

Pengembangan Soal Hots Bangun Datar Untuk Mengukur Nalar Dalam Pemecahan Masalah Matematika di Sekolah Dasar

Asep Yogie Sholihin^{1*}, Afaz Maarif Muslim², Aida Nurasyipha³, Syifa Rizky Auliya⁴, Hafiziani Eka Putri⁵, Teten Ginanjar Rahayu⁶

^{1,2,3,4,5,6}Program Studi Pendidikan Guru Sekolah Dasar, Universitas Pendidikan Indonesia, Jawa Barat, Indonesia

Email Corresponding Author: asepyogies@upi.edu

Info Artikel

Article history:

Received : 19/05/2026

Revised : 04/06/2026

Accepted : 06/06/2026

Published : 17/06/2026

Kata-kata kunci:

HOTS, bangun datar, penalaran matematis, pemecahan masalah, sekolah dasar, pengembangan instrumen

ABSTRAK

Penelitian ini bertujuan untuk mengembangkan instrumen soal berbasis Higher Order Thinking Skills (HOTS) pada materi bangun datar untuk mengukur kemampuan penalaran dan pemecahan masalah matematika siswa sekolah dasar. Penelitian menggunakan metode Research and Development (R&D) dengan model pengembangan 4D yang meliputi tahap define, design, develop, dan disseminate, namun penelitian dibatasi hingga tahap develop. Instrumen yang dikembangkan berupa 15 soal pilihan ganda berbasis HOTS yang disusun berdasarkan indikator penalaran (reasoning), pemecahan masalah (problem solving), dan pembuktian (proof) mengacu pada National Council of Teachers of Mathematics (NCTM) serta Taksonomi Bloom revisi Anderson dan Krathwohl pada level kognitif C4 dan C5. Validasi awal instrumen dilakukan oleh satu dosen ahli pendidikan matematika untuk memperoleh masukan terkait kesesuaian indikator, materi, dan konstruksi soal, kemudian dilanjutkan dengan uji coba kepada siswa kelas IV sekolah dasar. Hasil penelitian menunjukkan bahwa instrumen memperoleh persentase validitas sebesar 85,33% dengan kategori sangat valid. Hasil analisis reliabilitas menunjukkan koefisien sebesar 0,66 dengan kategori baik. Analisis validitas butir menunjukkan bahwa 7 dari 15 butir soal memenuhi kriteria valid, sedangkan butir lainnya memerlukan revisi lebih lanjut. Tingkat kesukaran soal didominasi kategori sedang, daya pembeda sebagian besar berada pada kategori baik hingga sangat baik, serta efektivitas pengecoh menunjukkan bahwa sebagian besar distraktor berfungsi dengan baik. Dengan demikian, instrumen yang dikembangkan menunjukkan kelayakan awal sebagai alat ukur kemampuan penalaran dan pemecahan masalah matematika siswa sekolah dasar, namun beberapa butir soal masih memerlukan revisi untuk meningkatkan kualitas pengukuran.

1. PENDAHULUAN

Kemampuan berpikir tingkat tinggi atau Higher Order Thinking Skills (HOTS) dan keterampilan pemecahan masalah merupakan kompetensi krusial yang harus dikuasai oleh siswa pada abad ke-21. Namun, realitas di lapangan menunjukkan bahwa tingkat keterampilan matematika siswa di Indonesia masih perlu banyak ditingkatkan. Hal ini tergambar jelas dari rilis hasil studi internasional Programme for International Student Assessment (PISA) 2022, di mana Indonesia menempati peringkat ke-69 secara global untuk literasi matematika (Kemendikbudristek, 2023). Rendahnya capaian skor PISA ini mengindikasikan bahwa

DOI: <https://doi.org/10.63976/jimat.v7i2.1398>

sebagian besar siswa di jenjang pendidikan dasar hingga menengah masih mengalami kesulitan besar ketika dihadapkan pada soal-soal non-rutin yang menuntut kemampuan penalaran logis, analisis konteks, dan strategi pemecahan masalah. Menanggapi permasalahan berskala nasional dan internasional tersebut, peneliti memandang bahwa perbaikan kualitas instrumen penilaian di tingkat Sekolah Dasar (SD) merupakan langkah awal yang sangat mendesak. Siswa SD perlu dibiasakan sejak dini untuk menghadapi tantangan berpikir tingkat tinggi, terutama pada materi yang konsepnya aplikatif dan sering dijumpai dalam kehidupan sehari-hari, seperti materi bangun datar. Proses berpikir kreatif merupakan salah satu aspek yang penting dikembangkan siswa agar mereka tidak sekadar menghafal rumus, melainkan mampu mengkonstruksi pemahaman (Hidajat, F. A. 2022). Pengenalan dan penggunaan evaluasi berbasis HOTS pada materi geometri dasar sangat penting untuk merangsang nalar siswa agar lebih adaptif dalam memecahkan masalah spasial dan logis.

Berbagai penelitian sebelumnya telah berupaya menyoroti peningkatan kemampuan pemecahan masalah siswa SD. Menurut Safirah dan Abdillah (2024), model Problem Based Learning (PBL) memberikan pengaruh yang signifikan terhadap kemampuan pemecahan masalah matematika siswa SD, ditunjukkan oleh peningkatan kemampuan siswa dalam memahami, menganalisis, dan menyelesaikan masalah matematika selama proses pembelajaran. Sejalan dengan hal tersebut, penelitian Khairani, Sukmawati, dan Nasrun (2023) membuktikan bahwa perlakuan pembelajaran tertentu dapat meningkatkan daya nalar matematis anak pada rentang usia kelas tinggi SD. Lebih spesifik pada bidang geometri, penelitian Ristanty dan Pratama (2022) menyimpulkan bahwa kemampuan siswa dalam memecahkan masalah bangun datar segiempat sangat dipengaruhi oleh tahapan berpikir geometri mereka. Meskipun berbagai penelitian telah banyak berfokus pada inovasi model pembelajaran dan media pembelajaran (Sumilat, Kumolontang, & Rompah, 2022), masih sedikit penelitian yang secara khusus berfokus pada upaya pengembangan instrumen tes berupa soal HOTS yang tervalidasi secara empiris untuk mengukur nalar dan pemecahan masalah bangun datar di jenjang SD. Gap penelitian inilah yang memunculkan kebaruan dari penelitian ini. Penelitian ini menawarkan kebaruan dengan merancang instrumen penilaian HOTS matematika yang secara spesifik mengukur indikator penalaran (reasoning) dan pemecahan masalah (problem solving) pada materi bangun datar kelas IV SD, dengan bersandar pada Taksonomi Bloom revisi Anderson dan Krathwohl pada tataran kognitif C4 (menganalisis) dan C5 (mengevaluasi). Produk soal yang dikembangkan tidak hanya sekadar alat uji, melainkan instrumen diagnostik yang dapat memetakan kesulitan nalar geometri siswa berdasarkan teori perkembangan kognitif dan teori Van Hiele mengenai level berpikir geometri pada anak.

Berdasarkan hal tersebut, penelitian ini difokuskan pada pengembangan dan pengujian kualitas instrumen penilaian matematika berbentuk soal pilihan ganda pada jenjang sekolah dasar. Penelitian ini bertujuan untuk mendeskripsikan proses pengembangan instrumen serta menguji tingkat validitas dan reliabilitas soal berbasis Higher Order Thinking Skills (HOTS) pada materi bangun datar sebagai alat ukur kemampuan penalaran dan pemecahan masalah siswa sekolah dasar. Instrumen yang dikembangkan diharapkan dapat menjadi alternatif alat evaluasi yang berkualitas bagi guru dalam melaksanakan penilaian pembelajaran matematika, sekaligus mendukung pembiasaan siswa dalam menghadapi asesmen literasi dan numerasi yang berorientasi pada standar evaluasi global.

