
STRATEGI PENYUSUNAN INSTRUMEN EVALUASI UNTUK MENGUKUR KEMAMPUAN BERPIKIR TINGKAT TINGGI (HIGHER ORDER THINKING SKILLS)

Hesti Novita Sari¹, Ernawati², Musalamatul Husna³, Rizki Wirayuda⁴.

Alamat e-mail: hestiovitassari327@gmail.com, ernawati@unuha.ac.id,

musalamatulhusna@gmail.com, rizkiwirayuda@student.unuha.ac.id

Bahasa dan Sastra Indonesia
Universitas Nurul Huda

Submit: 04-07-2026, Revisi: 07-07-2026, Terbit: 29-07-2026

DOI: 10.30739/peneroka.v6i2.5382

Abstrak

Penelitian ini bertujuan untuk mendeskripsikan strategi penyusunan instrumen evaluasi yang digunakan dalam mengukur kemampuan berpikir tingkat tinggi *Higher Order Thinking Skills* atau (HOTS), Implementasinya dalam pembelajaran, serta tantangan dan upaya peningkatan kualitas instrumen. Ruang lingkup penelitian difokuskan pada proses penyusunan instrumen evaluasi HOTS, mulai dari analisis capaian pembelajaran (CP), tujuan pembelajaran (TP). Perumusan indikator, penyusunan kisi-kisi, pengembangan stimulus, penulisan butir soal, penyusunan rubrik penskoran, validasi, hingga uji coba instrumen berdasarkan berbagai hasil penelitian yang relevan. Penelitian ini menggunakan metode *library research* dengan pendekatan kualitatif deskriptif. Data penelitian berupa data sekunder yang diperoleh dari artikel ilmiah, buku, prosiding, dan dokumen akademik yang dipublikasikan dalam sepuluh tahun terakhir, kemudian dianalisis menggunakan Teknik *content analysis* melalui reduksi data, klasifikasi, penyajian, interpretasi, dan penarikan simpulan. Hasil kajian menunjukkan bahwa penyusunan instrumen HOTS harus dilakukan secara sistematis dengan mengacu pada kompetensi pembelajaran, level kognitif C4-C6, stimulus kontekstual, rubrik penskoran, validasi ahli, dan uji coba empiris sehingga menghasilkan instrumen yang valid, reliabel, dan objektif. Implementasi instrumen HOTS mampu mengukur kemampuan peserta didik dalam menganalisis, mengevaluasi, dan menciptakan solusi secara lebih akurat. Namun, pengembangannya masih menghadapi kendala, terutama kompetensi guru dan kualitas penyusunan instrumen. Oleh karena ini, diperlukan pelatihan, pendampingan, validasi berkelanjutan, serta pemanfaatan teknologi digital untuk meningkatkan kualitas instrumen evaluasi HOTS.

Kata kunci: strategi, instrumen evaluasi pembelajaran, kemampuan berpikir tingkat tinggi.

Abstract

This study aims to describe the strategies for developing assessment instruments used to measure higher-order thinking skills (HOTS), their implementation in the classroom, and the challenges and efforts to improve the quality of these instruments. The scope of the study focuses on the process of developing HOTS assessment instruments, ranging from the analysis of learning outcomes (CP) and learning objectives (TP), the formulation of indicators, the development of a test blueprint, the

creation of stimuli, the writing of test items, the development of a scoring rubric, validation, to the pilot testing of the instruments based on various relevant research findings. This study employs a library research method with a descriptive qualitative approach. The research data consists of secondary data obtained from scientific articles, books, conference proceedings, and academic documents published within the last ten years, which were then analyzed using content analysis techniques through data reduction, classification, presentation, interpretation, and drawing of conclusions. The results of the study indicate that the development of HOTS instruments must be carried out systematically, taking into account learning competencies, cognitive levels C4–C6, contextual stimuli, scoring rubrics, expert validation, and empirical testing, in order to produce instruments that are valid, reliable, and objective. The implementation of HOTS instruments can measure students' abilities to analyze, evaluate, and create solutions more accurately. However, their development still faces challenges, particularly regarding teacher competence and the quality of instrument design. Therefore, training, mentoring, ongoing validation, and the use of digital technology are needed to improve the quality of HOTS assessment instruments.

Keywords: *strategies, learning evaluation instruments, high-level thinking skills.*

Pendahuluan

Perkembangan pendidikan pada abad ke-21 menuntut peserta didik bukan hanya harus memahami pengetahuan, tapi juga memiliki kemampuan berpikir tingkat tinggi atau *Higher Order Thinking Skills* (HOTS). HOTS sangat diperlukan untuk pembelajaran dalam ranah lingkungan sebab dalam masalah lingkungan adalah masalah yang kompleks sehingga siswa membutuhkan kompetensi menganalisis, mengevaluasi, dan mencipta yang diperlukan untuk menghadapi berbagai permasalahan dalam kehidupan real (Carel et al., 2021). HOTS mejadi salah satu kompetensi yang penting sehingga perlu dikembangkan lewat proses pembelajaran sebab memungkinkan membantu peserta didik untuk berpikir kritis, kreatif, dan mampu mengambil keputusan yang tepat. Di dukung oleh tegar (Saputra et al., 2025) HOTS merupakan kompetensi berpikir yang menuntut siswa untuk memodifikasi materi dan gagasan yang sudah dipelajari sebelumnya untuk menghasilkan sintesa yang baru.

