

## Pengembangan Instrumen Tes Kemampuan Pemecahan Masalah Matematika Berbasis *Higher Order Thinking Skills* (HOTS) Siswa SMA

Asti Arifuddin\*, Baso Intang Sappaile, Mansyur  
Universitas Negeri Makassar, Makassar, Indonesia

\*Corresponding Author: [astiarifuddin29@gmail.com](mailto:astiarifuddin29@gmail.com)

Dikirim: 05-05-2026; Direvisi: 26-05-2026; Diterima: 29-05-2026

**Abstrak:** Penelitian ini bertujuan untuk mengembangkan instrumen tes matematika berbasis *Higher Order Thinking Skills* (HOTS) yang valid dan reliabel untuk mengukur kemampuan pemecahan masalah matematika peserta didik kelas XI SMA. Penelitian ini merupakan penelitian pengembangan (*research and development*) yang mengacu pada model Mardapi, yang meliputi tahap: menentukan spesifikasi instrumen tes, menulis instrument, telaah ahli, uji coba terbatas, uji coba luas, analisis, revisi, dan perakitan instrumen. Subjek penelitian terdiri atas 75 peserta didik pada uji coba terbatas dan 125 peserta didik pada uji coba diperluas. Instrumen yang dikembangkan berupa 16 butir soal pilihan ganda berbasis HOTS pada level kognitif C4, C5, dan C6 dengan materi fungsi dan lingkaran. Hasil penelitian menunjukkan bahwa instrumen memiliki validitas teoritik yang sangat baik dengan nilai CVR 0,975 dan CVI sebesar 0,98. Secara empiris, 15 dari 16 butir soal dinyatakan valid ( $p < 0,05$ ), sedangkan 1 butir soal dieliminasi. Reliabilitas instrumen sebesar 0,849 dengan kategori sangat tinggi. Tingkat kesukaran soal menunjukkan distribusi proporsional (33% mudah, 47% sedang, dan 20% sukar), serta daya pembeda berada pada kategori baik dan sangat baik. Dengan demikian, instrumen tes matematika berbasis HOTS yang dikembangkan dinyatakan valid, reliabel, dan layak digunakan untuk mengukur kemampuan pemecahan masalah matematika peserta didik.

**Kata Kunci:** HOTS; instrumen tes; pemecahan masalah matematika; validitas; reliabilitas.

**Abstract:** This study aims to develop a valid and reliable mathematics test instrument based on Higher-Order Thinking Skills (HOTS) to measure the mathematical problem-solving abilities of eleventh-grade high school students. This study employs the research and development model proposed by Mardapi, which includes the stages of determining the test instrument specifications, developing the instrument items, expert review, limited-scale trial, expanded-scale trial, analysis, revision, and instrument assembly. The subjects of the study were 75 students in the limited trials and 125 students in the extended trials. The instrument consisted of 16 multiple-choice questions designed at the cognitive levels C4, C5, and C6, covering functions and circles. The results showed that the instrument had very high theoretical validity, with CVR 0,975 and CVI values of 0,98. Empirically, 15 of the 16 items were valid ( $p < 0.05$ ), while one item was eliminated. The reliability coefficient was 0.849, categorized as very high. The difficulty levels were distributed proportionally (33% easy, 47% moderate, and 20% difficult), and the discrimination index was mostly in the good and very good categories. Therefore, the HOTS-based mathematics test instrument developed is valid, reliable, and appropriate for measuring students' mathematical problem-solving abilities.

**Keywords:** HOTS; test instrument; mathematical problem solving; validity; reliability.

## PENDAHULUAN

Perkembangan pendidikan abad ke-21 menuntut peserta didik memiliki berbagai keterampilan penting, seperti kemampuan berpikir kritis, kreatif, komunikatif, dan kolaboratif, agar mampu beradaptasi dengan kemajuan ilmu pengetahuan dan teknologi yang berlangsung sangat cepat. Salah satu kemampuan yang menjadi perhatian dalam proses pembelajaran adalah *Higher Order Thinking Skills* (HOTS). Kemampuan ini penting dimiliki peserta didik karena pembelajaran tidak lagi hanya menekankan pada kemampuan mengingat informasi, tetapi juga kemampuan mengolah informasi, menganalisis, mengevaluasi, serta menggunakan pengetahuan untuk menyelesaikan berbagai persoalan dalam kehidupan sehari-hari. HOTS menggambarkan kemampuan berpikir tingkat tinggi yang melibatkan aktivitas analisis, evaluasi, dan penciptaan solusi terhadap suatu permasalahan.

Dalam pembelajaran matematika, HOTS memiliki peranan yang sangat penting karena matematika tidak hanya berhubungan dengan penguasaan konsep dan prosedur, tetapi juga berkaitan dengan kemampuan berpikir logis, sistematis, analitis, serta kemampuan menyelesaikan masalah. Sulistyawati et al. (2023) menjelaskan bahwa kemampuan pemecahan masalah matematika berperan penting dalam pembelajaran karena mendorong peserta didik menggunakan penalaran, analisis, dan strategi penyelesaian dalam menghadapi masalah kontekstual. Kemampuan berpikir kritis dan analitis yang termasuk dalam HOTS berperan penting dalam membantu peserta didik menghadapi permasalahan akademik yang kompleks serta mempersiapkan mereka menghadapi dunia kerja yang membutuhkan kemampuan berpikir kritis (Wahyudi et al., 2024).

Kemampuan pemecahan masalah matematika merupakan kemampuan peserta didik dalam menemukan solusi terhadap suatu permasalahan melalui beberapa tahapan, yaitu memahami masalah, menyusun strategi penyelesaian, melaksanakan strategi yang dipilih, dan melakukan pemeriksaan kembali terhadap hasil yang diperoleh. Proses tersebut menunjukkan bahwa pemecahan masalah matematika memerlukan kemampuan berpikir tingkat tinggi karena peserta didik harus melakukan analisis, penalaran, dan evaluasi untuk memperoleh penyelesaian yang tepat. Pentingnya pengembangan kemampuan berpikir tingkat tinggi juga terlihat pada pelaksanaan *Programme for International Student Assessment* (PISA). OECD, 2023 menyebutkan bahwa penilaian PISA menekankan kemampuan peserta didik dalam menggunakan pengetahuan dan penalaran matematika untuk menyelesaikan masalah dalam berbagai situasi kehidupan. Namun demikian, hasil PISA 2022 menunjukkan bahwa kemampuan matematika peserta didik Indonesia masih berada dibawah rata-rata negara anggota OECD, terutama pada aspek penalaran dan pemecahan masalah. Hasil tersebut mengindikasikan bahwa kemampuan berpikir tingkat tinggi peserta didik Indonesia masih perlu ditingkatkan.