2. METODE PENELITIAN

Penelitian ini menggunakan metode *Research and Development* (R&D) dengan model pengembangan 4D (*Four-D Model*) yang dikembangkan oleh Sivasailam Thiagarajan, Dorothy S. Semmel, dan Melvyn I. Semmel (1974). Bagian ini menjelaskan mengenai rancangan

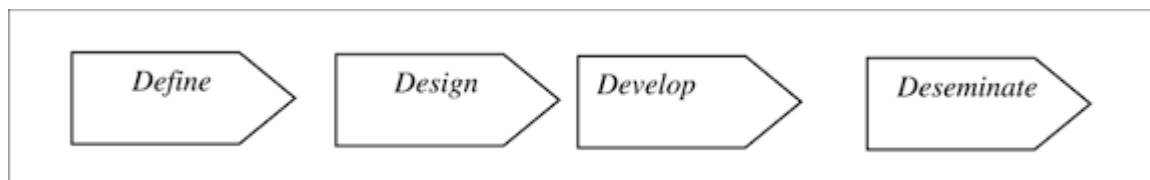
DOI: <https://doi.org/10.63976/jimat.v7i2.1398>

penelitian, subjek penelitian, instrumen yang digunakan, prosedur pengumpulan data, serta teknik analisis data. Penjelasan metode penelitian disusun secara sistematis agar memberikan gambaran yang jelas mengenai langkah-langkah penelitian dalam mengembangkan soal HOTS materi bangun datar sebagai alat ukur kemampuan penalaran dan pemecahan masalah siswa sekolah dasar, sebagai berikut:

Desain Penelitian

Penelitian ini menggunakan metode *Research and Development* (R&D) dengan model pengembangan 4D. Model 4D dipilih karena memiliki langkah-langkah sistematis dalam mengembangkan produk pendidikan, termasuk instrumen penilaian berupa soal HOTS. Selain itu, model ini sesuai untuk menghasilkan produk yang valid, praktis, dan memiliki kualitas empiris yang baik.

Gambar 1. Model Pengembangan 4D



Tahapan model 4D terdiri atas *Define*, *Design*, *Develop*, dan *Disseminate*. Namun, penelitian ini dibatasi sampai tahap *Develop*, karena tujuan penelitian difokuskan pada pengembangan dan pengujian kualitas soal HOTS materi bangun datar. Tahap *Define* dilakukan melalui analisis kebutuhan, kurikulum, karakteristik siswa, dan materi. Tahap *Design* meliputi penyusunan kisi-kisi, indikator, dan rancangan butir soal. Tahap *Develop* dilakukan melalui validasi ahli, revisi, uji coba lapangan, dan analisis hasil uji coba.

Subjek dan Sampel

Subjek penelitian ini terdiri atas validator ahli dan siswa sekolah dasar. Validator ahli berjumlah satu orang, yaitu dosen ahli pendidikan matematika. Validasi dilakukan sebagai validasi awal untuk memperoleh masukan terkait kesesuaian indikator, materi, konstruksi soal, dan aspek kebahasaan sebelum instrumen diuji cobakan. Validator berperan dalam menilai kelayakan soal dari aspek isi, konstruksi, bahasa, dan kesesuaian indikator HOTS. Populasi penelitian adalah seluruh siswa kelas IV sekolah dasar yang telah mempelajari materi bangun datar. Sampel penelitian dipilih menggunakan teknik *purposive sampling*, yaitu satu kelas IV pada salah satu sekolah dasar dengan jumlah sekitar 25-30 siswa. Pemilihan kelas IV didasarkan pada pertimbangan bahwa siswa telah memperoleh materi bangun datar serta memiliki kesiapan dalam mengerjakan soal berbasis penalaran dan pemecahan masalah.

Instrumen Penelitian

Instrumen yang digunakan pada penelitian ini meliputi:

1. Lembar Validasi Ahli/Judgement Expert
Digunakan untuk memperoleh penilaian dari validator terhadap kualitas soal yang dikembangkan. Aspek yang dinilai meliputi kesesuaian materi, konstruksi soal, penggunaan bahasa, dan indikator HOTS.
2. Tes Soal HOTS Bangun Datar
Berupa seperangkat soal pilihan ganda berjumlah 15 yang dikembangkan untuk mengukur kemampuan penalaran dan pemecahan masalah siswa.

3. Dokumentasi

Digunakan untuk memperoleh data pendukung berupa daftar siswa dan hasil pelaksanaan uji coba.

Teknik atau Prosedur Pengumpulan Data

1. Tahap *Define*

Mengumpulkan informasi melalui studi kurikulum, analisis materi bangun datar, identifikasi kebutuhan soal HOTS di sekolah dasar, serta menentukan indikator penalaran dan pemecahan masalah berdasarkan standar National Council of Teachers of Mathematics.

2. Tahap *Design*

Menyusun kisi-kisi soal berdasarkan indikator HOTS serta indikator penalaran dan pemecahan masalah mengacu pada National Council of Teachers of Mathematics, kemudian merancang butir soal HOTS.

3. Tahap *Develop*

- a. Soal divalidasi oleh para ahli menggunakan lembar validasi/Judgement Expert.
- b. Melakukan revisi berdasarkan saran validator.
- c. Soal yang telah direvisi diuji cobakan kepada siswa kelas IV sekolah dasar.
- d. Mengumpulkan hasil jawaban siswa untuk dianalisis dengan ANATES.

Teknik Analisis Data

Data dalam penelitian ini dianalisis secara kualitatif dan kuantitatif.

1. Analisis Kualitatif

Data berupa saran, komentar, dan masukan dari validator dianalisis secara deskriptif sebagai dasar revisi soal.

2. Analisis Kuantitatif

Skor hasil validasi ahli dihitung menggunakan rata-rata persentase untuk menentukan tingkat kelayakan soal. Data hasil uji coba siswa dianalisis menggunakan ANATES untuk mengetahui kualitas butir soal yang meliputi reliabilitas, tingkat kesukaran, dan daya pembeda

Analisis Validasi Instrumen Soal

Validasi instrumen dilakukan untuk mengetahui kelayakan soal HOTS sebelum diuji cobakan kepada siswa. Validasi dilakukan oleh validator ahli yang terdiri dari dosen pendidikan matematika dengan menilai aspek materi, konstruksi soal, bahasa, dan kesesuaian indikator HOTS. Karena validasi melibatkan satu validator, hasil yang diperoleh digunakan sebagai dasar validasi awal instrumen dan revisi produk sebelum pelaksanaan uji coba lapangan. Menurut Sugiyono, validasi ahli digunakan untuk menilai apakah instrumen penelitian telah sesuai dengan tujuan pengukuran dan layak digunakan dalam penelitian (Sugiyono, 2019). Penilaian validator dihitung menggunakan persentase kelayakan dengan rumus:

$$\text{Persentase skor} = \frac{\text{Total skor jawaban validator}}{\text{Jumlah skor maksimal}} \times 100\%$$

Tabel 1. Kriteria Validasi Instrumen

Persentase	Kategori
81% - 100%	Sangat Valid
61% - 80%	Valid
41% - 60%	Cukup Valid

21% - 40%	Kurang Valid
0% - 20%	Tidak Valid

Sumber: Sugiyono (2019)

Instrumen dinyatakan layak digunakan apabila memperoleh kategori valid atau sangat valid.