Untuk mendapati sejauh mana kemampuan berpikir tingkat tinggi peserta didik berkembang, diperlukan sebuah proses evaluasi yang tepat dan terarah. Evaluasi pembelajaran memiliki peran peting dalam mengukur ketercapaian tujuan pembeljaran dan memberikan informasi mengenai kemampuan peserta didik. Dalam konteks HOTS, evaluasi tidak hanya berfungsi untuk mengukur mengingat ataupun memahami materi, tetapi harus mampu mengukur kompetensi peserta didik dalam menelaah informasi, mengevaluasi berbagai alternatif pemecahan masalah, dan menghasilkan sebuah gagasan atau solusi baru. Didukung oleh (Mariam et al., 2020) kemampuan guru dalam menguasai dan membuat alat evaluasi berbasis HOTS dalam proses pembelajaran sebagai usaha meningkatkan kualitas peserta didik.

Namun, berdasarkan praktiknya penyusunan instrumen evaluasi yng mampu mengukur kemampuan berpikir tigkat tingi peserta didik masih menghadapi berbagai kendala. Tidak sedikit

instrumen evaluasi yang di gunakan dalam pembelajaran masih berfokus terhadap kemampuan berpikir tingkat rendah. Seperti yang di kemukakan oleh (Aini et al., 2023) analis soal yang sering dimanfaatkan dalam PAS oleh guru slalu mengandung bentuk soal LOTS dimana sebagian besar memberikan soal yang hanya melatih proses C1 (megingat), C2 (memahami) ataupun C3 (menerapkan) dari materi yang pernah diajarkan. Kemudian penelitian yang di lakukan oleh (Winari et al., 2021) diperoleh hasil bahwa menyusun soal HOTS tentu bukan perkara mudah, sebab soal yang dibuat harus diawali dngan stimulus yang menjadi pengantar soal HOTS tersebut. Akibatnya, kompetensi peserta didik saat menganalisis, mengevaluasi, dan menciptakan belum terukur secara optimal. Dilain itu, sebagian pendidik masih mengalami kesulitan dalam merumuskan indikator HOTS, yaitu menyusun stimulus yang kontekstual, dan megembangkan butir-butir soal yang sesuai degan karakteristik kemampuan berpikir tingkat tinggi. Kondisi ini menunjukkan bahwa kualitas instrumen evaluasi masih perlu di tingkatkan agar sejalan dengan tuntutan pemblajaran abad ke-21 yang berorientasi dalam pengembangan HOTS.

Guru masih sering menyamakan soal HOTS dengan soal yang sulit, padahal tingkat kesulitan berbeda dengan tingkat kompleksitas berpikir yang diukur dalam HOTS (Driana et al., 2021). Menurut Ropianingsih et al., (2021) menyatakan bahwa instrumen penilaian yang diberikan guru kepada peserta didik masih berada pada level C1-C3 sehingga dampaknya dapat dilihat terhadap hasil belajar peserta didik yang belum mampu menyesuaikan permasalahan yang lebih kompleks. Didukung oleh penelitian yang di lakukan Azizah & Silfianah, (2024) di peroleh hasil tes bahwa mayoritas siswa masih menunjukkan tingkat pemahaman yang rendah dalam menyelesaikan soal-soal HOTS.

Menurut aderson dan krathwohl dalam revisi taksonomi bloom, HOTS berada pada tingkatan kognitif yang meliputi kemampuan (C4), (C5), dan (C6). Kompetensi tersebut menuntut peserta didik untuk mengolah informasi secara mendalam, menilai berbagai alternatif berdasarkan pertimbangan tertentu, serta menghasilkan gagasan atau solusi baru. Penelitian Pinasti dan Ernawati, (2026) menunjukkan bahwa instrumen soal dalam buku text Bahasa Indonesia kelas X didominasi oleh soal berbasis HOTS sebesar 62,5%, sehingga pembelajaran tidak lagi bertendensi pada hafalan, tetapi mendorong peserta didik untuk menganalisis, mengevaluasi dan menciptakan gagasan baru. Oleh itu, penyusunan evaluasi perlu di lakukan secara sistematis dengan memperhatikan indikator HOTS agar instrumen yang di kembangkan mampu mengukur kemampuan berpikir tingkat tinggi secara valid dan reliabel.

Beberapa penelitian terdahulu telah mengkaji upaya penyusunan maupun pengembangan instrumen evaluasi berbasis HOTS. Septiana et al., (2024) meneliti strategi guru dalam mengembangkan instrumen berbasis HOTS pada mata pelajaran Bahasa Indonesia. Sejalan dengan itu, Yusmilda et al., (2023) mengembangkan instrumen penilaian tes berbasis HOTS pada jenjang

pendidikan dasar dalam menghadapi tuntutan era (*super smart society*). Selanjutnya, Putri et al., (2023) menyusun instrumen penilaian HOTS daalam pembelajaran IPS di Sekolah dasar (SD), dari ketiga penelitian tersebut terlihat bahwa strategi penyusunan instrumen evaluasi HOTS yang ada saat ini masih bersifat parsial, yaitu terikat pada mata pelajaran atau jenjang pendidikan tertentu, sehingga belum ada rumusan strategi yang bersifat konseptual dan dapat diaplikasikan secara umum. Oleh karena itu, penelitian ini memiliki kebaruan (*novelty*) berupa perumusan strategi penyusunan instrumen evaluasi HOTS yang bersifat konseptual, sistematis, integratif, dan tidak terikat pada satu mata pelajaran maupun jenjang pendidikan tertentu. Strategi ini disusun berdasarkan sintesis kendala-kendala umum yang dihadapi pendidik di lapangan serta dilandaskan pada taksonomi Bloom revisi Anderson dan Krathwohl (C4–C6), sehingga dapat menjadi rujukan yang relevan bagi pendidik lintas bidang studi dan jenjang pendidikan.

