



## Pengembangan Instrumen Tes *Higher Order Thinking Skill* Berbasis Literasi Matematis pada Materi Bangun Ruang Sisi Datar

Amalia Khairunnisa<sup>1\*</sup>, Sitti Inaya Masrura<sup>2</sup>, Nur Fahri Tadjuddin<sup>3</sup>

<sup>1,2,3</sup>Program Studi Pendidikan Matematika, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Sulawesi Barat, Indonesia

-----  
Corresponding Author:

Author Name\*: Amalia Khairunnisa

Email\*: [khairunnisaamalia261@gmail.com](mailto:khairunnisaamalia261@gmail.com)

Diterima: 22 Desember 2024. Disetujui: 25 April 2025. Dipublikasikan: 1 Mei 2025

### ABSTRACT

Penelitian ini bertujuan untuk mengembangkan instrumen tes *Higher Order Thinking Skill* berbasis literasi matematis pada materi Bangun Ruang Sisi Datar dengan tingkat validitas, reliabilitas, kesukaran, dan daya pembeda yang baik. Jenis penelitian ini adalah penelitian pengembangan (R&D) dengan mengacu pada model formative research tesser (preliminary, self evaluation, prototyping, dan field test). Instrumen tes yang dibuat terdiri dari kisi-kisi, lembar tes sebanyak 12 butir soal uraian, kunci jawaban, dan pedoman penilaian. Berdasarkan hasil uji ahli dan uji empiris, hasil penelitian menunjukkan bahwa instrumen tes memiliki kualitas. Hasil uji oleh ahli menunjukkan bahwa instrumen tes valid, efektif dan praktis baik dari aspek konten, konstruk, maupun bahasa. Sedangkan hasil uji empiris oleh siswa menunjukkan bahwa semua butir soal valid dengan tingkat reliabilitas sangat tinggi, tingkat kesukaran sebagian sukar dan sedang, serta daya pembeda yang memenuhi kriteria baik dan sangat baik.

**Keywords:** pengembangan, instrumen, HOTS, literasi matematis

### PENDAHULUAN

Mempelajari matematika memerlukan perhatian khusus karena matematika berkaitan erat dengan banyak bagian kehidupan kita [1]. Salah satu kemampuan utama matematika adalah menyelesaikan berbagai permasalahan, baik yang sederhana dalam perhitungan maupun yang lebih kompleks melalui pemodelan matematika. Matematika sebagai suatu disiplin ilmu bertujuan untuk meningkatkan kemampuan siswa dalam berpikir kritis, sistematis, logis, analitis dan kreatif, serta memperkuat keterampilan kerja yang efektif, yang termasuk dalam kategori keterampilan berpikir tingkat tinggi [2].

Kemampuan berpikir tingkat tinggi atau HOTS (*Higher Order Thinking Skills*) merupakan kemampuan berpikir yang bukan hanya sekadar menghafal atau mengulangi informasi yang telah diketahui tetapi juga mencakup kemampuan untuk menghubungkan, memanipulasi, dan mentransformasi pengetahuan serta pengalaman yang dimiliki guna berpikir kritis dan kreatif dalam membuat keputusan dan menyelesaikan masalah dalam situasi baru [3]. Tujuan utama dari HOTS yaitu untuk mengembangkan kemampuan siswa berpikir pada tingkat yang lebih tinggi. Ini mencakup kemampuan berpikir kritis dalam menilai berbagai informasi, memecahkan masalah secara kreatif dengan menggunakan pengetahuan yang dimiliki, dan membuat pilihan dalam situasi yang kompleks [4]. Dalam konteks pembelajaran, berpikir tingkat tinggi terjadi ketika siswa mampu menghubungkan dan mentransformasi

pengetahuan yang sudah mereka miliki dengan masalah-masalah yang belum pernah diajarkan. Sehingga ketika seorang siswa dianggap mampu menyelesaikan masalah dan menerapkan pengetahuan yang diperoleh sebelumnya dalam situasi baru yang belum pernah dihadapi. Kemampuan tersebut merupakan keterampilan berpikir tingkat tinggi [5]. Dalam hal ini HOTS (*Higher Order Thinking Skills*) melalui proses kognitif yang mencakup analisis (C4), evaluasi (C5), dan mencipta (C6) [6].