Analisis Butir Soal

Analisis ini dilakukan menggunakan aplikasi Anates. Melalui analisis tersebut dapat diketahui kualitas setiap butir soal yang meliputi validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Hasil analisis digunakan sebagai dasar untuk menentukan kelayakan soal HOTS yang dikembangkan. Menurut Suharsimi Arikunto, instrumen yang baik harus memenuhi beberapa kriteria, di antaranya valid, reliabel, memiliki tingkat kesukaran yang sesuai, serta mampu membedakan kemampuan siswa. Hasil analisis tersebut digunakan untuk menentukan apakah soal layak digunakan sebagai instrumen pengukur kemampuan penalaran dan pemecahan masalah siswa sekolah dasar.

1. Validitas Soal

Validitas digunakan untuk mengetahui sejauh mana suatu butir soal mampu mengukur apa yang seharusnya diukur. Semakin tinggi nilai validitas suatu soal, maka semakin baik soal tersebut dalam mengukur kemampuan peserta didik sesuai tujuan pembelajaran. Validitas umumnya ditunjukkan melalui nilai korelasi antara skor butir soal dengan skor total tes.

Tabel 2. Validitas Soal

Indeks Korelasi	Kriteria
0,80-1,00	Sangat Tinggi
0,60- 0,79	Tinggi
0,40-0,59	Cukup
0,20-0,39	Rendah
0,00-0,19	Sangat Rendah

Sumber: Riduwan (2013)

2. Reliabilitas Soal

Reliabilitas digunakan untuk mengetahui tingkat konsistensi instrumen tes. Instrumen yang reliabel akan memberikan hasil yang relatif tetap apabila digunakan pada kelompok dengan karakteristik yang sama.

Tabel 3. Reliabilitas Soal

Interval Koefisien	Tingkat Hubungan
0,00-0,19	Sangat Buruk
0,20-0,39	Buruk

0,40-0,59	Cukup
0,60-0,79	Baik
0,80-1,00	Sangat Baik

Sumber: Arikunto (2018)

3. Tingkat Kesukaran

Analisis tingkat kesukaran dilakukan untuk mengetahui kategori soal, apakah termasuk mudah, sedang, atau sukar. Tingkat kesukaran yang baik akan membantu instrumen dalam mengukur kemampuan siswa secara proporsional.

Tabel 4. Tingkat Kesukaran Soal

Kriteria	Tingkat Kesukaran
0-15%	Sangat Sukar
16%-30%	Sukar
31%-70%	Sedang
71%-85%	Mudah
86-100%	Sangat Mudah

Sumber: Arikunto (2018)

4. Daya Pembeda

Daya pembeda digunakan untuk mengetahui kemampuan suatu butir soal dalam membedakan siswa yang memiliki kemampuan tinggi dan siswa yang memiliki kemampuan rendah.

Tabel 5. Daya Pembeda

Kriteria	Tingkat Kesukaran
<9%	Sangat Buruk
10-19%	Buruk
20-29%	Cukup
30-49%	Baik
>50%	Sangat Baik

Sumber: Arikunto (2018)

5. Efektivitas Pengecoh

Efektivitas pengecoh dilakukan untuk mengetahui apakah pilihan jawaban pengecoh pada soal pilihan ganda telah berfungsi dengan baik. Distraktor dikatakan berfungsi apabila dipilih oleh siswa, terutama siswa yang memiliki kemampuan rendah.

Tabel 6. Efektivitas Pengecoh

Pengecoh Berfungsi	Kriteria
3	Sangat Baik
2	Baik
1	Kurang Baik
0	Tidak Baik

Sumber: Basri et all (2021)

3. HASIL DAN PEMBAHASAN

Hasil Penelitian

Pada bagian ini disajikan hasil penelitian mengenai pengembangan instrumen tes yang telah melalui tahap validasi ahli dan uji coba instrumen. Hasil penelitian meliputi analisis kualitatif dan kuantitatif terhadap instrumen yang dikembangkan. Analisis kualitatif diperoleh dari kritik dan saran validator, sedangkan analisis kuantitatif diperoleh dari hasil uji validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh pada butir soal.

Hasil Analisis Kualitatif Validator

Validasi ahli dilakukan untuk memperoleh masukan dan saran terhadap instrumen yang dikembangkan. Data kualitatif diperoleh dari kritik dan saran validator mengenai aspek materi, bahasa, dan penyajian soal. Berdasarkan hasil validasi ahli, diperoleh beberapa saran dan masukan terhadap instrumen yang dikembangkan. Validator menyarankan perbaikan pada redaksi beberapa soal agar lebih jelas dan mudah dipahami peserta didik. Selain itu, terdapat saran untuk memperbaiki distraktor pada beberapa butir soal agar lebih berfungsi dengan baik. Validator juga menyarankan agar soal dilengkapi dengan ilustrasi dan gambar pendukung untuk meningkatkan daya tarik serta membantu peserta didik memahami konteks soal. Seluruh saran dari validator kemudian digunakan sebagai dasar revisi instrumen sebelum dilakukan uji coba. Setelah dilakukan revisi sesuai masukan validator, instrumen dinyatakan layak untuk diuji cobakan.

Hasil Analisis Kuantitatif Instrumen

Analisis kuantitatif dilakukan terhadap hasil uji coba instrumen untuk mengetahui kualitas butir soal yang dikembangkan. Analisis tersebut meliputi uji validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh.

Hasil Uji Validasi

Berdasarkan hasil penilaian validator terhadap 15 aspek penilaian dengan rentang skor 1-5, diperoleh total skor sebesar 64 dari skor maksimal 75. Persentase validitas dihitung menggunakan rumus:

$$\text{Persentase skor} = \frac{64}{75} \times 100\% = 85,33\%$$

Berdasarkan hasil penilaian validator, diperoleh persentase kelayakan sebesar 85,33% yang menunjukkan bahwa instrumen layak digunakan setelah dilakukan beberapa revisi sesuai saran validator. Hasil ini merupakan hasil validasi awal oleh satu ahli pendidikan matematika. Hasil tersebut menunjukkan bahwa proses validasi ahli berperan penting dalam memastikan kesesuaian instrumen dengan tujuan pengukuran. Temuan ini sejalan dengan Hamidah dan Wulandari (2021) yang menyatakan bahwa instrumen berbasis HOTS perlu melalui proses validasi sebelum digunakan agar kualitas butir soal sesuai dengan kemampuan yang hendak diukur.

Hasil Validitas Butir Soal

Berdasarkan hasil analisis validitas butir soal, terdapat 7 butir soal yang memenuhi kriteria valid dan signifikan, yaitu nomor 2, 6, 7, 8, 10, 11, dan 14. Sementara itu, butir soal nomor 3, 4, 5, 9, 12, 13, dan 15 menunjukkan validitas yang rendah atau tidak signifikan sehingga memerlukan revisi sebelum digunakan kembali. Adapun butir soal nomor 1 menghasilkan nilai NAN yang mengindikasikan adanya permasalahan pada data respons atau karakteristik butir soal sehingga perlu diperbaiki atau dieliminasi. Oleh karena itu, tidak seluruh butir soal dapat langsung digunakan tanpa revisi lebih lanjut.

Hasil Reliabilitas Soal

Berdasarkan hasil analisis reliabilitas, diperoleh koefisien reliabilitas sebesar 0,66. Nilai tersebut berada pada rentang 0,61 - 0,80 yang termasuk dalam kategori baik. Ini menunjukkan bahwa soal yang dikembangkan memiliki tingkat konsistensi yang baik. Soal relatif stabil dalam mengukur kemampuan berpikir tingkat tinggi siswa apabila digunakan pada kelompok yang setara.