Berdasarkan uraian tersebut, artikel ini bertujuan untuk mendeskripsikan strategi penyusunan instrumen evaluasi yang dapat digunakan untuk mengukur kemampuan berpikir tingkat tinggi (*Higher Order Thinking Skills* atau HOTS) secara efektif sehingga mampu menghasilkan informasi yang akurat tentang kompetensi peserta didik.

Metode Penelitian

Penelitian ini menerapkan metode kajian pustaka (*library research*) yang dipadukan dengan pendekatan kualitatif deskriptif. Kajian pustaka merupakan sebuah metode yang mengandalkan aneka sumber tertulis sebagai data utama guna menghasilkan bagian dari proposal penelitian, di mana data-datanya dihimpun melalui sumber informasi berupa buku, artikel, jurnal, media daring, serta dokumen-dokumen lain (Saefullah, 2024). Pendekatan kualitatif deskriptif diterapkan untuk memaparkan dan memaknai berbagai konsep, strategi, maupun temuan penelitian yang terkait dengan proses penyusunan instrumen evaluasi yang kemudian menjadi indikator *Higher Order Thinking Skills* (HOTS). Kerangka konseptual dalam penelitian ini didasarkan pada Taksonomi Bloom versi revisi Anderson dan Krathwohl, terutama pada ranah kemampuan berpikir tingkat tinggi yang meliputi (C4), (C5), dan (C6).

Data yang digunakan bersifat sekunder, bersumber dari artikel ilmiah, jurnal nasional maupun internasional, buku, prosiding, dan dokumen akademik lain yang berkaitan dengan pengembangan instrumen evaluasi HOTS. Proses pengumpulan data dilaksanakan secara sistematis melalui beberapa tahap. Tahap pertama, peneliti menetapkan fokus kajian, yakni strategi penyusunan instrumen evaluasi untuk mengukur kemampuan berpikir tingkat tinggi. Tahap kedua, peneliti melakukan penelusuran dan identifikasi terhadap berbagai sumber pustaka yang relevan. Tahap ketiga, sumber-sumber yang telah diperoleh diseleksi berdasarkan kesesuaian tema, tingkat kredibilitas, serta kecocokannya dengan tujuan penelitian. Tahap keempat, literatur yang terpilih dihimpun dan didokumentasikan sebagai bahan analisis. Prioritas diberikan pada literatur terbitan

sepuluh tahun terakhir agar dapat menggambarkan perkembangan mutakhir mengenai konsep, penyusunan, penerapan, dan evaluasi instrumen HOTS.

Analisis data dilaksanakan dengan teknik analisis isi (*content analysis*) melalui beberapa tahapan berurutan. Tahap pertama, peneliti melaksanakan reduksi data dengan menyaring informasi yang sesuai dengan fokus kajian. Tahap kedua, data yang telah terpilih dikelompokkan berdasarkan tema, meliputi perumusan kompetensi dan indikator, penyusunan kisi-kisi, pengembangan stimulus, penulisan butir soal, penyusunan pedoman penskoran, validasi, uji coba, hingga implementasi instrumen HOTS. Tahap ketiga, data yang telah dikelompokkan tersebut dipaparkan secara deskriptif agar pola dan temuan bertendensi lebih mudah dipahami. Tahap keempat, peneliti melakukan interpretasi dengan membandingkan hasil-hasil penelitian guna menemukan kesamaan, perbedaan, dan tendensi terkait strategi penyusunan instrumen evaluasi HOTS. Tahap akhir berupa penarikan simpulan yang didasarkan pada hasil sintesis atas seluruh data yang telah dianalisis. Legalitas data diperkuat lewat triangulasi sumber, yaitu dengan membandingkan informasi dari beragam jurnal, buku, dan dokumen ilmiah yang relevan.

Pembahasan

1. Perumusan Instrumen Evaluasi untuk Mengukur *Higher Order Thinking Skills* (HOTS)

Berdasarkan telaah terhadap sejumlah penelitian, proses perumusan instrumen HOTS ternyata tidak dimulai dari penyusunan butir soal, melainkan diawali dengan menganalisis kompetensi yang hendak diukur, seperti analisis Capaian Pembelajaran (CP) dan Tujuan Pembelajaran (TP). Setelah itu, dilanjutkan dengan menentukan indikator soal, menyusun kisi-kisi, menyusun stimulus, menyusun butir soal, hingga menyusun kurikulum (Litna et al., 2021). Penelitian tersebut menjelaskan bahwa penyusunan instrumen diawali dengan penetapan spesifikasi tes yang mencakup kompetensi, indikator, serta bentuk instrumen yang selaras dengan tujuan pembelajaran.

Selanjutnya Tujuan Pembelajaran (TP) merupakan penjabaran operasional dari Capaian Pembelajaran (Bait et al., 2025). Menurut Hartono et al, (2023) memperlihatkan bahwa analisis kurikulum dan tujuan pembelajaran merupakan dasar dalam merumuskan instrumen evaluasi berpikir tingkat tinggi (HOTS). Penyusunan instrumen HOTS dilakukan lewat penetapan (CP), indikator pembelajaran, materi, indikator soal, level kognitif Taksonomi Bloom (C4–C6), bentuk soal, kunci jawaban, dan pedoman penskoran yang disusun secara sistematis dalam kisi-kisi.