Upaya untuk meningkatkan kemampuan berpikir tingkat tinggi juga menjadi perhatian pemerintah melalui kebijakan Asesmen Nasional. Asesmen Kompetensi Minimum (AKM) dikembangkan untuk mengukur kemampuan literasi dan numerasi peserta didik melalui soal-soal kontekstual yang menuntut kemampuan bernalar dan pemecahan masalah (Rohim et al., 2021). Oleh sebab itu, proses pembelajaran dan penilaian di sekolah perlu diarahkan pada pengembangan kemampuan berpikir tingkat tinggi peserta didik. Akan tetapi, kondisi di lapangan menunjukkan bahwa penilaian matematika di sekolah masih didominasi oleh soal-soal rutin dan



prosedural yang lebih banyak mengukur kemampuan berpikir tingkat rendah (*Lower Order Thinking Skills/LOTS*). Hasil kajian Ryandi & Santri, (2024) menunjukkan bahwa instrumen evaluasi yang digunakan guru umumnya masih berorientasi pada kemampuan memahami dan belum banyak mengukur kemampuan HOTS peserta didik. Selain itu, guru juga masih mengalami kesulitan dalam menyusun soal berbasis HOTS sesuai dengan indikator kemampuan berpikir tingkat tinggi. Kondisi tersebut menunjukkan bahwa pengembangan instrumen tes berbasis HOTS menjadi kebutuhan yang penting untuk mendukung pengukuran kemampuan pemecahan masalah matematika peserta didik secara lebih tepat dan menyeluruh.

Temuan tersebut diperkuat oleh beberapa hasil penelitian terdahulu. Sulistyawati et al. (2023) mengungkapkan bahwa kemampuan pemecahan masalah matematika peserta didik masih tergolong rendah karena peserta didik mengalami kesulitan dalam memahami dan menyelesaikan soal berbasis HOTS. Penelitian Litna et al. (2021) juga menunjukkan bahwa pengembangan instrumen tes HOTS matematika masih diperlukan untuk menghasilkan alat evaluasi yang memenuhi standar kualitas tes dan mampu mengukur kemampuan berpikir tingkat tinggi secara tepat. Selain itu, penelitian Mildariah et al. (2025) menjelaskan bahwa instrumen HOTS yang valid dan reliabel dapat membantu guru memperoleh informasi yang lebih akurat mengenai kemampuan berpikir tingkat tinggi peserta didik. Gumelar & Istiyono, (2021) juga menegaskan bahwa pengembangan dan validasi instrumen HOTS menjadi bagian penting dalam mendukung asesmen matematika yang sesuai dengan tuntutan pembelajaran abad ke-21.

Berbagai penelitian mengenai HOTS dan kemampuan pemecahan masalah matematika telah banyak dilakukan. Akan tetapi, sebagian besar penelitian lebih menitikberatkan pada penerapan model pembelajaran berbasis HOTS dan analisis kemampuan peserta didik dalam menyelesaikan soal HOTS. Penelitian yang secara khusus mengembangkan instrumen tes matematika berbasis HOTS untuk mengukur kemampuan pemecahan masalah matematika siswa SMA masih relatif terbatas, terutama instrumen yang dikembangkan berdasarkan indikator pemecahan masalah matematika dan diuji secara menyeluruh dari aspek validitas, reliabilitas, tingkat kesukaran, serta daya pembeda. Padahal, kualitas instrumen asesmen memiliki peranan penting dalam menentukan ketepatan hasil evaluasi pembelajaran. Instrumen yang baik harus mampu mengukur kompetensi peserta didik secara valid dan konsisten. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan instrumen tes matematika berbasis HOTS yang valid dan reliabel untuk mengukur kemampuan pemecahan masalah matematika peserta didik SMA. Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan instrumen asesmen matematika berbasis HOTS yang sesuai dengan tuntutan pembelajaran abad ke-21 dan kebijakan AKM. Selain itu, hasil penelitian ini diharapkan dapat menjadi referensi bagi guru dalam menyusun instrumen evaluasi yang mampu mengukur kemampuan berpikir tingkat tinggi dan kemampuan pemecahan masalah matematika peserta didik secara lebih komprehensif.

## KAJIAN TEORI

### *Higher Order Thinking Skills (HOTS)*

Kemampuan berpikir tingkat tinggi atau *Higher Order Thinking Skills (HOTS)* merupakan kemampuan berpikir kompleks yang melibatkan aktivitas menganalisis,



mengevaluasi, dan menciptakan solusi terhadap suatu permasalahan. Desiriah & Setyarsih, (2021) menjelaskan bahwa HOTS merupakan kemampuan peserta didik dalam menggunakan dan menginterpretasikan pengetahuan untuk menyelesaikan permasalahan melalui proses analisis, evaluasi, hingga menghasilkan ide atau solusi baru. Dalam pembelajaran matematika, HOTS berperan penting karena matematika tidak hanya berorientasi pada penguasaan konsep, tetapi juga pada kemampuan bernalar dan menyelesaikan masalah secara sistematis. Kemampuan HOTS berkaitan dengan kemampuan peserta didik dalam mengolah informasi secara mendalam, menghubungkan berbagai konsep, serta menghasilkan solusi yang tepat terhadap suatu masalah. HOTS mencakup kemampuan berpikir kritis, penalaran, pemecahan masalah, serta kemampuan menggunakan pengetahuan dalam berbagai konteks. Dengan demikian, HOTS tidak hanya mengukur kemampuan mengingat informasi, tetapi juga kemampuan memahami, menginterpretasikan, dan menerapkan pengetahuan dalam kehidupan sehari-hari. Dalam pembelajaran matematika, HOTS dapat dilihat melalui kemampuan peserta didik dalam menyelesaikan soal-soal non-rutin yang membutuhkan penalaran logis dan strategi penyelesaian yang tepat. Saraswati & Agustika (2020) menyatakan bahwa kemampuan berpikir tingkat tinggi dalam matematika ditunjukkan melalui kemampuan menganalisis masalah, mengevaluasi strategi penyelesaian, dan menemukan solusi secara kreatif. Soal berbasis HOTS umumnya menuntut peserta didik untuk mengaitkan konsep matematika dengan situasi nyata sehingga proses berpikir yang dilakukan menjadi lebih mendalam dan bermakna. Soal-soal HOTS berada pada level kognitif menganalisis (*analyzing-C4*), mengevaluasi (*evaluating-C5*), dan mengkreasi (*creating-C6*).