Hasil Analisis Tingkat Kesukaran

Hasil analisis menunjukkan bahwa dari 15 butir soal yang dianalisis, diperoleh variasi tingkat kesukaran soal yang terdiri atas kategori sangat mudah, mudah, sedang, dan sukar. Berdasarkan hasil analisis, terdapat 2 butir soal berkategori sangat mudah, yaitu soal nomor 1 dan 4. Selanjutnya, terdapat 1 butir soal berkategori mudah, yaitu soal nomor 15. Sebanyak 9 butir soal berada pada kategori sedang, yaitu soal nomor 2, 3, 5, 6, 7, 9, 10, 13, dan 14. Sementara itu, terdapat 3 butir soal yang termasuk kategori sukar, yaitu soal nomor 8, 11, dan 12. Hal ini menunjukkan bahwa mayoritas butir soal berada pada kategori sedang, sehingga tingkat kesukaran soal secara umum tergolong cukup baik untuk digunakan dalam pengukuran hasil belajar peserta didik.

Hasil Analisis Daya Pembeda

Hasil analisis menunjukkan bahwa dari 15 butir soal yang dianalisis berdasarkan daya pembeda, terdapat 6 butir soal berkategori sangat baik, yaitu soal nomor 2, 7, 8, 10, 11, dan 14 dengan indeks daya pembeda sebesar 62,50%-87,50%. Sebanyak 3 butir soal berkategori baik, yaitu soal nomor 6, 9, dan 15 dengan indeks daya pembeda sebesar 37,50%-50,00%. Selain itu, terdapat 3 butir soal berkategori cukup, yaitu soal nomor 4, 5, dan 13 dengan indeks daya pembeda sebesar 25,00%. Sementara itu, terdapat 3 butir soal berkategori sangat buruk, yaitu soal nomor 1, 3, dan 12 dengan indeks daya pembeda sebesar 0,00%. Dengan demikian, hasil analisis menunjukkan bahwa sebagian besar butir soal berada pada kategori baik hingga sangat baik dalam aspek daya pembeda. Hal ini menunjukkan bahwa sebagian besar soal telah mampu membedakan peserta didik yang memiliki kemampuan tinggi dan kemampuan rendah. Namun demikian, butir soal nomor 1, 3, dan 12 memerlukan revisi karena belum mampu membedakan kemampuan peserta didik secara optimal, sehingga perlu dilakukan perbaikan pada aspek konstruksi soal, tingkat kesukaran, maupun kualitas distraktor sebelum digunakan secara lebih luas.

Hasil Analisis Efektivitas Pengecoh

Hasil analisis menunjukkan bahwa kualitas pengecoh pada sebagian besar butir soal, Berdasarkan hasil analisis efektivitas pengecoh, diperoleh 12 butir soal (80,00%) berkategori sangat baik, 2 butir soal (13,33%) berkategori baik, dan 1 butir soal (6,67%) berkategori tidak baik. Butir soal yang berkategori baik terdapat pada nomor 5 dan 10 karena hanya dua distraktor yang berfungsi dengan baik. Sementara itu, butir soal nomor 1 berkategori tidak baik karena seluruh distraktornya tidak dipilih oleh peserta didik. Secara umum, hasil tersebut menunjukkan bahwa sebagian besar pengecoh pada instrumen yang dikembangkan telah berfungsi dengan baik dalam menarik perhatian peserta didik yang belum menguasai materi, sehingga mampu

DOI: <https://doi.org/10.63976/jimat.v7i2.1398>

mendukung kualitas instrumen yang digunakan untuk mengukur kemampuan penalaran dan pemecahan masalah matematika.

Pembahasan

Pembahasan pada penelitian ini disusun berdasarkan tahapan model pengembangan 4D yang meliputi *define*, *design*, *develop*, dan *disseminate*. Namun, penelitian ini dibatasi hanya sampai tahap *develop*. Pembatasan tersebut dilakukan karena tujuan penelitian difokuskan pada proses pengembangan dan pengujian kualitas instrumen soal HOTS bangun datar untuk mengukur kemampuan nalar dalam pemecahan masalah matematika siswa sekolah dasar. Tahap *disseminate* tidak dilakukan karena penelitian ini berfokus pada pengembangan awal dan pengujian kualitas instrumen, sehingga penyebarluasan produk secara lebih luas belum menjadi tujuan utama penelitian. Dengan demikian, pembahasan penelitian difokuskan pada tiga tahap pengembangan, yaitu *define*, *design*, dan *develop*.

Tahap Define

Pada tahap *define* dilakukan analisis kebutuhan, kurikulum, karakteristik siswa, dan materi pembelajaran. Hasil analisis menunjukkan bahwa pengembangan instrumen HOTS diperlukan karena pembelajaran matematika pada jenjang sekolah dasar tidak hanya menekankan kemampuan menghafal rumus, tetapi juga perlu mengembangkan kemampuan penalaran dan pemecahan masalah matematis siswa. Kemampuan tersebut penting karena siswa perlu dibiasakan menghadapi persoalan kontekstual yang membutuhkan proses berpikir tingkat tinggi. Temuan tersebut sejalan dengan penelitian Intan, Kuntarto, dan Alirmansyah (2020) yang menunjukkan bahwa kemampuan siswa sekolah dasar dalam mengerjakan soal HOTS masih berada pada kategori yang perlu ditingkatkan karena sebagian besar siswa mengalami kesulitan dalam menganalisis informasi dan menentukan strategi penyelesaian masalah. Selain itu, penelitian Dike, Fitriana, dan Adpriyadi (2022) juga menyatakan bahwa pengembangan instrumen penilaian berorientasi HOTS diperlukan agar siswa terbiasa menyelesaikan soal yang menuntut kemampuan berpikir tingkat tinggi. Temuan tersebut juga diperkuat oleh Latifah, Husaini, dan Nisa (2023) yang menjelaskan bahwa pengembangan instrumen penilaian berbasis HOTS diperlukan untuk membiasakan peserta didik berpikir kritis, menganalisis informasi, dan menyelesaikan masalah secara sistematis dalam proses pembelajaran. Tahap *define* juga mencakup penentuan indikator kemampuan yang akan diukur, yaitu *reasoning* (penalaran), *problem solving* (pemecahan masalah), dan *proof* (pembuktian). Ketiga indikator tersebut dipilih karena dapat merepresentasikan kemampuan berpikir matematis siswa dalam menyelesaikan permasalahan secara logis dan sistematis. Indikator tersebut kemudian dikaitkan dengan level kognitif HOTS berdasarkan Taksonomi Bloom revisi Anderson dan Krathwohl, yaitu C4 (*menganalisis*) dan C5 (*mengevaluasi*). Hal ini sejalan dengan National Council of Teachers of Mathematics yang menekankan pentingnya kemampuan penalaran, pemecahan masalah, dan pembuktian dalam pembelajaran matematika. Berdasarkan hasil analisis kebutuhan dan kajian penelitian sebelumnya, tahap *define* menjadi landasan utama dalam penyusunan kisi-kisi serta rancangan instrumen pada tahap berikutnya. Dengan demikian, instrumen yang dikembangkan diharapkan mampu mengukur kemampuan berpikir tingkat tinggi siswa secara lebih terarah dan sesuai dengan tujuan penelitian.