Berdasarkan penelitian tersebut, tahap analisis menjadi landasan utama sebelum penyusunan kisi-kisi dan penulisan soal dilakukan. Pada tahap ini, peneliti menganalisis kompetensi pembelajaran, karakteristik peserta didik, materi, dan kebutuhan pembelajaran sehingga instrumen yang dikembangkan sesuai dengan tujuan evaluasi.

Tabel 1. komponen Penyusunan Instrumen Evaluasi HOST

Komponen	Fungsi
Capaian pembelajaran (CP)	Menentukan kompetensi yang akan diukur
Tujuan Pembelajaran	Menjabarkan capaian pembelajaran
Indikator soal	Menjadi acuan penyusun soal
Materi	Menentukan ruang lingkup soal
Level kognitif	Menentukan tingkat berpikir (C4-C6)
Bentuk soal	Menyesuaikan teknik evaluasi
Kunci jawaban	Acuan Penilaian
Pedoman Penskoran	Menjamin Objektivitas penilaian

Butir Soal

Pada tahap ini, butir soal dikembangkan dengan mengacu pada indikator yang telah dirumuskan sebelumnya, guna mengukur kemampuan berpikir tingkat tinggi peserta didik. Penyusunan butir soal disesuaikan dengan tujuan pembelajaran, dapat berbentuk pilihan ganda maupun benar-salah (Widodo et al., 2024). Penulisan butir soal menjadi komponen utama dalam penyusunan instrumen HOTS. Pada instrumen berbasis HOTS, butir soal dirancang bukan hanya untuk menaksir kemampuan mengingat, melainkan untuk manskir kemampuan peserta didik dalam menyelesaikan soal atas level kognitif (C4–C6) sesuai dengan Taksonomi Bloom revisi.

Penulisan butir soal HOTS diawali dengan penyusunan stimulus yang kontekstual. Stimulus berfungsi sebagai sumber informasi awal yang harus dipahami peserta didik sebelum menjawab pertanyaan. Instrumen HOTS dirancang agar peserta didik tidak sekedar mengingat konsep, tapi juga menggunakan pengetahuan yang dimiliki untuk menyelesaikan masalah dalam berbagai konteks kehidupan (Wahidin, 2023). Variasi bentuk stimulus juga membantu instrumen mengukur kemampuan berpikir tingkat tinggi secara lebih autentik (Herwanti et al., 2026). Penggunaan stimulus yang kontekstual dan bervariasi terbukti dapat meningkatkan kualitas instrumen HOTS (Viyanti et al., 2023).

Tabel 2. Contoh Butir Soal HOTS Berdasarkan Level Kognitif Taksonomi Bloom Revisi

Level Kognitif	Indikator Kemampuan	Contoh Butir Soal
	Menganalisis informasi, mengidentifikasi hubungan,	Bacalah teks argumentasi analisis kekuatan dan kelemahan

C4 (Menganalisis)	atau menemukan kelemahan suatu argumen	argumen penulis berdasarkan fakta yang disajikan dalam teks
C5 (Mengevaluasi)	Memberikan penilaian berdasarkan kriteria atau bukti yang tersedia	Perhatikan dua teks argumentasi berikut evaluasilah teks yang memiliki argumen lebih kuat disertai alasan yang logis berdasarkan data yang tersedia.
C6 (Mencipta)	Menghasilkan gagasan solusi, atau karya baru berdasarkan informasi yang diperoleh	Berdasarkan permasalahan lingkungan pada teks berikut, susunlah sebuah argumen yang disertai fakta dan solusi yang benar dan tepat untuk menyelesaikan permasalahan tersebut.

Berdasarkan Tabel tersebut, perbedaan butir soal pada setiap level HOTS terletak ditingkat tuntutan proses berpikir yang harus dilakukan oleh peserta didik. Soal pada level **C4 (Menganalisis)** , **C5 (Mengevaluasi)** , **C6 (Mencipta)**

Rubik atau pedoman penskoran.

Menurut Batubara dan Sudrajat, (2019) setelah penulisan butir soal, tahapan berikutnya adalah menyusun kunci jawaban atau pedoman penskoran agar proses penilaian berlangsung secara objektif dan sesuai dengan indikator yang telah ditetapkan.

Tabel 3. Contoh Rubrik Penskoran Instrumen Evaluasi HOTS

Aspek yang Dinilai	Kriteria Penilaian	Skor
Ketepatan Analisis	Analisis sangat tepat, logis dan didukung fakta yang relevan	4
	Analisis Tepat tetapi kurang lengkap	3
	Analisis kurang tepat dan alasan kurang lengkap	2
	Analisis tidak sesuai atau tidak menjawab	1
Argumentasi	Argumentasi yang logis, sistematis dan didukung bukti	4
	Argumentasi cukup logis namun bukti kurang lengkap	3
	Argumentasi kurang logis	2
	Tidak memberikan argumentasi	1
Kesimpulan	Kesimpulan hasil analisis dan didukung alasan	4
	Kesimpulan cukup sesuai	3
	Kesimpulan kurang sesuai	2

Validasi Instrumen

Setelah butir soal dan pedoman penskoran disusun, tahap yang dilakukan berikutnya adalah validasi instrumen. Menurut Zainuddin et al., (2020) dan Suarti, (2022) Seluruh butir soal dinyatakan layak digunakan karena memenuhi kriteria materi, konstruksi, dan bahasa, serta terbukti valid dan reliabel setelah uji coba lapangan. Hal ini sejalan dengan penelitian Rusdianto et al. (2020) yang menyatakan bahwa instrumen penilaian berbasis HOTS valid secara teoritis berdasarkan judgment ahli, serta valid dan reliabel secara empiris berdasarkan hasil uji coba lapangan.