Selain pentingnya kemampuan berpikir tingkat tinggi, kemampuan literasi matematika juga memiliki peran yang signifikan dalam membantu siswa menyelesaikan masalah yang melibatkan penerapan matematika dalam kehidupan sehari-hari [7]. Literasi matematis melibatkan kemampuan individu untuk menggunakan matematika secara efektif dalam berbagai situasi kehidupan sehari-hari, termasuk kemampuan merumuskan, menggunakan, dan menafsirkan konsep matematika [8]. Sehingga dalam rangka mengembangkan keterampilan berpikir tingkat tinggi dalam matematika, penting bagi siswa untuk memperoleh literasi matematis yang kuat.

Namun jika dilihat kembali berdasarkan hasil evaluasi PISA dan TIMSS, skor yang diperoleh oleh pelajar Indonesia masih tergolong rendah. Pada hasil survei PISA tahun 2018 menunjukkan bahwa Indonesia berada di peringkat 74 dari 79 negara yang berpartisipasi dalam survei tersebut. Sementara itu, pada tahun 2015, tes TIMSS melibatkan siswa kelas IV

memperoleh skor sebesar 397. Berdasarkan kategori tingkat kemampuan dalam TIMSS, skor 625 menunjukkan kemampuan tingkat lanjut, skor 550 menunjukkan kemampuan tingkat tinggi, skor 475 menunjukkan kemampuan tingkat menengah, dan skor 400 menunjukkan kemampuan tingkat rendah. Secara keseluruhan, kemampuan siswa Indonesia dapat dikategorikan pada tingkat rendah [9]. Hal ini mengartikan bahwa kemampuan siswa masih tergolong rendah.

Kurangnya latihan menjawab soal-soal bertipe HOTS, khususnya yang bergantung pada literasi matematika, merupakan salah-satu dari banyak faktor yang menghambat hasil belajar siswa. Tugas-tugas yang diselesaikan siswa biasanya lebih berfokus pada hafalan daripada pengembangan kemampuan berpikir kritis [10]. Pembelajaran matematika saat ini lebih terfokus pada pemahaman prosedural, komunikasi satu arah, suasana kelas yang monoton, dan keterampilan berpikir tingkat rendah (LOTS). Pembelajaran juga sangat bergantung pada buku paket yang umumnya menyajikan soal-soal rutin dan pertanyaan penilaian dengan tingkat kesulitan rendah [11].

Hasil observasi yang telah dilakukan oleh peneliti dan wawancara dengan salah satu guru matematika di SMP Negeri 1 Campalagian menunjukkan bahwa kemampuan siswa dalam menyelesaikan soal HOTS berbasis literasi matematis masih rendah. Hal ini disebabkan karena guru cenderung hanya menggunakan bahan ajar atau instrumen yang ada di buku paket yang tersedia. Guru hanya beberapa kali memberikan soal-soal HOTS kepada siswa, terutama soal HOTS berbasis literasi matematika yang jarang diberikan. Soal-soal HOTS yang diberikan merupakan hasil pengembangan sendiri, namun sulit untuk mengembangkan soal yang menuntut kemampuan berpikir tingkat tinggi dan literasi matematis siswa secara mandiri. Guru juga menyatakan bahwa pemahaman siswa terhadap materi bangun ruang sisi datar masih perlu ditingkatkan. Hasil observasi menunjukkan bahwa soal-soal yang diberikan oleh guru lebih fokus pada pemahaman rumus yang tercantum dalam buku paket, sehingga siswa kesulitan menyelesaikan tugas-tugas yang memerlukan kemampuan berpikir tingkat tinggi. Oleh karena itu, peneliti menyimpulkan bahwa masih sedikit acuan atau contoh instrumen tes HOTS berbasis literasi matematis yang tersedia untuk mengeksplorasi kemampuan berpikir tingkat tinggi dan literasi matematis siswa.

Berdasarkan uraian diatas, diperlukan pengembangan lebih jauh mengenai instrumen tes yang mengintegrasikan kemampuan berpikir tingkat tinggi dan literasi matematis secara kontekstual pada materi bangun ruang sisi datar yang valid untuk mengukur kemampuan siswa secara mendalam. Sehingga diharapkan dapat digunakan sebagai bahan evaluasi guru dalam pembelajaran matematika. Serta dapat memberikan dampak positif terhadap pembelajaran matematika di sekolah. Dengan demikian, tujuan dari penelitian ini yaitu untuk mengetahui validitas, kepraktisan, dan efektifitas dari pengembangan

instrumen tes *higher order thinking skill* berbasis literasi matematis pada materi bangun ruang sisi datar yang reliabel dengan tingkat kesukaran dan daya pembeda yang memadai

## METODE PENELITIAN

Jenis penelitian yang akan dilakukan adalah Penelitian dan Pengembangan (*Research and Development*) dengan menggunakan model *Formatif Research Tessmer* yang terdiri dari 4 tahapan, dimulai dari: tahap *preliminary*, tahap *self evaluation*, tahap *prototyping*, dan tahap *field test*.