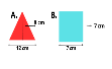


Tahap Design

Pada tahap *design* dilakukan penyusunan kisi-kisi soal dan rancangan instrumen HOTS berdasarkan indikator kemampuan penalaran, pemecahan masalah, dan pembuktian matematis


DOI: <https://doi.org/10.63976/jimat.v7i2.1398>

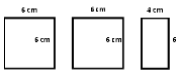
menurut NCTM. Penyusunan soal juga disesuaikan dengan level kognitif HOTS yaitu C4 (menganalisis) dan C5 (mengevaluasi).

Tabel 12. Kisi-Kisi Soal HOTS Bangun Datar

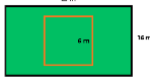
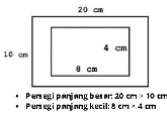
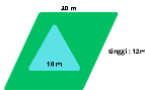
No	Indikator Kemampuan	Indikator	Soal	Alternatif Jawaban	Kriteria Penskoran	Level Kognitif
1.	Penalaran (Reasoning)	Peserta didik mampu mengamati dan membandingkan dua bangun datar yang berbeda (segitiga dan persegi) secara logis dengan menghitung luasnya masing-masing untuk menentukan kebenaran hubungan antara keduanya.	 <p>Manakah Pernyataan yang BENAR tentang kedua bangun datar tersebut?</p>	A. Luas bangun datar A lebih besar dari bangun datar B B. Luas kedua bangun datar tersebut sama C. Luas bangun datar B lebih besar dari bangun datar A D. Kedua luas bangun datar tidak bisa dibandingkan karena memiliki bentuk yang berbeda	Skor 1: Benar Skor 0: Salah	C4
2.	Pemecahan Masalah	Peserta didik mampu menggunakan strategi pengurangan luas persegi panjang (taman) dengan persegi (kolam) untuk menemukan solusi area sisa dalam konteks kehidupan sehari-hari.	 <p>Resti memiliki taman berbentuk persegi panjang dengan panjang 10 meter dan lebar 6 meter. Di dalam taman, ada kolam berbentuk persegi pada tengah-tengah taman dengan panjang sisi 3 meter. Sisa taman akan ditanami dengan bunga mawar supaya terlihat indah, berapa sisa luas yang akan ditanami bunga tersebut?</p>	A. 49 m ² B. 50 m ² C. 51 m ² D. 52 m ²	Skor 1: Benar Skor 0: Salah	C4
3.	Pemecahan Masalah	Peserta didik mampu menerapkan konsep keliling persegi panjang (lapangan) dan pembagian interval (jarak antar lampu) untuk menentukan jumlah objek secara presisi.	 <p>Lampu Pak Rusdi berbentuk persegi panjang dengan panjang 28 meter dan lebar 16 meter. Pak Rusdi ingin memasang lampu di sekeliling lapangan dengan jarak 4 meter untuk setiap lampu. Berapa banyak lampu yang dibutuhkan Pak Rusdi?</p>	A. 20 Lampu B. 22 Lampu C. 24 Lampu D. 18 Lampu	Skor 1: Benar Skor 0: Salah	C4
4.	Pembuktian (Proof)	Peserta didik mampu mengevaluasi kesalahan logika pada luas persegi dengan	4. Perhatikan gambar berikut!	A. Stiker itu sebenarnya sudah benar, karena luas persegi dihitung	Skor 1: Benar Skor 0: Salah	C5

DOI: <https://doi.org/10.63976/jimat.v7i2.1398>

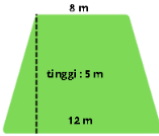
No	Indikator Kemampuan	Indikator	Soal	Alternatif Jawaban	Kriteria Penskoran	Level Kognitif
		cara membuktikan bahwa nilai 32 cm^2 adalah hasil perkalian sisi $\times 4$ (keliling), bukan sisi \times sisi (luas).	<p style="text-align: center;">8 cm</p>  <p style="text-align: center;">Luas = 32 cm^2</p> <p>Kamu baru saja membeli sebungkus kertas lipat (origami). Pada plastik bungkusnya, terdapat stiker label seperti pada gambar. Ternyata, ada kesalahan cetak pada bagian "Luas" di stiker tersebut. Jika kamu diminta menjelaskan kesalahannya, alasan yang paling tepat adalah...</p>	<p>dengan 8×4.</p> <p>B. Stiker itu salah, karena angka 32 itu adalah keliling. Luas kertas seharusnya 8×8.</p> <p>C. Stiker itu salah, karena luas kertas tersebut seharusnya dihitung dengan $8 + 8$.</p> <p>D. Stiker itu salah, karena luas kertas tersebut seharusnya dihitung dengan 8×2.</p>		
5.	Penalaran (Reasoning)	Peserta didik mampu menganalisis hubungan perubahan panjang sisi persegi terhadap perubahan luas bangun secara logis serta mengevaluasi kebenaran suatu pernyataan matematis berdasarkan hasil perhitungan luas.	<p>Terdapat dua potong kertas karton berbentuk persegi dengan ukuran yang berbeda di atas meja:</p> <p>Karton A memiliki panjang sisi 4 cm. Karton B memiliki panjang sisi 8 cm.</p> <p>Untuk menguji pemahaman kelas, Pak Guru menuliskan sebuah tebak-tebakan di papan tulis:</p> <p>"Karena panjang sisi Karton B dua kali lebih panjang dari Karton A, maka luas Karton B pasti dua kali lipat lebih besar dari luas Karton A."</p> <p>Jika kamu diminta menjawab tebak-tebakan Pak Guru tersebut, tanggapan yang paling tepat adalah...</p>	<p>A. Benar, karena panjang sisinya dikali dua, jadi luas kartonnya pasti otomatis ikut dikali dua juga.</p> <p>B. Salah, karena kalau dihitung, luas Karton A itu 16 cm^2, sedangkan Karton B itu 64 cm^2. Jadi luasnya malah 4 kali lipat lebih besar, bukan 2 kali lipat!</p> <p>C. Salah, karena cara mencari luas karton itu pinggirannya harus ditambahkan semua, bukan dikalikan.</p> <p>D. Benar, karena memang ukurannya dua kali lipat, karena 8 cm dibagi 4 cm hasilnya 2 cm</p>	<p>Skor 1: Benar</p> <p>Skor 0: Salah</p>	C5
6.	Pembuktian (Proof)	Peserta didik mampu membangun argumen atau contoh penyangkal untuk membuktikan bahwa kesamaan keliling pada dua bangun tidak secara otomatis menjamin kesamaan	<p>Di bagian akhir buku latihan matematika, terdapat sebuah kalimat jebakan untuk menguji ketelitian sebagai berikut:</p> <p>"Dua bangun</p>	<p>A. Benar. Kalau kelilingnya sama-sama 24 cm, ukuran luasnya pasti otomatis ikut sama persis.</p> <p>B. Salah. Bentuk bangunnya kan bisa beda-beda (misalnya persegi dan persegi</p>	<p>Skor 1: Benar</p> <p>Skor 0: Salah</p>	C5

