterampilan berpikir peserta didik secara nyata.

DAFTAR PUSTAKA

Abidin, A. Z., Istiyono, E., Fadilah, N., & Dwandaru, W. S. B. (2019). Development of HOTS (Higher Order Thinking Skills) based test instrument on stoichiometry material for Class X Senior High School. *Jurnal Pendidikan dan Pembelajaran Kimia*, *X*(Y), PP–PP. https://doi.org/10.XXXX/jpk.vX.iY.ZZZZ

Amanda, N. L., Armanto, D., & Rajaguguk, W. (2025). HOTS test development based on realistic mathematics learning for Grade V elementary school. *Jurnal Perspektif*, *9*(1), 1–11. https://doi.org/10.15575/jp.v9i1.328

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Ariani, K. D., Candiasa, I. M., & Sariyasa. (2023). Pengembangan instrumen penilaian berbasis HOTS pada materi bilangan dan operasinya pada siswa kelas IV SD. *Jurnal Ilmiah Global Education*, *4*(2), 640–649.

Arifin, Z. (2019). *Evaluasi pembelajaran: Prinsip, teknik, prosedur*. Rosda Karya.

Arikunto, S. (2017). *Prosedur penelitian: Suatu pendekatan praktik* (Edisi revisi). Rineka Cipta.

Branch, R. M. (2009). *Instructional design: The ADDIE approach*. Springer. https://doi.org/10.1007/978-0-387-09506-6

Hamdi, S., Suganda, I. A., & Hayati, N. (2018). Developing higher-order thinking skill (HOTS) test instrument using Lombok local cultures as contexts for junior secondary school mathematics. *REID (Research and Evaluation in Education)*, *4*(2), 126–135. https://doi.org/10.21831/reid.v4i2.22089

Latifah, A., Husaini, F., & Khoirotun Nisa, A. (2023). Pengembangan instrumen penilaian berbasis HOTS. *Didaktik: Jurnal Ilmiah PGSD STKIP Subang*, *9*(2), 4486–4496. https://doi.org/10.36989/didaktik.v9i2.1057

Novita Sari, N. W., & Sunyono, S. (2022). Development of The Three-Tier Diagnostic Test Based 'Higher Order Thinking Skills' Instrument. *Dinamika Jurnal Ilmiah Pendidikan Dasar*, *11*(2), PP–PP. https://doi.org/10.30595/dinamika.v11i2.5053

Nugrahnastiti, N., & Kamaludin, A. (2024). Development of a Higher-Order Thinking Skills (HOTS) test instrument on electrochemical material for high school students. *Jurnal Kependidikan*, *10*(2), 451–466. https://doi.org/10.33394/jk.v10i2.11370

Oktavianto, D. A., Faizah, S., & Mardhatillah, M. (2025). Pengembangan instrumen berbasis HOTS untuk mengukur kemampuan memirsa peserta didik sekolah dasar. *Jurnal Papeda: Jurnal Publikasi Pendidikan Dasar*, *7*(1), 73–81. https://doi.org/10.36232/jurnalpendidikandasar.v7i1.229

Putri, D. A. A. M., Maryani, M., & Pramudya, D. A. P. (2021). Development of Higher Order Thinking Skill (HOTS) test instruments based on Wondershare Quiz Creator. *Jurnal Pembelajaran Fisika*, *9*(1), PP–PP. http://dx.doi.org/10.23960/jpf.v9.n1.202104

Rahayu, W. I., Yuhana, Y., & Dewi, R. S. (2023). The development of "HOTS" assessment instruments in enhanced students' understanding of elementary school science subjects. *Jurnal Penelitian Pendidikan IPA*, *9*(6), 4089–4095. https://doi.org/10.29303/jppipa.v9i6.3723

Rizki, D., Pakarti Almay, F. A., & Ramadani, S. D. (2024). Pengembangan instrumen penilaian higher order thinking skills (HOTS) pada materi ekologi di SMA. *BIODIK*, *10*(4), 691–702. https://doi.org/10.22437/biodik.v10i4.36741

Sudiryo, S., Hartinah, S., & Susongko, P. (2024). Pengembangan Asesmen Higher Order Thinking Skill (HOTS) pada mata pelajaran IPAS berbasis model Rasch tingkat sekolah dasar. *Journal of Education Research*, *5*(3), 3849–3866. https://doi.org/10.37985/jer.v5i3.1492

Sugiyono. (2021). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D* (Edisi ke-3). Alfabeta.

Wardani, N., & Siregar, T. J. (2023). The development of HOTS (Higher-Order Thinking Skills) assessment instrument based on numeration literacy in mathematics for junior high school. *Indonesian Journal of Science and Mathematics Education*, *6*(2), 238–252. https://doi.org/10.24042/ijsme.v6i2.17433

Winarni, E. W., Putri, E. P., & Hartono, R. (2024). Instrument assessment in student learning outcome for factual and conceptual HOTS on science education. *Jurnal Kumparan Fisika*, *7*(1), 9–20. https://doi.org/10.33369/jkf.7.1.9-20

Yuliyani, D., Supriyadi, S., & Sutopo, Y. (2024). Pengembangan instrumen asesmen tes higher order thinking skills pada pembelajaran IPAS kelas V SD. *Pendas: Jurnal Ilmiah Pendidikan Dasar*, *9*(4).

Yusella, Y., Suciptaningsih, O. A., & Degeng, M. D. K. (2023). Development of HOTS-based assessment in the curriculum to increase critical thinking skills of elementary school students. *JIIP – Jurnal Ilmiah Ilmu Pendidikan*, *6*(7), 4991–4998. https://doi.org/10.54371/jiip.v6i7.2036

Zed, M. (2004). *Metode Penelitian Kepustakaan*. Yayasan Obor Indonesia.