Pengembangan HOTS dalam pembelajaran juga berkaitan erat dengan proses asesmen. Instrumen asesmen berbasis HOTS diperlukan agar kemampuan berpikir tingkat tinggi peserta didik dapat diukur secara tepat. Penerapan tes berbasis HOTS dapat menjadi salah satu pendekatan yang efektif dalam mengembangkan kemampuan berpikir tingkat tinggi peserta didik (Billa et al., 2024). Mildariah et al. (2025) menjelaskan bahwa instrumen HOTS yang valid dan reliabel dapat membantu guru memperoleh informasi yang lebih akurat mengenai kemampuan berpikir peserta didik. Instrumen tersebut harus dirancang dengan memperhatikan indikator kemampuan berpikir tingkat tinggi sehingga mampu mengukur kemampuan analisis, evaluasi, dan pemecahan masalah peserta didik.

Berdasarkan uraian tersebut, dapat disimpulkan bahwa HOTS merupakan kemampuan berpikir kompleks yang melibatkan aktivitas menganalisis, mengevaluasi, dan menciptakan solusi terhadap suatu masalah. Dalam pembelajaran matematika, HOTS berperan penting dalam mengembangkan kemampuan berpikir kritis, kreatif, dan kemampuan pemecahan masalah matematis peserta didik.

### **Kemampuan Pemecahan Masalah Matematika**

Kemampuan pemecahan masalah matematika merupakan salah satu kompetensi penting dalam pembelajaran matematika. Kemampuan ini menunjukkan kemampuan peserta didik dalam memahami masalah, menentukan strategi penyelesaian, melaksanakan penyelesaian, serta mengevaluasi hasil yang diperoleh. Pemecahan masalah menjadi bagian penting dalam pembelajaran matematika karena dapat membantu peserta didik mengembangkan kemampuan berpikir logis, kritis, dan sistematis. Pemecahan masalah merupakan suatu usaha untuk menemukan solusi



dari kesulitan yang dihadapi ketika jawaban tidak dapat diperoleh secara langsung. Polya mengemukakan empat langkah utama dalam pemecahan masalah, yaitu memahami masalah (*understanding the problem*), menyusun rencana penyelesaian (*devising a plan*), melaksanakan rencana (*carrying out the plan*), dan memeriksa kembali hasil penyelesaian (*looking back*). Langkah-langkah tersebut menunjukkan bahwa kemampuan pemecahan masalah melibatkan aktivitas berpikir yang sistematis dan terstruktur.

Kemampuan pemecahan masalah matematika berbasis HOTS menuntut peserta didik memiliki kemampuan berpikir kritis dan analitis dalam menyelesaikan masalah matematika (Rismawati et al., 2022). Peserta didik tidak hanya dituntut memahami konsep, tetapi juga mampu menerapkan konsep tersebut dalam situasi yang berbeda. Selain itu, kemampuan pemecahan masalah matematika juga mencerminkan kemampuan peserta didik dalam menggunakan pengetahuan matematika untuk menyelesaikan masalah kontekstual. Kemampuan pemecahan masalah menjadi salah satu indikator penting dalam penilaian matematika karena menunjukkan kemampuan peserta didik dalam menggunakan penalaran matematika pada berbagai situasi kehidupan nyata. Sulistyawati et al. (2023) menyatakan bahwa peserta didik masih mengalami kesulitan dalam menyelesaikan soal berbasis HOTS karena kurang terbiasa menghadapi soal yang membutuhkan kemampuan analisis dan penalaran. Kondisi tersebut menunjukkan bahwa pembelajaran dan asesmen matematika perlu diarahkan pada pengembangan kemampuan berpikir tingkat tinggi dan kemampuan pemecahan masalah matematika.

Kemampuan berpikir tingkat tinggi dan kemampuan pemecahan masalah matematika memiliki hubungan yang sangat erat dalam pembelajaran matematika. HOTS menuntut peserta didik untuk mampu menganalisis informasi, mengevaluasi berbagai alternatif penyelesaian, dan menghasilkan solusi yang tepat terhadap suatu masalah. Kemampuan tersebut sejalan dengan proses pemecahan masalah matematika yang melibatkan aktivitas memahami masalah, menentukan strategi penyelesaian, serta mengevaluasi hasil yang diperoleh. Soal berbasis HOTS umumnya dirancang dalam bentuk masalah non-rutin dan kontekstual sehingga mendorong peserta didik untuk berpikir kritis dan kreatif. Sulistyawati et al. (2023) menyatakan bahwa penggunaan soal HOTS dapat membantu meningkatkan kemampuan pemecahan masalah matematis peserta didik karena peserta didik dilatih untuk menggunakan penalaran dan strategi penyelesaian secara sistematis. Selain itu, pengembangan instrumen berbasis HOTS juga menjadi bagian penting dalam proses evaluasi pembelajaran matematika. Mildariah et al. (2025) menjelaskan bahwa instrumen HOTS yang valid dan reliabel dapat membantu mengukur kemampuan berpikir tingkat tinggi dan kemampuan pemecahan masalah peserta didik secara lebih akurat. Dengan demikian, pengembangan instrumen tes berbasis HOTS diperlukan agar proses asesmen mampu mengukur kompetensi peserta didik secara komprehensif sesuai dengan tuntutan pembelajaran abad ke-21.