Tabel 4. Aspek validitas Instrumen HOTS

Aspek Validasi	Tujuan Validasi
Validasi Isi	Menilai kesesuaian soal dengan CP, TP indikator, dan materi
Validitas Konstruksi	Memaastikan soal mengukur kemampuan HOTS (C4,C5,C6)
Validitas bahasa	Memastikan bahasa soal jelas, komunikatif, dan tidak ambigu

Berdasarkan Tabel tersebut, validasi instrumen tidak hanya menilai kesesuaian materi dengan tujuan pembelajaran, tetapi memastikan bahwa setiap butir soal mampu mengukur kemampuan berpikir tingkat tinggi dan menggunakan bahasa yang mudah dipahami peserta didik.

Uji Coba Instrumen

Setelah uji coba dinyatakan valid oleh ahli, tahap selanjutnya adalah ujicoba instrumen kepada peserta didik.

Tabel 5. Aspek Uji Coba Instrumen

Aspek Uji Coba	Tujuan
Tingkat kesukaran	Mengetahui apakah soal terlalu mudah, sedang, atau terlalu sulit
Daya Pembeda	Mengetahui kemampuan soal dalam membedakan peserta didik yang berkemampuan tinggi dan rendah
Validitas	Menguji apakah alat ukur benar benar mengukur variabel yang hendak diteliti
Realibilitas	Menguji konsistensi dan kestabilan alat ukur jika digunakan dalam waktu atau pada b kelompok sampel yang berbeda

2. Strategi Penyusunan Instrumen Evaluasi untuk *Mengukur Higher Order Thinking Skills* (HOTS)

Studi kepustakaan Rahman et al. (2026) menunjukkan bahwa pengembangan instrumen HOTS pada umumnya menggunakan model ADDIE dan Borg & Gall, serta sebagian besar telah melalui validasi isi (expert judgment). Namun, penggunaan analisis psikometrik modern seperti Model Rasch masih terbatas, sehingga kualitas instrumen yang dihasilkan cenderung bervariasi. Temuan ini diperkuat oleh Ratu et al. (2026) yang menyatakan bahwa strategi penyusunan instrumen HOTS perlu diawali dengan identifikasi kompetensi pembelajaran, dilanjutkan dengan penyusunan soal kontekstual yang mampu merangsang kemampuan berpikir kritis, analitis, dan kreatif peserta didik. Sejalan dengan itu, Faulina dan Jamil (2026) menyatakan bahwa evaluasi pembelajaran berbasis HOTS akan lebih efektif jika diterapkan melalui penilaian autentik yang bersifat formatif, reflektif, dan berorientasi pada proses.

Menurut Iqbal et al., (2024) evaluasi pembelajaran berbasis HOTS dilakukan secara komprehensif dengan menggunakan berbagai bentuk penilaian, seperti tes tertulis, tes lisan, penilaian kinerja, proyek, portofolio, dan penilaian diri. Menurut Sholeha et al., (2024) strategi penyusunan instrumen evaluasi HOTS perlu memperhatikan keterpaduan antara tujuan pembelajaran, aktivitas belajar, bentuk penilaian, dan mekanisme umpan balik sehingga instrumen yang dikembangkan mampu mengukur sekaligus mengembangkan kemampuan berpikir tingkat tinggi peserta didik secara optimal.

3. Implementasi Instrumen Evaluasi untuk *Mengukur Higher Order Thinking Skills* (HOTS)

Literature review Manullang et al. (2025) menunjukkan bahwa soal berbasis stimulus kontekstual lebih efektif mengukur kemampuan berpikir tingkat tinggi (C4–C6) dibanding soal yang bersifat mengingat. Instrumen yang paling banyak digunakan berupa soal uraian, pilihan ganda beralasan, dan tugas pemecahan masalah, karena dinilai mampu menggambarkan kemampuan berpikir peserta didik secara komprehensif. Temuan ini sejalan dengan hasil validitas instrumen HOTS pada materi keanekaragaman hayati yang dikembangkan Rahmawati dan Trimulyono (2021), yang menunjukkan bahwa instrumen tervalidasi ahli secara konsisten memiliki tingkat kelayakan tinggi dalam mengukur aspek analisis dan evaluasi peserta didik.

Implementasi instrumen evaluasi HOTS dilakukan dengan memasukkan soal-soal yang berada pada tingkat kognitif (C4), (C5), dan (C6) ke dalam proses pembelajaran. Penggunaan instrumen ini tidak hanya berfungsi sebagai alat evaluasi pada akhir pembelajaran, tetapi juga dimanfaatkan sebagai evaluasi formatif untuk memantau perkembangan kemampuan berpikir peserta didik. Melalui evaluasi tersebut, guru dapat mengetahui tingkat penguasaan materi sekaligus kemampuan

peserta didik dalam memecahkan masalah, menyusun argumen, serta mengambil keputusan. Hasil evaluasi selanjutnya dijadikan sebagai dasar untuk memperbaiki strategi pembelajaran agar proses belajar berlangsung lebih efektif dan berfokus pada pengembangan kemampuan berpikir tingkat tinggi (Magdalena, 2023).