Penelitian ini melibatkan dua kelompok subjek: uji coba skala kecil dan uji coba skala besar. Untuk uji coba skala kecil, melibatkan 3 siswa dari kelas VIII C dan 6 siswa dari kelas VIII B. Sedangkan uji coba skala besar melibatkan 27 siswa dari kelas VIII A. Subjek uji coba yang dipilih mengacu pada ciri pemilihan subjek menurut Tessmer [12]. Sementara itu, pemilihan kelas didasarkan pada observasi dan informasi dari guru mata pelajaran, yang menunjukkan bahwa siswa di kelas-kelas tersebut memiliki kemampuan yang bervariasi dan karakteristik yang sesuai dengan desain pengembangan instrumen tes.

Instrumen yang digunakan dalam penelitian ini terdiri dari (1) Instrumen studi lapangan dengan memanfaatkan pedoman wawancara dan observasi untuk mengetahui kemampuan berpikir tingkat tinggi dan literasi matematis siswa. Data yang diperoleh dari wawancara dan observasi menjadi dasar untuk mengembangkan instrumen tes. (2) Instrumen validasi oleh ahli dan praktisi lapangan berupa angket yang diberikan kepada 2 dosen ahli matematika, 2 guru matematika dan 1 guru bahasa indonesia SMP Negeri 1 Campalagian sebagai validator untuk menilai kualitas instrumen tes. Angket validasi akan dijadikan panduan untuk melakukan revisi dan perbaikan sebelum diberikan kepada siswa untuk diujikan. (3) Analisis instrumen tes oleh siswa untuk mengetahui tingkat kelayakan instrumen tes berdasarkan hasil jawaban siswa, mencakup uji validitas, reliabilitas, tingkat kesukaran, dan daya pembeda.

### 1. Analisis Validitas

Teknik analisis data hasil angket validasi dengan menggunakan skala likert untuk melihat kelayakan instrumen tes yang telah dikembangkan. Setiap butir angket diberikan pilihan jawaban Sangat Sesuai (SS) dengan nilai 4, Sesuai (S) dengan nilai 3, Kurang Sesuai (KS) dengan nilai 2, dan Tidak Sesuai (TS) dengan nilai 1. Untuk menghitung total skor hasil penilaian oleh validator menggunakan rumus:

$$P = \frac{\text{skor total hasil pengumpulan data}}{\text{skor total kriterium}} \times 100\%$$

Keterangan:

Skor kriterium = skor paling tinggi tiap butir × jumlah butir × jumlah responden

P = persentase kelayakan instrumen tes

**Tabel 1.** Skala persentase kelayakan

Skor Persentase (%)	Interpretasi
0 ≤ skor ≤ 25	Tidak Layak
25 < skor ≤ 50	Kurang Layak
50 < skor ≤ 75	Cukup Layak
75 < skor ≤ 100	Layak

Teknik analisis data hasil skor jawaban siswa diuraikan sebagai berikut:

a. Uji validitas butir soal

Validitas item secara umum dikatakan valid jika memiliki dukungan yang besar terhadap skor total. Berikut rumus menghitung validitas.

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{((N \sum X^2 - (\sum X)^2)(N \sum Y^2 - (\sum Y)^2))}}$$

Keterangan:

r<sub>xy</sub>= koefisien korelasi antara variabel X dan variabel Y

X= skor item

Y= skor total

N= banyaknya responden

jika r hitung > r tabel maka soal valid

jika r hitung < r tabel maka soal tidak valid

b. Uji reliabilitas

Reliabilitas soal tes menunjukkan sejauh mana hasil pengukuran dapat diandalkan. Berikut rumus menghitung reliabilitas:

$$r_{11} = \left( \frac{n}{n-1} \right) \left( \frac{S^2 - \sum pq}{S^2} \right)$$

Keterangan:

r<sub>11</sub>= reliabilitas tes secara keseluruhan

p= proporsi subjek yang menjawab item dengan benar

q= proporsi subjek yang menjawab item dengan salah

∑pq= jumlah hasil perkalian antara p dan q

n = banyaknya item

S= standar deviasi dari tes

**Tabel 2.** Klasifikasi interpretasi reliabilitas

Nilai R	Interpretasi
0,00-0,200	Sangat Rendah
0,201-0,400	Rendah
0,401-0,600	Cukup
0,601-0,800	Tinggi
0,801-1,00	Sangat Tinggi

c. Uji tingkat kesukaran

Rumus menghitung tingkat kesukaran butir soal yakni:

$$\text{tingkat kesukaran} = \frac{\text{rata-rata}}{\text{skor maksimum tiap soal}} = \frac{\text{jumlah skor tiap soal}}{\text{jumlah siswa} \times \text{skor maksimum tiap soal}}$$