### **Instrumen Tes Berbasis HOTS**

Instrumen tes merupakan alat yang digunakan untuk memperoleh data mengenai kemampuan atau kompetensi peserta didik secara sistematis. Dalam proses evaluasi pembelajaran, instrumen memiliki peranan penting karena kualitas hasil pengukuran sangat ditentukan oleh kualitas instrumen yang digunakan. Instrumen evaluasi merupakan alat yang digunakan untuk memperoleh informasi mengenai



hasil belajar peserta didik sehingga dapat menjadi dasar dalam pengambilan keputusan pada proses pembelajaran (Magdalena et al., 2020). Oleh karena itu, instrumen yang baik harus mampu mengukur kompetensi yang hendak diukur secara tepat dan konsisten.

Instrumen HOTS tidak hanya menilai kemampuan mengingat konsep, tetapi juga kemampuan peserta didik dalam menggunakan pengetahuan untuk menyelesaikan masalah kontekstual. Soal berbasis HOTS dirancang untuk mengukur kemampuan penalaran, berpikir kritis, dan pemecahan masalah melalui stimulus kontekstual yang berkaitan dengan kehidupan nyata. Instrumen tes yang baik harus memenuhi karakteristik validitas dan reliabilitas. Validitas menunjukkan tingkat ketepatan instrumen dalam mengukur kompetensi yang hendak diukur, sedangkan reliabilitas menunjukkan konsistensi hasil pengukuran. Selain itu, kualitas instrumen juga dapat ditinjau dari tingkat kesukaran dan daya pembeda butir soal. Tingkat kesukaran menunjukkan tingkat kemudahan atau kesulitan soal, sedangkan daya pembeda menunjukkan kemampuan soal dalam membedakan peserta didik yang memiliki kemampuan tinggi dan rendah.

### **Model Pengembangan Instrumen Mardapi**

Pengembangan instrumen dalam penelitian ini mengacu pada model pengembangan instrumen yang dikemukakan oleh Mardapi. Menurut Mardapi et al. (2011), pengembangan instrumen dilakukan melalui beberapa tahapan, yaitu penyusunan spesifikasi tes, penulisan butir soal, telaah ahli, uji coba instrumen, analisis butir soal, revisi, dan perakitan instrumen final. Tahapan tersebut bertujuan untuk menghasilkan instrumen yang memenuhi kriteria kualitas tes, baik dari aspek validitas maupun reliabilitas. Melalui proses telaah ahli dan analisis empiris, instrumen yang dikembangkan dapat disesuaikan dengan kompetensi yang hendak diukur sehingga hasil pengukuran menjadi lebih akurat. Model Mardapi juga memberikan perhatian terhadap kualitas setiap butir soal melalui analisis tingkat kesukaran, daya pembeda, dan efektivitas instrumen secara keseluruhan. Dalam penelitian ini, model pengembangan Mardapi digunakan karena sesuai dengan tujuan penelitian, yaitu mengembangkan instrumen tes matematika berbasis HOTS yang valid, reliabel, dan mampu mengukur kemampuan pemecahan masalah matematika peserta didik secara komprehensif.

## **METODE PENELITIAN**

Penelitian ini menerapkan pendekatan *Research and Development* (R&D) dengan tujuan menghasilkan instrumen tes berbasis *Higher Order Thinking Skills* (HOTS) untuk mengukur kemampuan pemecahan masalah matematika siswa SMA. Pendekatan yang digunakan adalah deskriptif kuantitatif, dengan fokus pada tahapan perancangan, pengujian, serta penyempurnaan instrumen secara sistematis. Proses pengembangan mengacu pada tahapan yang dikemukakan oleh Mardapi, mulai dari penyusunan kisi-kisi, penulisan butir soal, validasi oleh ahli, uji coba, analisis butir, hingga revisi dan penyusunan instrumen akhir.

Penelitian dilaksanakan di SMA Negeri 4 Bulukumba dalam rentang waktu empat bulan. Subjek penelitian melibatkan validator ahli serta siswa SMA yang dipilih secara purposif. Data diperoleh melalui lembar penilaian ahli, hasil pengerjaan tes oleh siswa, serta tanggapan kualitatif sebagai bahan pendukung.



Analisis data dilakukan dengan memadukan pendekatan kuantitatif dan kualitatif. Secara kuantitatif, dilakukan pengujian validitas isi menggunakan *indeks Aiken*, analisis karakteristik butir, serta pengujian reliabilitas dengan *Alpha Cronbach*. Sementara itu, data kualitatif dimanfaatkan untuk memperkaya interpretasi sekaligus menyempurnakan instrumen yang dikembangkan

## HASIL DAN PEMBAHASAN

### Proses dan Hasil Pengembangan Instrumen Tes Berbasis HOTS

Hasil pengembangan dalam penelitian ini berupa instrumen tes matematika berbasis *Higher Order Thinking Skills* (HOTS) yang telah melalui tahapan perancangan dan uji coba sehingga memenuhi kriteria validitas dan reliabilitas. Pengembangan instrumen mengacu pada langkah-langkah pengembangan instrumen menurut Mardapi yang dilakukan secara sistematis melalui beberapa tahapan, yaitu:

#### Menentukan Spesifikasi Instrumen Tes

Tahap awal dalam pengembangan instrumen adalah menyusun spesifikasi tes. Pada tahap ini ditentukan tujuan pengukuran, yaitu untuk mengukur kemampuan pemecahan masalah matematika siswa berbasis HOTS. Selanjutnya dilakukan analisis terhadap Capaian Pembelajaran (CP) dan materi matematika kelas XI SMA Kurikulum Merdeka. Hasil analisis menunjukkan bahwa materi fungsi dan lingkaran memiliki potensi yang kuat untuk dikembangkan menjadi soal-soal HOTS pada level kognitif C4 (menganalisis), C5 (mengevaluasi), dan C6 (mencipta). Pemilihan kedua materi tersebut tidak hanya didasarkan pada kesesuaiannya dengan kurikulum, tetapi juga pada kemampuannya dalam memfasilitasi proses berpikir tingkat tinggi. Temuan ini menunjukkan bahwa tidak semua materi matematika secara otomatis dapat digunakan untuk mengukur HOTS, sehingga diperlukan seleksi materi secara konseptual agar instrumen benar-benar merepresentasikan kemampuan analisis, evaluasi, dan kreasi.