No	Indikator Kemampuan	Indikator	Soal	Alternatif Jawaban	Kriteria Penskoran	Level Kognitif
		luas.	memiliki keliling sama, yaitu 24 cm. Luas kedua bangun tersebut sudah pasti sama." Pendapat yang paling tepat untuk menanggapi kalimat di buku tersebut adalah...	panjang). Walaupun kelilingnya sama-sama 24 cm, kalau dihitung luasnya bisa berbeda. C. Salah. Luas bangun datar sama sekali tidak bisa dihitung kalau yang diketahui cuma angka kelilingnya saja. D. Benar. Karena angka kelilingnya sama, pasti rumusnya sama, jadi hasil luasnya tidak mungkin beda.		
7.	Pemecahan Masalah	Peserta didik mampu memvisualisasi bangun datar dengan menyusun gabungan dua persegi dan satu persegi panjang menjadi sebuah model bangun datar baru yang logis.	 <p>Gufon memiliki 3 potongan kertas seperti ilustrasi di bawah ini:</p> <p>2 bangun persegi dengan panjang sisi 6 cm. 1 bangun persegi panjang dengan ukuran 6 cm x 4 cm. Gufon menyusun ketiga bangun tersebut dengan cara menempelkan sisi-sisi yang panjangnya 6 cm secara sejajar. Bangun datar baru apakah yang akan terbentuk dari gabungan ketiganya?</p> <p>A. Sebuah persegi panjang besar berukuran 16 cm x 6 cm. B. Sebuah persegi besar berukuran 12 cm x 12 cm. C. Bentuk menyerupai huruf L yang tidak beraturan. D. Trapesium siku-siku.</p>	<p>A. Sebuah persegi panjang besar berukuran 16 cm x 6 cm. B. Sebuah persegi besar berukuran 12 cm x 12 cm. C. Bentuk menyerupai huruf L yang tidak beraturan. D. Trapesium siku-siku.</p>	<p>Skor 1: Benar Skor 0: Salah</p>	C4
8.	Penalaran (Reasoning)	Peserta didik mampu membandingkan luas antara persegi dan persegi panjang melalui perhitungan matematis untuk	Mulyono memiliki dua jenis kertas lipat (origami) untuk membuat kerajinan. Kertas A berbentuk persegi		<p>Skor 1: Benar Skor 0: Salah</p>	C4

DOI: <https://doi.org/10.63976/jimat.v7i2.1398>

No	Indikator Kemampuan	Indikator	Soal	Alternatif Jawaban	Kriteria Penskoran	Level Kognitif
		menentukan validitas pernyataan perbandingan	dengan panjang sisi 8 cm. Kertas B berbentuk persegi panjang dengan ukuran panjang 10 cm dan lebar 6 cm. Jika mulyono ingin menggunakan kertas yang memiliki permukaan paling luas, kertas mana yang harus ia pilih dan apa alasannya?			
9.	Pemecahan Masalah	Peserta didik mampu menghitung luas sisa suatu lahan dengan cara mengurangi luas bagian di dalamnya.	 <p>Sebuah lahan berbentuk persegi panjang berukuran 25 meter \times 16 meter. Di dalam lahan tersebut akan dibuat jalur berbentuk persegi dengan sisi 6 m. Berapa luas lahan yang dapat digunakan?</p>	A. 364 m ² C. 400 m ² B. 376 m ² D. 406 m ²	Skor 1: Benar Skor 0: Salah	C4
10.	Penalaran (Reasoning)	Peserta didik mampu menganalisis dampak perubahan skala (perbesaran 2 kali) pada persegi panjang kecil terhadap ketersediaan luas area sisa di dalam persegi panjang besar.	 <p>Persegi panjang besar: 20 cm \times 10 cm Persegi panjang kecil: 8 cm \times 4 cm</p> <p>Jika persegi panjang kecil diperbesar menjadi 2 kali panjang dan lebar, maka luas daerah yang tersisa akan ... ?</p>	A. tetap B. bertambah C. mengurangi D. Tidak Berubah	Skor 1: Benar Skor 0: Salah	C4
11	Pemecahan Masalah	Peserta didik mampu merancang langkah perhitungan untuk mencari luas lahan yang tidak terpakai dengan cara mengurangi luas persegi panjang (lapangan) dengan persegi (area tengah).	 <p>Sebuah taman kota berbentuk jajargenjang memiliki panjang alas 20 m dan tinggi 12 m. Di dalam</p>	A. 30 m ² B. 180 m ² C. 210 m ² D. 240 m ²	Skor 1: Benar Skor 0: Salah	C4

DOI: <https://doi.org/10.63976/jimat.v7i2.1398>

No	Indikator Kemampuan	Indikator	Soal	Alternatif Jawaban	Kriteria Penskoran	Level Kognitif
			taman tersebut, akan dibangun sebuah kolam air mancur berbentuk segitiga dengan panjang alas 10 m dan tinggi 6 m. Berapa luas area taman yang tidak digunakan untuk kolam air mancur?			
12	Pembuktian (Proof)	Peserta didik mampu menguji validitas argumen subjektif tentang kesamaan luas persegi dan persegi panjang dengan menyajikan bukti objektif melalui hasil rumus matematis	Rani ingin membuat dua taman di halaman rumahnya: Taman A berbentuk persegi dengan sisi 8 m, Taman B berbentuk persegi panjang dengan panjang 10 m dan lebar 6 m. Rani mengatakan bahwa kedua taman memiliki luas yang sama karena ukurannya hampir mirip. Bagaimana tanggapanmu atas pendapat Rani?	A. Benar, karena bentuk tidak memengaruhi luas B. Salah, karena luas taman A lebih besar C. Salah, karena luas taman B lebih besar D. Tidak dapat ditentukan.	Skor 1: Benar Skor 0: Salah	C4
13	Pembuktian (Proof)	Peserta didik mampu mengevaluasi dan menyanggah prosedur perhitungan luas trapesium yang salah (asal mengalikan sisi) dengan merujuk pada konsep rata-rata sisi sejajar x tinggi.	Di sudut halaman sekolah, terdapat lapangan kasti berbentuk trapesium. Di papan informasi pinggir lapangan, terdapat gambar penunjuk ukuran seperti ini:  Melihat gambar tersebut, seorang siswa mencoba menghitung luas lapangan di buku tulisnya dengan cara mengalikan semua angka yang tertera, yaitu: $12 \times 8 \times 5$. Pendapat yang paling tepat mengenai hasil hitungan siswa tersebut adalah...	A. Benar, karena untuk mencari luas memang harus mengalikan semua angka yang ada di gambar. B. Salah, karena luas trapesium harusnya dihitung dengan mencari rata-rata sisi sejajar (8 dan 12) terlebih dahulu, baru kemudian dikalikan dengan tingginya (5). C. Salah, karena jika ketiga angka itu langsung dikalikan, itu adalah cara untuk mencari volume (isi ruangan), bukan mencari luas lapangan yang datar. D. Benar, karena bentuk lapangan tersebut sebenarnya bisa dibagi menjadi persegi panjang dan segitiga	Skor 1: Benar Skor 0: Salah	C4
14	Pembuktian (Proof)	Peserta didik mampu mengevaluasi kesalahan konsep pada perhitungan luas layang-layang dan	Sebuah layang-layang memiliki panjang diagonal 10 cm dan 8 cm. Seorang siswa	A. Benar, karena rumus mencari daerah di dalam bangun datar adalah panjang dikali lebar.	Skor 1: Benar Skor 0: Salah	C5