Menurut Kurnia* et al. (2022), implementasi instrumen evaluasi HOTS yang telah melewati tahapan validasi, uji reliabilitas, analisis tingkat kesukaran, dan analisis daya pembeda mampu menghasilkan informasi yang lebih akurat mengenai kemampuan peserta didik dalam menganalisis, mengevaluasi, dan memecahkan masalah.

4. Tantangan dan Upaya Peningkatan Kualitas Instrumen Evaluasi untuk Mengukur *Higher Order Thinking Skills* (HOTS)

Berdasarkan penelitian Risnita et al. (2024), pemahaman guru dalam menyusun instrumen HOTS masih belum optimal. Ryandi dan Santri (2024) juga menyatakan bahwa penyusunan instrumen HOTS masih didominasi oleh soal berpikir tingkat rendah sehingga belum mampu mengukur kemampuan analisis, evaluasi, dan kreasi peserta didik secara optimal.

Temuan tersebut diperkuat oleh Tyas et al., (2025) yang menyatakan bahwa perkembangan teknologi digital membuka peluang besar dalam meningkatkan kualitas instrumen evaluasi. Hasil Systematic Literature Review menunjukkan bahwa pemanfaatan platform digital, aplikasi asesmen, dan sistem penilaian berbasis komputer membantu guru menyusun instrumen yang lebih variatif, mempercepat pemberian umpan balik, serta mempermudah analisis kualitas butir soal. Untuk mengatasi berbagai tantangan dalam penyusunan instrumen evaluasi HOTS, diperlukan beberapa upaya sebagai berikut:

1) Peningkatan Kompetensi Guru

Kompetensi guru dapat ditingkatkan melalui pelatihan, workshop, dan Musyawarah Guru Mata Pelajaran (MGMP). Kegiatan tersebut terbukti membantu guru mengembangkan indikator, kisi-kisi, dan butir soal HOTS sesuai dengan karakteristik kemampuan berpikir tingkat tinggi.

2) Pendampingan Guru

Guru perlu memperoleh pendampingan berkelanjutan dalam menyusun instrumen HOTS, termasuk melalui pendekatan Item Response Theory (IRT) dan pelatihan analisis butir soal. Pendampingan ini bertujuan untuk meningkatkan kompetensi guru dalam mengembangkan instrumen yang valid dan reliabel.

3) Pemanfaatan Teknologi

Pemanfaatan perangkat lunak analisis butir soal mampu membantu guru menganalisis validitas, reliabilitas, tingkat kesukaran, dan daya pembeda soal sehingga instrumen evaluasi HOTS yang dihasilkan memiliki kualitas yang lebih baik.

Simpulan

Berdasarkan hasil kajian kepustakaan, dapat disimpulkan bahwa penyusunan instrumen evaluasi untuk mengukur *Higher Order Thinking Skills* (HOTS) harus dilakukan secara sistematis dan berlandaskan pada kompetensi pembelajaran yang ingin dicapai. Proses penyusunannya diawali dengan analisis Capaian Pembelajaran (CP), Tujuan Pembelajaran (TP), dan indikator pencapaian kompetensi, kemudian dilanjutkan dengan penyusunan kisi-kisi, penyusunan stimulus yang kontekstual, penulisan butir soal pada level kognitif C4, C5, dan C6, penyusunan rubrik penskoran, validasi ahli, serta uji coba instrumen untuk memastikan validitas, reliabilitas, tingkat kesukaran, dan daya pembeda soal. Rangkaian tersebut menghasilkan instrumen lebih valid, reliabel, objektif, dan mampu menaksir kompetensi peserta didik dalam menganalisis, mengevaluasi serta menciptakan solusi terhadap berbagai masalah.

Penelitian ini menunjukkan bahwa strategi penyusunan instrumen HOTS dapat menjadi acuan bagi guru, dosen, dan pengembang asesmen dalam merancang evaluasi pembelajaran yang selaras dengan tuntutan abad ke-21. Penerapan strategi tersebut diharapkan mampu meningkatkan kualitas evaluasi serta mengembangkan kemampuan berpikir kritis, kreatif, dan pemecahan masalah peserta didik.

Kelebihan penelitian ini terletak pada sintesis berbagai penelitian terkini yang memberikan gambaran komprehensif mengenai penyusunan, implementasi, dan pengembangan instrumen HOTS. Namun, karena menggunakan metode library research, penelitian ini hanya mengandalkan data sekunder sehingga belum menguji efektivitasnya secara empiris. Oleh karena itu, penelitian selanjutnya disarankan mengembangkan dan menguji instrumen HOTS pada berbagai jenjang pendidikan dan mata pelajaran untuk mendapati bukti empirik yang lebih kuat.

Daftar Rujukan

- Aini, R., Karma, I. N., & Affandi, L. H. (2023). Kesulitan guru dalam menyusun soal evaluasi berbasis higher order thinking skills dalam pembelajaran kurikulum 2013. *Jurnal Educatio FKIP UNMA*, 9(4), 2062-2069. <https://doi.org/10.31949/educatio.v9i4.6035>
- Azizah, I. N., & Silfianah, I. (2024). Pengembangan Instrumen Penilaian Kognitif Berbasis Higher Order Thinking Skills (HOTS) Berbantu Aplikasi Quizizz pada Materi Ikatan Kimia dan Gaya Antarmolekul. *UNESA Journal of Chemical Education*, 13(2), 159-170. <https://doi.org/10.26740/ujced.v13n2.p159-170>
- Bait, E. H., Mulyasari, E., Hendriawan, D., Arwasih, A., & Ulwan, M. N. (2025). Kurikulum Merdeka dan Dinamika Tujuan Pendidikan: Integrasi Capaian Pembelajaran (CP), Tujuan Pembelajaran (TP), dan Alur Tujuan Pembelajaran (ATP). *Kalam Cendekia: Jurnal Ilmiah Kependidikan*, 13(1). <https://doi.org/10.20961/jkc.v13i1.97505>