Capaian Pembelajaran Matematika kelas XI Fase F menekankan bahwa peserta didik mampu menentukan fungsi invers, komposisi fungsi, dan transformasi fungsi untuk memodelkan situasi dunia nyata menggunakan fungsi yang sesuai. Selain itu, pada materi geometri peserta didik diharapkan mampu menerapkan teorema lingkaran, menentukan panjang busur, serta luas juring lingkaran untuk menyelesaikan berbagai permasalahan kontekstual. Berdasarkan hasil analisis tersebut, dirumuskan 8 tujuan pembelajaran dan 16 indikator soal. Indikator yang dikembangkan telah memenuhi karakteristik HOTS serta memuat aspek materi, konstruksi, bahasa, dan karakteristik HOTS. Dari keseluruhan indikator, terdapat 12 indikator pada level C4, 3 indikator pada level C5, dan 1 indikator pada level C6. Dominasi indikator pada level C4 menunjukkan bahwa instrumen lebih banyak mengukur kemampuan analisis sebagai fondasi berpikir tingkat tinggi. Namun demikian, keberadaan indikator pada level C5 dan C6 tetap penting sebagai representasi kemampuan evaluasi dan kreasi yang lebih kompleks. Pada tahap ini juga ditentukan bentuk instrumen yang digunakan, yaitu tes pilihan ganda. Pemilihan bentuk tes ini didasarkan pada pertimbangan efisiensi pengukuran dan kemudahan analisis hasil tes. Selain itu, penelitian ini juga bertujuan untuk membuktikan bahwa bentuk soal pilihan ganda tetap dapat digunakan untuk mengukur HOTS apabila disusun dengan stimulus kontekstual dan distraktor yang fungsional. Selanjutnya,



disusun kisi-kisi instrumen yang memuat CP, materi, tujuan pembelajaran, indikator soal, level kognitif, bentuk soal, dan nomor butir soal. Penyusunan kisi-kisi dilakukan secara sistematis agar setiap butir soal memiliki keterkaitan yang jelas dengan tujuan pengukuran.

### **Menulis Instrumen**

Berdasarkan kisi-kisi yang telah disusun, langkah selanjutnya adalah menulis butir soal. Instrumen dikembangkan dalam bentuk pilihan ganda dengan satu jawaban benar dan beberapa pengecoh (distraktor). Setiap butir soal dirancang dengan memperhatikan karakteristik HOTS, yaitu menggunakan stimulus kontekstual dan menuntut kemampuan analisis serta pemecahan masalah. Dalam penyusunannya, soal-soal dikembangkan menggunakan konteks kehidupan sehari-hari yang dekat dengan pengalaman peserta didik, seperti pertanian, perikanan, UMKM, dan pariwisata. Penggunaan konteks nyata bertujuan agar soal lebih bermakna serta mampu mendorong peserta didik menghubungkan konsep matematika dengan situasi kehidupan nyata.

### **Menelaah Instrumen (Telaah Ahli)**

Instrumen yang telah disusun kemudian ditelaah oleh lima validator ahli melalui proses *expert judgment*. Telaah dilakukan terhadap tiga aspek utama, yaitu aspek materi, konstruksi, dan bahasa. Menurut Mardapi et al. (2011), kegiatan pada telaah instrumen adalah meneliti, a) apakah butir pertanyaan sesuai dengan indikator, b) apakah bahasa yang digunakan sudah komunikatif dan menggunakan tata bahasa yang benar, dan c) apakah butir pertanyaan tidak bias, d) apakah jumlah butir sudah tepat sehingga tidak menjemukan menjawabnya. Hasil telaah menunjukkan bahwa secara umum seluruh butir soal telah sesuai dengan indikator yang diukur, memiliki konstruksi yang baik, dan menggunakan bahasa yang komunikatif. Hasil penilaian validator kemudian dianalisis menggunakan rumus Lawshe. Tahap perancangan menghasilkan draf awal instrumen tes matematika berbasis HOTS yang secara teoritis telah memenuhi kriteria kelayakan dan siap untuk diuji cobakan.

### **Tahap Uji Coba Tes**

Tahap uji coba bertujuan memperoleh bukti empiris mengenai kualitas instrumen yang dikembangkan. Uji coba dilakukan melalui dua tahap, yaitu uji coba terbatas dan uji coba diperluas.

### **Uji Coba Terbatas**

Uji coba terbatas dilakukan terhadap 75 siswa kelas XI SMA Negeri 11 Bulukumba yang memiliki karakteristik serupa dengan subjek penelitian. Pemilihan subjek dilakukan menggunakan teknik *purposive sampling* dengan bantuan guru mata pelajaran matematika. Tahap ini bertujuan untuk mengevaluasi keterbacaan soal, kejelasan bahasa, kesesuaian konteks, waktu pengerjaan, serta respons awal peserta didik terhadap instrumen. Berdasarkan hasil uji coba, sebagian besar peserta didik mampu memahami maksud soal dengan baik. Bahasa yang digunakan dinilai komunikatif dan sesuai dengan tingkat perkembangan peserta didik. Soal-soal yang menggunakan konteks kehidupan sehari-hari dinilai menarik dan membantu peserta didik memahami permasalahan yang disajikan. Sebagian besar peserta didik juga menunjukkan ketertarikan terhadap soal yang berbasis konteks nyata dan dekat dengan lingkungan mereka. Waktu pengerjaan selama 90 menit dinilai cukup untuk menyelesaikan seluruh butir soal. Meskipun demikian, beberapa peserta didik



mengaku belum terbiasa dengan tipe soal HOTS yang menuntut kemampuan analisis dan pemecahan masalah. Hal ini menunjukkan bahwa instrumen yang dikembangkan telah mengarah pada pengukuran kemampuan berpikir tingkat tinggi.

### Uji Coba Diperluas

Setelah dilakukan revisi berdasarkan hasil uji coba terbatas, instrumen kemudian diujicobakan secara lebih luas kepada 125 siswa kelas XI SMA Negeri 4 Bulukumba. Pada tahap ini digunakan 16 butir soal yang telah direvisi. Pelaksanaan uji coba dilakukan dalam kondisi yang terkontrol dengan waktu pengerjaan selama 90 menit. Data hasil pengerjaan peserta didik kemudian dianalisis untuk mengetahui validitas, reliabilitas, tingkat kesukaran, dan daya pembeda butir soal.