No	Indikator Kemampuan	Indikator	Soal	Alternatif Jawaban	Kriteria Penskoran	Level Kognitif
		membuktikan bahwa perkalian diagonal tanpa pembagian dua adalah prosedur yang keliru.	menghitung luasnya dengan cara langsung mengalikan kedua diagonal tersebut ($10 \times 8 = 80 \text{ cm}^2$). Evaluasi yang paling tepat untuk perhitungan siswa tersebut adalah...	B. Salah, karena 80 cm^2 adalah luas persegi panjang luar yang menutupi layang-layang tersebut. Luas layang-layang yang asli seharusnya hanya setengah dari persegi panjang itu. C. Salah, karena diagonal layang-layang seharusnya ditambahkan terlebih dahulu, bukan dikalikan. D. Benar, karena tidak ada sisi lain yang bisa dihitung selain kedua diagonal tersebut.		
15	Pemecahan Masalah	Peserta didik mampu mengonstruksi kemungkinan bentuk geometri baru (seperti trapesium atau jajar genjang) dari potongan satu persegi panjang dan dua segitiga siku-siku	Doni memiliki mainan kayu berbentuk bangun datar sebagai berikut: 1 bangun persegi panjang ($8 \text{ cm} \times 4 \text{ cm}$) 2 bangun segitiga siku-siku dengan alas 4 cm dan tinggi 4 cm Doni ingin menyusun bangun-bangun tersebut menjadi bangun baru. Bangun datar baru yang paling mungkin dibentuk oleh doni adalah...	A. Persegi B. Jajar genjang C. Trapesium D. Lingkaran	Skor 1: Benar Skor 0: Salah	C4

Pada tahap *design* dilakukan penyusunan kisi-kisi dan rancangan instrumen berdasarkan indikator *reasoning*, *problem solving*, dan *proof* serta level HOTS C4, dan C5. Penyusunan instrumen menggunakan konteks kehidupan sehari-hari seperti taman, lapangan, kertas lipat, dan berbagai bentuk bangun datar di sekitar siswa. Hal tersebut sejalan dengan penelitian Astawayasa (2022) yang menunjukkan bahwa pengembangan asesmen HOTS pada matematika sekolah dasar perlu disusun berdasarkan indikator berpikir tingkat tinggi agar soal yang dihasilkan dapat mengukur kemampuan analisis, evaluasi, dan kreativitas siswa secara lebih optimal. Penelitian tersebut juga menekankan bahwa penggunaan konteks nyata pada soal matematika membantu siswa menghubungkan konsep dengan pengalaman sehari-hari. Selain itu, penelitian Zain et al (2024) menunjukkan bahwa kemampuan berpikir matematis, khususnya pada materi geometri, memiliki hubungan erat dengan kemampuan penalaran dan pemecahan masalah siswa. Hal tersebut menunjukkan bahwa penyusunan soal perlu disesuaikan dengan kemampuan berpikir yang ingin diukur. Berdasarkan hal tersebut,

DOI: <https://doi.org/10.63976/jimat.v7i2.1398>

penyusunan kisi-kisi pada penelitian ini tidak hanya berfokus pada penguasaan materi, tetapi juga mengarahkan siswa pada kemampuan berpikir tingkat tinggi yang lebih kompleks.

Tahap *Develop*

Pada tahap *develop* dilakukan validasi ahli, revisi instrumen, dan uji coba soal kepada siswa sekolah dasar. Hasil validasi awal oleh satu ahli pendidikan matematika menunjukkan bahwa instrumen memperoleh persentase kelayakan sebesar 85,33% dan layak digunakan setelah dilakukan revisi sesuai masukan validator. Selanjutnya, instrumen yang telah direvisi diuji cobakan kepada siswa kelas IV SD Negeri Sukasari 1 untuk mengetahui kualitas butir soal. Pelaksanaan uji coba dilakukan secara langsung di kelas dengan memberikan soal HOTS bangun datar kepada siswa.

Gambar 2. Dokumentasi Pelaksanaan Uji Coba Soal HOTS



Setelah pelaksanaan uji coba, hasil jawaban siswa dianalisis menggunakan aplikasi ANATES untuk mengetahui kualitas butir soal yang meliputi validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh.

Tabel 13. Rekapitulasi Hasil Analisis Butir Soal

No. Soal	Daya Pembeda (%)	Kategori	Tingkat Kesukaran	Korelasi	Keterangan
1	0,00	Sangat Buruk	Sangat Mudah	NAN	Tidak Valid
2	75,00	Sangat Baik	Sedang	0,548	Valid
3	0,00	Sangat Buruk	Sedang	0,139	Tidak Valid
4	25,00	Cukup	Sangat Mudah	0,218	Tidak Valid
5	25,00	Cukup	Sedang	0,166	Tidak Valid
6	50,00	Baik	Sedang	0,494	Valid
7	62,50	Sangat Baik	Sedang	0,603	Valid
8	87,50	Sangat Baik	Sukar	0,739	Sangat Valid
9	50,00	Baik	Sedang	0,442	Tidak Valid
10	62,50	Sangat Baik	Sedang	0,592	Valid
11	75,00	Sangat Baik	Sukar	0,654	Sangat Valid

12	0,00	Sangat Buruk	Sukar	0,089	Tidak Valid
13	25,00	Cukup	Sedang	0,139	Tidak Valid
14	87,50	Sangat Baik	Sedang	0,606	Sangat Valid
15	37,50	Baik	Mudah	0,481	Tidak Valid

Pada tahap *develop* dilakukan validasi ahli, revisi instrumen, dan uji coba kepada siswa sekolah dasar. Berdasarkan rekapitulasi hasil analisis butir soal, tidak seluruh soal yang dikembangkan memenuhi kriteria kualitas instrumen yang baik. Soal nomor 1 menghasilkan nilai korelasi NAN dengan daya pembeda sebesar 0,00%, yang mengindikasikan bahwa butir soal tersebut belum mampu membedakan peserta didik berkemampuan tinggi dan rendah. Temuan tersebut didukung oleh hasil tingkat kesukaran yang menunjukkan bahwa soal nomor 1 berada pada kategori sangat mudah (100%), sehingga tidak mampu menghasilkan variasi respons peserta didik. Selain itu, soal nomor 3 dan 12 juga menunjukkan daya pembeda sebesar 0,00% serta validitas yang rendah sehingga memerlukan revisi lebih lanjut. Beberapa butir soal lainnya juga menunjukkan efektivitas pengecoh yang belum optimal karena masih terdapat distraktor yang tidak berfungsi. Meskipun demikian, sebagian besar butir soal telah memenuhi kriteria validitas, memiliki daya pembeda yang baik hingga sangat baik, serta didukung oleh tingkat reliabilitas instrumen sebesar 0,66 yang termasuk kategori baik. Oleh karena itu, ketiga butir soal tersebut memerlukan revisi pada aspek redaksi, konteks soal, maupun kualitas alternatif jawaban. Temuan ini sejalan dengan penelitian Yuparing, Wiyono, dan Sutadji (2023) yang menyatakan bahwa pengembangan instrumen HOTS matematika sekolah dasar dapat menghasilkan instrumen yang valid, reliabel, dan layak digunakan dalam proses pembelajaran. Hasil penelitian ini juga didukung oleh Nurussama et al. (2025) yang menemukan bahwa instrumen yang dikembangkan melalui tahap validasi ahli dan uji coba lapangan mampu memenuhi kriteria kelayakan sebagai alat ukur kemampuan matematis siswa. Namun, hasil penelitian ini menunjukkan bahwa analisis kualitas setiap butir soal tetap diperlukan untuk mengidentifikasi kelemahan instrumen dan menjadi dasar perbaikan sebelum digunakan secara lebih luas. Dengan demikian, instrumen yang dikembangkan telah menunjukkan kelayakan awal sebagai alat ukur kemampuan penalaran dan pemecahan masalah matematika, tetapi beberapa butir soal masih memerlukan revisi agar kualitas pengukurannya menjadi lebih optimal.