- Batubara, U. N., & Sudrajat, A. (2019). Teknik penyusunan instrumen penilaian higher order thinking skill (HOTS) dalam pembelajaran sejarah. *Lentera Pendidikan: Jurnal Ilmu Tarbiyah dan Keguruan*, 22(2), 335-344. <https://doi.org/10.24252/lp.2019v22n2i15>
- Carel, G., Jusniani, N., & Monariska, E. (2021). Kemampuan higher order thinking skills dalam pembelajaran metakognitif ditinjau dari persepsi siswa. *PYTHAGORAS : Jurnal Matematika dan Pendidikan Matematika*, 16(2), 204-216. <https://doi.org/10.21831/pythagoras.v16i2.37926>
- Driana, E., Susilowati, A., Ernawati, & Ghani, A. R. A. (2021). Assessing students' higher-order thinking skills: Knowledge and practices of chemistry teachers in vocational senior secondary schools. *Jurnal Pendidikan Teknologi dan Kejuruan*, 27(1), 37–47. <https://doi.org/10.21831/jptk.v27i1.32882>
- Faulina, R. F., & Jamil, S. (2026). Peran Evaluasi Pembelajaran Berbasis Hots Dalam Mendukung Keterampilan Berpikir Kritis Peserta Didik: Systematic Literature Review: The Role of HOTS-Based Learning Evaluation in Supporting Students' Critical Thinking Skills: A Systematic Literature Review. *Jurnal Education*, 2(2), 134-143. <https://doi.org/10.59562/education.v2i2.13185>
- Hartono^①, R., Winarni^②, E. W., & Koto^③, I. (2023). Pengembangan Instrumen Penilaian Higher Order Thingking Skills (HOTS) pada Pembelajaran IPA Siswa Kelas IV SDN 21 Bengkulu Tengah. *Jurnal KAPEDAS–Kajian Pendidikan Dasar*, 2(1), 205-217. <https://doi.org/10.33369/kapedas.v2i1.26539>
- Iqbal, M. G., Khasanah, S. U., & Luthfi, A. (2025). Strategi dan Evaluasi High Order Thingking Skill HOTS dalam Pembelajaran PAI (Studi Kasus di SMAN 114 Jakarta). *Ideguru: Jurnal Karya Ilmiah Guru*, 10(1), 23-29. <https://doi.org/10.51169/ideguru.v10i1.1428>
- Herwanti, K., Nugrohadhi, S., Nugraha, S. C. P., & Asri, D. U. (2026). Stimulus Fenomena Lokal sebagai Asesmen Diagnostik HOTS IPA Terpadu. *JURNAL PENDIDIKAN MIPA*, 16(2), 240-246. <https://doi.org/10.37630/jpm.v16i2.4121>
- Litna, K. O. (2021). Pengembangan instrumen tes higher order thinking skills (HOTS) matematika SMA kelas X (Doctoral dissertation, UNIVERSITAS PENDIDIKAN GANESHA). <https://doi.org/10.23887/jpepi.v11i1.278>
- Kurnia*, L. D., Haryati, S., & Linda, R. (2022). Pengembangan Instrumen Evaluasi Higher Order Thinking Skills Menggunakan Quizizz Pada Materi Termokimia untuk Meningkatkan Kemampuan Berpikir Tingkat Tinggi Peserta Didik. *Jurnal Pendidikan Sains Indonesia*, 10(1), 176–190. <https://doi.org/10.24815/jpsi.v10i1.21727>
- Magdalena, M. (2023). Implementasi Instrument Assessment HOTS (Higher Order Thinking Skills) Tema 7 di Kelas 1 Sekolah Dasar. *Jurnal Tonggak Pendidikan Dasar: Jurnal Kajian Teori dan Hasil Pendidikan Dasar*, 2(2), 179-192. <https://doi.org/10.22437/jtpd.v2i2.29330>
- Manullang, S. S. C., Gea, S. J. B. P., & Nahadi, N. (2025). Higher-Order Thinking Skills (HOTS)-Based Assessment in Chemistry Education: A Systematic Literature Review of Its Implementation in Senior High Schools. *Journal of The Indonesian Society of Integrated Chemistry*, 17(1), 38-45. <https://doi.org/10.22437/jisic.v17i1.44953>
- Mariam, P., Nurhayati, Y., & Irmawan, I. (2020). Penerapan evaluasi pembelajaran berbasis hots. *Jurnal Pengabdian Tri Bhakti*, 2(2), 171-178. <https://doi.org/10.36555/tribhakti.v2i2.1696>