### Menganalisis Instrumen Butir Soal

Data hasil uji coba dianalisis secara kuantitatif untuk mengetahui kualitas instrumen. Analisis meliputi uji validitas butir soal, uji reliabilitas tes, analisis tingkat kesukaran, dan analisis daya pembeda.

### Revisi Instrumen

Berdasarkan hasil analisis butir soal, dilakukan revisi terhadap instrumen. Butir soal yang tidak memenuhi kriteria validitas dieliminasi, sedangkan butir soal yang memenuhi kriteria dipertahankan untuk tahap selanjutnya.

### Perakitan Instrumen Final

Tahap akhir pengembangan instrumen adalah merakit kembali seluruh butir soal yang telah memenuhi kriteria menjadi instrumen final. Pada tahap ini dilakukan penyempurnaan format dan penataan urutan soal sehingga diperoleh instrumen tes matematika berbasis HOTS yang siap digunakan. Secara keseluruhan, proses pengembangan instrumen dalam penelitian ini menekankan dua aspek utama, yaitu validasi secara teoritis melalui telaah ahli dan validasi empiris melalui uji coba lapangan. Kedua proses tersebut dilakukan untuk memastikan bahwa instrumen yang dihasilkan memiliki kualitas yang baik dan layak digunakan sebagai alat ukur kemampuan pemecahan masalah matematika siswa.

### Hasil Uji Validitas Teoritik dan Empiris Instrumen Tes

Validitas teoritik dalam penelitian ini dilakukan melalui penilaian para ahli (*expert judgment*) terhadap butir soal yang dikembangkan. Penilaian dilakukan untuk mengetahui kesesuaian instrumen dari aspek materi, konstruksi, dan bahasa. Hasil penilaian kemudian dianalisis menggunakan rumus Lawshe untuk memperoleh nilai *Content Validity Ratio* (CVR).

$$CVR (\text{Content Validity Ratio}) = \frac{n_e - \frac{N}{2}}{\frac{N}{2}}$$

Keterangan:

( $n_e$ ) = jumlah validator yang menyatakan butir soal esensial

( $N$ ) = jumlah validator

Berdasarkan hasil perhitungan CVR, diperoleh hasil sebagai berikut :



**Tabel 1.** Nilai *Content Validity Ratio* (CVR) Tiap Butir Soal

Jumlah Butir	Rata-rata CVR	Nilai CVI	Kategori
16	0,975	0,98	Sangat Baik

Hasil validasi ahli menunjukkan bahwa sebagian besar butir soal memperoleh nilai CVR sebesar 1,00, sedangkan satu butir soal memperoleh nilai CVR sebesar 0,60 karena terdapat satu validator yang menyatakan butir tersebut belum sepenuhnya esensial. Berdasarkan nilai CVR tiap butir soal diperoleh rata-rata nilai CVR sebesar 0,975. Nilai CVI yang diperoleh adalah sebesar 0,98 dengan kategori sangat baik. Temuan ini menunjukkan bahwa instrumen yang dikembangkan telah memenuhi validitas isi dan telah merepresentasikan konstruk kemampuan pemecahan masalah matematika berbasis HOTS secara memadai.

Meskipun demikian, validitas teoritik saja belum cukup untuk menjamin kualitas instrumen. Oleh karena itu, dilakukan uji validitas empiris berdasarkan data hasil uji coba lapangan dengan menggunakan bantuan aplikasi JASP versi 0.96. Kriteria pengujian yang digunakan adalah apabila nilai signifikansi (*p-value*) < 0,05 maka butir soal dinyatakan valid.

**Tabel 2.** Hasil Analisis Validitas Empiris

Jumlah Butir Valid	Jumlah Butir Tidak Valid
15	1

Hasil analisis menunjukkan bahwa dari 16 butir soal yang diujicobakan, sebanyak 15 butir soal memiliki nilai signifikansi < 0,001 sehingga dinyatakan valid. Sementara itu, terdapat satu butir soal yang memiliki nilai signifikansi sebesar 0,070 sehingga dinyatakan tidak valid. Butir soal yang tidak valid yaitu butir soal 13 kemudian dieliminasi dan tidak digunakan pada tahap analisis selanjutnya. Dengan demikian, jumlah butir soal yang digunakan pada tahap berikutnya adalah sebanyak 15 butir soal. Tingginya validitas empiris instrumen menunjukkan bahwa konstruksi butir soal telah sesuai dengan indikator kemampuan yang diukur. Validitas empiris merupakan bukti bahwa suatu instrumen mampu mengukur konstruk yang hendak diukur berdasarkan data hasil uji coba lapangan (Santosa et al., 2025). Selain itu, Pratiwi et al. (2024) menyatakan bahwa validitas yang tinggi dapat dicapai melalui proses validasi ahli yang dilakukan secara sistematis dan ketat. Dengan demikian, instrumen tes berbasis HOTS yang dikembangkan telah memenuhi kriteria validitas teoritik dan empiris sehingga layak digunakan untuk mengukur kemampuan pemecahan masalah matematika siswa.

### Analisis Reliabilitas Instrumen Tes

Uji reliabilitas dilakukan terhadap 15 butir soal yang telah memenuhi kriteria validitas empiris. Perhitungan reliabilitas menggunakan koefisien *Alpha Cronbach* dengan bantuan aplikasi JASP versi 0.96.

**Tabel 3.** Hasil Uji Reliabilitas Butir Soal

Frequentist Scale Reliability Statistics				
			95% CI	
Coefficient	Estimate	Std. Error	Lower	Upper
Cronbach's $\alpha$	0.849	0.018	0.814	0.884

Note. Of the observations, 125 complete cases were used.

Hasil analisis menunjukkan bahwa instrumen memperoleh nilai koefisien reliabilitas sebesar 0,849. Nilai tersebut berada pada rentang 0,80–1,00 yang



termasuk dalam kategori sangat tinggi. Nilai reliabilitas yang tinggi menunjukkan bahwa instrumen memiliki konsistensi internal yang sangat baik dalam mengukur kemampuan pemecahan masalah matematika siswa. Artinya, setiap butir soal dalam instrumen memiliki keterkaitan yang kuat dalam mengukur konstruk yang sama.