KESIMPULAN

Berdasarkan hasil penelitian dan pengembangan instrumen soal *Higher Order Thinking Skills* (HOTS) pada materi bangun datar untuk mengukur kemampuan nalar dalam pemecahan masalah matematika siswa sekolah dasar, dapat disimpulkan bahwa proses pengembangan dilakukan menggunakan model 4D (*Define, Design, Develop, dan Disseminate*), namun penelitian dibatasi hingga tahap *Develop*. Pada tahap *Define* dilakukan analisis kebutuhan, kurikulum, karakteristik siswa, dan materi pembelajaran. Tahap *Design* dilakukan dengan menyusun kisi-kisi serta merancang soal berdasarkan indikator kemampuan penalaran, pemecahan masalah, dan pembuktian matematis sesuai indikator NCTM serta level HOTS C4 dan C5. Selanjutnya, pada tahap *Develop* dilakukan validasi ahli, revisi instrumen, dan uji coba kepada siswa sekolah dasar. Hasil validasi awal oleh satu ahli pendidikan matematika menunjukkan persentase kelayakan sebesar 85,33%. Berdasarkan hasil tersebut serta revisi yang dilakukan sesuai masukan validator, instrumen dinyatakan layak digunakan pada tahap

DOI: <https://doi.org/10.63976/jimat.v7i2.1398>

uji coba. Hasil analisis reliabilitas menunjukkan koefisien sebesar 0,66 yang termasuk kategori baik, sehingga instrumen memiliki konsistensi yang baik. Selain itu, sebagian besar butir soal berada pada tingkat kesukaran sedang, memiliki daya pembeda yang baik, serta pengecoh yang secara umum telah berfungsi dengan baik. Hasil penelitian menunjukkan bahwa instrumen memperoleh persentase validasi ahli sebesar 85,33% dan reliabilitas sebesar 0,66. Namun, hasil analisis validitas butir menunjukkan bahwa tidak seluruh soal memenuhi kriteria valid. Dari 15 butir soal yang dikembangkan, terdapat 7 butir soal yang memenuhi kriteria valid, sedangkan butir lainnya memerlukan revisi atau perbaikan lebih lanjut. Oleh karena itu, instrumen yang dikembangkan memiliki potensi untuk digunakan sebagai alat evaluasi setelah dilakukan revisi terhadap butir-butir yang belum memenuhi kriteria validitas.

REFERENSI

- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Arikunto, S. (2018). *Dasar-dasar evaluasi pendidikan* (Edisi revisi). Jakarta: Bumi Aksara.
- Astawayasa, K. G., Widana, I. W., & Adi, I. N. R. (2022). Pengembangan asesment HOTS mata pelajaran matematika sekolah dasar. *Jurnal Ilmiah Pendidikan Citra Bakti*, 9(1), 129-140. <https://doi.org/10.38048/jipcb.v9i1.659>
- Dike, D., Fitriana, F., & Adpriyadi, A. (2022). Instrumen penilaian tes berorientasi HOTS pada pembelajaran tematik di sekolah dasar. *Edumedia: Jurnal Keguruan dan Ilmu Pendidikan*, 6(2). <https://doi.org/10.51826/edumedia.v6i2.667>
- Hamidah, M. H., & Wulandari, S. S. (2021). Pengembangan instrumen penilaian berbasis HOTS menggunakan aplikasi Quizizz. *Efisiensi: Kajian Ilmu Administrasi*, 18(1), 105-124. <https://doi.org/10.21831/efisiensi.v18i1.36997>
- Hulwani, A. Z., & Yuhana, Y. (2024). Pengembangan Instrumen Tes Matematika Berbasis Higher Order Thinking Skills (HOTS) Untuk Siswa SMP. *Jurnal Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*, 5(2), 1074-1081. <https://doi.org/10.46306/lb.v5i2.648>
- Hidajat, F. A. (2022). *Buku Ajar Pengembangan Berpikir Tingkat Tinggi dan Berpikir Kreatif Matematis*. Penerbit NEM.
- Intan, F. M., Kuntarto, E., & Alirmansyah. (2020). Kemampuan siswa dalam mengerjakan soal HOTS (*Higher Order Thinking Skills*) pada pembelajaran matematika di kelas V sekolah dasar. *Jurnal Pendidikan Dasar Indonesia*, 5(1), 6-10.
- Kemendikbudristek. (2023). *Hasil Programme for International Student Assessment (PISA) 2022 Indonesia*. Jakarta: Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi.
- Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi. (2021). *Matematika untuk SD/MI Kelas IV*. Jakarta: Pusat Kurikulum dan Perbukuan.
- Khairani, M., Sukmawati, & Nasrun. (2023). Pengaruh Model Pembelajaran Problem Based Learning (PBL) terhadap kemampuan penalaran dan pemecahan masalah matematika siswa kelas V SDN 1 Lejang Kabupaten Pangkep. *Al-Madrasah: Jurnal Ilmiah Pendidikan Madrasah Ibtidaiyah*, 7(1), 458-471. <https://doi.org/10.35931/am.v7i1.1885>
- Latifah, A., Husaini, F., & Nisa, A. K. (2023). Pengembangan instrumen penilaian berbasis HOTS. *Didaktik: Jurnal Ilmiah PGSD STKIP Subang*, 9(2), 4486-4496. <https://doi.org/10.36989/didaktik.v9i2.1057>
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

DOI: <https://doi.org/10.63976/jimat.v7i2.1398>

- Nurussama, A., Fitri, K. N., Haeriska, N., Zaidan, R., & Rahayu, T. G. (2025). Pengembangan instrumen tes kemampuan komunikasi matematis dalam materi bangun ruang dan bangun datar di sekolah dasar. *Pendas: Jurnal Ilmiah Pendidikan Dasar*, 10(2), 367-382. <https://doi.org/10.23969/jp.v10i2.26286>
- Riduwan. (2013). *Skala pengukuran variabel-variabel penelitian*. Bandung: Alfabeta.
- Ristanty, D. W., & Pratama, F. W. (2022). Analisis kemampuan pemecahan masalah matematika pada materi segiempat berdasarkan teori Van Hiele. *Jurnal Cendekia: Jurnal Pendidikan Matematika*, 6(2), 1648-1658. <https://doi.org/10.31004/cendekia.v6i2.1400>
- Safirah, A. D., & Abdillah, M. I. (2024). Pengaruh Model Problem Based Learning terhadap kemampuan pemecahan masalah matematika siswa SD. *ARSEN: Jurnal Penelitian Pendidikan*, 1(2), 102-109. <https://doi.org/10.30822/arsen.v1i2.3119>
- Sugiyono. (2019). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Bandung: Alfabeta.
- Sumilat, J. M., Kumolontang, D., & Rompah, Y. (2022). Pengaruh penggunaan media video pada pembelajaran matematika materi bangun datar. *Edukatif: Jurnal Ilmu Pendidikan*, 4(5), 7159–7167. <https://doi.org/10.31004/edukatif.v4i5.4017>
- Thiagarajan, S., Semmel, D. S., & Semmel, M. I. (1974). *Instructional Development for Training Teachers of Exceptional Children: A Sourcebook*. Minneapolis: University of Minnesota.
- Yuparing, F. R., Wiyono, B. B., & Sutadji, E. (2023). Instrumen Asesmen Higher Order Thinking Skills (Hots) Pada Mata Pelajaran Matematika Kelas IV SD Negeri 4 Tanggung. *Jurnal Pendidikan Indonesia*, 4(4). <https://doi.org/10.59141/japendi.v4i04.1703>
- Zain, N. K., Marhayati, M., & Abdussakir, A. (2024). Pengembangan Quartet Mathematics Card pada Materi Geometri. *Jurnal Pendidikan Matematika Dan Sains*, 12(2), 198-210. <https://doi.org/10.21831/jpms.v12i2.78842>