- Pinasti, R., & Ernawati, E. (2026). Tingkatan Kognitif Soal Buku Cerdas Cergas Bahasa Indonesia Kelas X Berdasarkan Taksonomi Bloom (Revisi) . *DEIKTIS: Jurnal Pendidikan Bahasa Dan Sastra*, 6(2), 2154-2161. <https://doi.org/10.53769/deiktis.v6i2.3683>
- Rahman, F., Al-Musyarraf, M. S., Moulana, M. R., & Pradita, E. K. (2026). Teknik pengembangan instrumen penilaian berbasis HOTS mengacu pada taksonomi Anderson & Krathwohl: Studi kepustakaan 2015–2025. *Jurnal Jendela Pendidikan*, 6(01), 59-65. <https://doi.org/10.57008/jjp.v6i01.2132>
- Rahmawati, D. E., & Trimulyono, G. (2022). Validitas instrumen penilaian Higher Order Thinking Skills (HOTS) pada materi keanekaragaman hayati. *Berkala Ilmiah Pendidikan Biologi (BioEdu)*, 11(1), 138-147. <https://doi.org/10.26740/bioedu.v11n1.p138-147>
- Ratu, P. D. H., Tualaka, F., & Ndoen, T. (2026). Implementasi Taksonomi Bloom Revisi (Anderson & Krathwohl) dalam Pengembangan Instrumen Berbasis HOTS. *Educational Journal*, 1(4), 1649-1655. <https://doi.org/10.63822/bhyspa39>
- Risnita, R., Sastradika, D., & Al Farisi, S. (2024). Teacher Skills in Preparing Evaluation Instruments: Case Study of Teacher's Understanding to Construct Hots Instrument. *Jurnal Pendidikan Sains Universitas Muhammadiyah Semarang*, 12(1), 20-30. <https://doi.org/10.26714/jps.12.1.2024.20-30>
- Rusudianto, A. R., Susanta, A. S., & Muktadir, A. M. (2020). Pengembangan Instrumen Penilaian Berbasis Higher Order Thinking (Hot) Pelajaran Matematika Kelas IV Sekolah Dasar. *Jurnal Pembelajaran Dan Pengajaran Pendidikan Dasar*, 3(1), 1-19.
- Ryandi, R. B., & Santri, D. D. (2024). Literature review on development of mathematics Higher Order Thinking Skills (HOTS) assessment instruments. *Journal of Curriculum and Pedagogic Studies (JCPS)*, 3(2), 57-69. <https://doi.org/10.30631/jcps.v3i2.2908>
- Ropianingsih, N. N. T., Suarjana, I. M., & Dibia, I. K. (2021). Hot Based Assessment Instrument to Assess Math Learning Achievement. *Journal for Lesson and Learning Studies*, 4(2), 271-280. <https://doi.org/10.23887/jlls.v4i2.36734>
- Saputra, T., Purnomo, E. A., & Joko, I. (2025). Peningkatan Kemampuan Higher Order Thinking Skills (HOTS) Melalui Optimaliasi Metakognitif Siswa. *Jurnal Ilmiah Matematika (JIMAT)*, 6(1), 396-408. <https://doi.org/10.63976/jimat.v6i1.854>
- Sholeha, D. F., Kusasi, M., & Bakti, I. (2023). Pengembangan Instrumen Soal Higher-Order Thinking Skills (HOTS) Berbasis Google Form untuk Mengukur Hasil Belajar Peserta Didik pada Materi Hidrolisis Garam. *JCAE (Journal of Chemistry And Education)*, 7(1), 32-48. <https://doi.org/10.20527/jcae.v7i1.2548>
- Suarti, N. W. (2022). Pengembangan instrumen penilaian tematik berbasis HOTS untuk peserta didik kelas IV sekolah dasar. *Indonesian Journal of Educational Development (IJED)*, 2(4), 534-548. <https://doi.org/10.5281/ZENODO.6178829>
- Saefullah, A. S. (2024). Ragam penelitian kualitatif berbasis kepustakaan pada studi agama dan keberagaman dalam Islam. *Al-Tarbiyah: Jurnal Ilmu Pendidikan Islam*, 2(4), 195–211. <https://doi.org/10.59059/al-tarbiyah.v2i4.1428>
- Tyas, K. E. T., Lestari, W., Subali, B., & Ellianawati, E. (2025). Analisis Tren Penggunaan Instrumen Penilaian Berbasis Digital: Systematic Literature Review (2020-2025). *Jurnal Dikdas Bantara*, 72-85.. <https://doi.org/10.32585/dikdasbantara.v8i1.6708>

- Laraswati, M. (2022). PENGEMBANGAN INSTRUMEN SOAL HOTS DENGAN VARIASI STIMULUS BERBASIS IT PADA MATERI KINEMATIKA. <http://digilib.unila.ac.id/id/eprint/65309>
- Wahidin, W. (2023). Pelaksanaan evaluasi pembelajaran berbasis higher order thinking skills (HOTS). JURNAL ILMU PENDIDIKAN & SOSIAL (SINOVA), 1(3), 107-116. <https://doi.org/10.71382/sinova.v1i3.32>
- Widodo, S., Rilianti, A. P., & Hs, A. M. (2024). Pembuatan Butir Soal SD Berbasis HOTS (High Order Thinking Skills). Jurnal Pena Karakter, 6(2), 31-36. <https://doi.org/10.62426/209aqv82>
- Winarti, W., Hairida, H., & Lestari, I. (2021). Deskripsi kemampuan guru membuat soal berdasarkan pada kurikulum 2013 di Sekolah Menengah Atas Kabupaten Landak. Jurnal Ilmiah Wahana Pendidikan, 7(2), 108-115. Landak. <https://doi.org/10.5281/zenodo.4659018>
- Zainuddin, M., Sutansi, S., & Untari, E. (2020). Pengembangan Instrumen Evaluasi Pembelajaran Tematik Berbasis HOTS (Higher order Thinking skill) dengan Penekanan Karakter. Briliant: Jurnal Riset Dan Konseptual, 5(4), 739-748. <https://doi.org/10.28926/briliant.v5i4.565>