Cahyani et al. (2024) menyatakan bahwa Instrumen dinyatakan memiliki reliabilitas sangat tinggi apabila nilai *Alpha Cronbach* berada pada rentang 0,80–1,00. Selain itu, Pratiwi et al. (2024) juga menyatakan bahwa instrumen HOTS yang dikembangkan berdasarkan indikator yang jelas dan terstruktur cenderung menghasilkan reliabilitas yang tinggi. Tingginya reliabilitas instrumen menunjukkan bahwa instrumen yang dikembangkan memiliki stabilitas dan konsistensi pengukuran yang baik sehingga layak digunakan sebagai alat ukur kemampuan pemecahan masalah matematika siswa.

### **Analisis Tingkat Kesukaran Butir Soal**

Analisis tingkat kesukaran dilakukan untuk mengetahui tingkat kemudahan atau kesulitan setiap butir soal. Tingkat kesukaran dihitung berdasarkan proporsi peserta didik yang menjawab benar pada setiap butir soal. Berdasarkan hasil analisis, diperoleh 5 butir soal berkategori mudah, 7 butir soal berkategori sukar soal, yaitu 30% mudah, 50% sedang, dan 20% sukar. Hal ini menunjukkan bahwa instrumen yang dikembangkan telah memiliki keseimbangan tingkat kesukaran yang baik. Temuan ini sejalan dengan pendapat Usman (2021) yang menyatakan bahwa instrumen yang baik perlu memiliki distribusi tingkat kesukaran yang proporsional agar mampu mengukur kemampuan peserta didik secara menyeluruh. Salah satu proporsi yang umum digunakan adalah 3:5:2 untuk kategori mudah, sedang, dan sukar. Soal yang baik adalah soal yang tidak terlalu mudah dan tidak terlalu sukar.

Distribusi tingkat kesukaran yang seimbang menunjukkan bahwa instrumen mampu mengakomodasi variasi kemampuan peserta didik, baik pada tingkat rendah, sedang, maupun tinggi. Dengan demikian, instrumen yang dikembangkan layak digunakan untuk mengukur kemampuan pemecahan masalah matematika secara lebih komprehensif.

### **Analisis Daya Pembeda Butir Soal**

Daya pembeda merupakan kemampuan suatu butir soal dalam membedakan peserta didik yang memiliki kemampuan tinggi dan rendah. Dalam penelitian ini, perhitungan daya pembeda dilakukan dengan membandingkan kelompok atas dan kelompok bawah yang masing-masing ditentukan sebesar 27% dari jumlah peserta tes. Berdasarkan hasil analisis, diperoleh 3 butir soal berkategori sangat baik, 8 butir soal berkategori baik, dan 4 butir soal berkategori cukup. Tidak terdapat butir soal yang termasuk kategori jelek. Hasil tersebut menunjukkan bahwa secara umum instrumen yang dikembangkan telah memiliki kemampuan yang baik dalam membedakan peserta didik berdasarkan tingkat kemampuannya. Butir soal yang berada pada kategori baik dan sangat baik menunjukkan bahwa soal mampu mengidentifikasi perbedaan kemampuan siswa secara efektif. Hal ini sejalan dengan penelitian Nurhayati et al. (2024) yang menyatakan bahwa daya pembeda merupakan kemampuan suatu butir soal dalam membedakan peserta didik yang memiliki kemampuan tinggi dan rendah. Butir soal dengan daya pembeda yang baik menunjukkan bahwa instrumen mampu mengidentifikasi perbedaan kemampuan peserta didik secara efektif. Hasil penelitian ini juga didukung oleh (Kurnia et al., 2022) yang menyatakan bahwa instrumen HOTS mampu mengukur kemampuan



berpikir tingkat tinggi peserta didik melalui analisis validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Demikian juga Febrianti et al. (2021) menyatakan bahwa instrumen HOTS yang memiliki daya beda yang baik mampu membedakan peserta didik berdasarkan tingkat kemampuan berpikirnya. Meskipun terdapat beberapa butir soal yang berada pada kategori cukup, butir-butir tersebut masih dapat digunakan dengan melakukan revisi pada aspek tertentu agar daya pembedanya meningkat. Secara keseluruhan, hasil analisis menunjukkan bahwa instrumen tes berbasis HOTS yang dikembangkan telah memenuhi kriteria validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Dengan demikian, instrumen tersebut layak digunakan sebagai alat ukur kemampuan pemecahan masalah matematika siswa kelas XI SMA. Selain relevan digunakan dalam penelitian, instrumen ini juga berpotensi dimanfaatkan oleh guru sebagai alat evaluasi pembelajaran yang mampu mengukur kemampuan berpikir tingkat tinggi peserta didik secara lebih optimal.

## KESIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa instrumen tes matematika berbasis *Higher Order Thinking Skills* (HOTS) yang dikembangkan telah memenuhi kriteria instrumen yang baik dan layak digunakan untuk mengukur kemampuan pemecahan masalah matematika peserta didik SMA. Instrumen yang dikembangkan dinyatakan valid, reliabel, serta mampu mengukur kemampuan berpikir tingkat tinggi peserta didik sesuai dengan indikator HOTS pada level analisis, evaluasi, dan kreasi. Pengembangan instrumen tes berbasis HOTS ini diharapkan dapat menjadi alternatif bagi guru dalam melaksanakan evaluasi pembelajaran matematika yang lebih bermakna dan sesuai dengan tuntutan pembelajaran abad ke-21. Selain itu, instrumen yang dihasilkan juga dapat mendukung pelaksanaan asesmen yang berorientasi pada pengembangan kemampuan berpikir kritis dan pemecahan masalah matematis peserta didik.

## DAFTAR PUSTAKA

- Billa, S., Purnamasari, R., & Suchyadi, Y. (2024). Pengembangan Instrumen Tes Pilihan Ganda Berbasis HOTS menggunakan Aplikasi Quizizz dan QR-Code pada Pembelajaran Matematika Materi Pecahan. *Didaktik: Jurnal Ilmiah PGSD FKIP Universitas Mandiri*, 10(4), 335–343.
- Cahyani, F. D., Sari, N. K., & Pujiyana. (2024). Hubungan Interaksi Teman Sebaya dengan Motivasi Belajar Siswa di Sekolah Dasar. *PENDIKDAS: Jurnal Pendidikan Sekolah Dasar*, 05(02), 6–14.
- Desiriah, E., & Setyarsih, W. (2021). Tinjauan Literatur Pengembangan Instrumen Penilaian Kemampuan Berpikir Tingkat Tinggi (HOTS) Fisika di SMA. *Orbita: Jurnal Hasil Kajian, Inovasi, Dan Aplikasi Pendidikan Fisika*, 7(1), 79–89.
- Febrianti, S. A. D., Widiana, I. W., & Yudiana, K. (2021). Higher-Order Thinking Skill ( HOTS ) Instrument-Based Cognitive Evaluation in Grade V Elementary School Students. *Thinking Skills and Creativity Journal*, 4(2), 48–56.



- Gumelar, A., & Istiyono, E. (2021). Development And Validation Instrument To Measure Teacher Knowledge About HOTS Mathematics Assessment. *Isted*, 1(1), 1–7. <https://doi.org/10.4108/eai.17-7-2021.2312394>.
- 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 Pendahuluan. *Jurnal Pendidikan Sains Indonesia*, 10(1), 176–190. <https://doi.org/10.24815/jpsi.v10i1.21727>.
- Litna, K. O., Mertasari, N. M. S., & Sudirtha, G. (2021). Pengembangan Instrumen Tes Higher Order Thinking Skills (HOTS) Matematika SMA Kelas X. *Jurnal Penelitian Dan Evaluasi Pendidikan Indonesia*, 11(1), 10–20.
- Magdalena, I., Fauzi, H. N., & Putri, R. (2020). Pentingnya Evaluasi Dalam Pembelajaran dan Akibat Memanipulasinya. *Jurnal Pendidikan Sains*, 2(2), 244–257. <https://ejournal.stitpn.ac.id/index.php/bintang>.
- Mardapi, D., Kumaidi, & Kartowagiran, B. (2011). Pengembangan Instrumen Pengukur Hasil Belajar Nirbias dn Terskala Baku. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 15(2), 326–341.
- Mildariah, K., Maimunah, & Anggraini, R. D. (2025). Development of Instruments of Higher Order Thinking Skills ( HOTS ) for Phase D Students. *Jurnal Gantang*, 10(2), 291–301. <https://doi.org/10.31629/jg.v10i2.7387> Abstract
- Nurhayati, S., Halini, & Indriani, T. (2024). Analisis Daya Pembeda Butir Pada Materi Persamaan dan Pertidaksamaan Nilai Mutlak dengan Metode Teori.pdf. *Jurnal MathEdu (Mathematic Education Journal) Vol. 7 No. 1 Maret 2024 Analisis*, 7(1), 57–63. <http://journal.ipts.ac.id/index.php/MathEdu>.
- OECD. (2023). *PISA 2022 Results (Volume 1): The State of Learning and Equity in Education, PISA: Vol. I*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>.
- Pratiwi, U. M., Yuhana, Y., Pujiastuti, H., & Novaliyosi. (2024). Pengembangan Instrumen Literasi Numerasi Berbasis HOTS pada Materi Aritmatika Sosial. *Jurnal Riset Pembelajaran Matematika*, 6(2), 81–88. <http://journal.unirow.ac.id/index.php/jrpm>.
- Rismawati, M., Rahmawati, P., & Rindiani, A. B. (2022). Analisis Kemampuan Berpikir Tingkat Tinggi dalam Pemecahan Masalah Matematika Berbasis Higher Order Thinking Skill ( HOTS ). *Jurnal Cendekia: Jurnal Pendidikan Matematik*, 06(02), 2134–2143.
- Rohim, D. C., Rahmawati, S., & Ganestri, I. D. (2021). Konsep Asesmen Kompetensi Minimum Meningkatkan Kemampuan Literasi Numerasi Sekolah Dasar untuk Siswa. *Jurnal Varidika*, 33(1), 54–62. <https://doi.org/10.23917/varidika.v33i1.14993>.
- 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://ejournal.lp2m.uinjambi.ac.id/ojp/index.php/jcps>.



- Santosa, Y. T., Maulida, D. W., Ferdianto, J., Sutarni, S., & Hidayati, Y. M. (2025). Evaluating the Quality of Mid-Semester Mathematics Summative Assessment in Secondary School: A Psychometric Analysis of Test Items. *IJECA (International Journal of Education & Curriculum Application)*, 8(3), 431–449. <https://doi.org/10.31764/ijeca.v8i3.36276>.
- Saraswati, P. M. S., & Agustika, G. N. S. (2020). Kemampuan Berpikir Tingkat Tinggi Dalam Menyelesaikan Soal HOTS Mata Pelajaran Matematika. *Jurnal Ilmiah Sekolah Dasar*, 4(2), 257–269. <https://ejournal.undiksha.ac.id/index.php/JISD/index%0AKemampuan>.
- Sulistiyawati, D., Hadi, W., Hidayat, A., & Muhammad, R. R. (2023). The Impact of Problem Based Learning Augmented With HOTS Problems on Students Mathematical Problem Solving Abilities. *Kalamatika: Jurnal Pendidikan Matematika*, 8(2), 219–236. <https://doi.org/10.22236/KALAMATIKA.vol8no2.2023pp219-236>.
- Usman, M. (2021). Kualitas Soal Ujian Akhir Sekolah Mata Fisika SMA Negeri Se-Kota Baubau. *Syntax Idea*, 3(10), 2171–2185. <https://doi.org/10.36418/syntax-idea.v3i10.1527>.
- Wahyudi, H., Azhar, M., & Yusra, N. (2024). Research Trends On High Order Thinking Skills ( HOTS ) In Arabic Language Education Journals In Indonesia Tren Penelitian High Order Thinking Skills ( HOTS ) Pada Jurnal Pendidikan Bahasa Arab Di Indonesia. *Ijazarabi: Journal of Arabic Learning*, 7(3), 1093–1104. <https://doi.org/10.18860 /ijazarabi. V7i3.24752>.