**Tabel 3.** Klasifikasi interpretasi tingkat kesukaran

Nilai P	Interpretasi
P = 0,00	Sangat Sukar
0,00 < P ≤ 0,30	Sukar
0,30 < P ≤ 0,70	Sedang
0,70 < P ≤ 1,00	Mudah
P = 1,00	Sangat Mudah

d. Uji daya pembeda

Rumus menghitung daya pembeda tiap butir soal sebagai berikut:

$$DP = \frac{B_A}{J_A} - \frac{B_B}{J_B}$$

Keterangan:

DP = Daya pembeda

BA = Banyaknya kelompok atas yang menjawab benar

BB = Banyaknya kelompok bawah yang menjawab benar

JA = Banyaknya subjek kelompok atas

JB = Banyaknya subjek kelompok bawah

**Tabel 4.** Klasifikasi interpretasi daya pembeda

Nilai P	Interpretasi
DP ≤ 0,00	Sangat Jelek
0,00 < DP ≤ 0,20	Jelek
0,20 < DP ≤ 0,40	Cukup
0,40 < DP ≤ 0,70	Baik
0,70 < DP ≤ 1,00	Sangat Baik

2. Analisis Kepraktisan

Analisis kepraktisan dilakukan dengan menghitung rata-rata nilai yang diperoleh dari angket respon oleh 2 observer yakni guru matematika, dengan menggunakan rumus:

$$P = \frac{\sum x}{\sum xi} \times 100\%$$

Keterangan:

P = persentase skor

∑x = total nilai skor yang diperoleh

$\sum xi$  = total nilai skor tertinggi

**Tabel 5.** Kriteria kepraktisan

Skor Persentase (%)	Interpretasi
81,0 < P ≤ 100,0	Sangat Praktis
61,0 < P ≤ 80,0	Praktis
41,0 < P ≤ 60,0	Kurang Praktis
21,0 < P ≤ 40,0	Tidak Praktis
0,0 < P ≤ 20,0	Sangat Tidak Praktis

3. Analisis Keefektifan

Analisis keefektifan dilakukan melalui hasil angket respon siswa terhadap instrumen tes yang telah dikerjakan. Tes dilakukan dalam skala kelompok kecil yaitu sebanyak 6 siswa. Selanjutnya hasil skor angket respon siswa dihitung dengan menggunakan rumus:

$$P = \frac{f}{N} \times 100\%$$

Keterangan:

P = persentase skor

f = banyaknya skor yang diperoleh

N = skor maksimal

**Tabel 6.** Kriteria persentase respon siswa

Skor Persentase (%)	Interpretasi
P ≥ 85,1	Sangat Baik
70,1 < P ≤ 85,0	Baik
50,1 < P < 70,0	Kurang Baik
P < 50,0	Tidak Baik

**HASIL DAN PEMBAHASAN**

Produk yang dihasilkan dari penelitian pengembangan ini berupa instrument tes HOTS (*Higher Order Thinking Skill*) berbasis kemampuan literasi matematis siswa SMP. Dengan menggunakan model *Formative Research Tessmer* yang terdiri dari tahap *preliminary, self evaluation, prototyping, dan field test*.

**1. Tahap Persiapan (Preliminary)**

Langkah awal dalam proses ini adalah mengumpulkan sejumlah referensi terkait penelitian, seperti literasi matematika dan instrumen tes kemampuan berpikir tingkat tinggi. Setelah sumber referensi sudah ada, selanjutnya dilakukan penelusuran perangkat pembelajaran dan wawancara kepada salah satu guru matematika kelas VIII mengenai kurikulum yang digunakan, tingkat kemampuan berpikir tingkat tinggi dan kemampuan literasi matematis siswa, serta bahan ajar yang digunakan.

**2. Tahap Evaluasi Diri (Self Evaluation)**

Pada tahap ini instrumen tes dirancang berdasarkan hasil wawancara dan penelusuran perangkat pembelajaran pada tahap sebelumnya. Instrumen tes yang dirancang mencakup kisi-kisi tes, soal tes sebanyak 12 butir soal uraian, kriteria jawaban soal tes, lembar kerja, dan pedoman penskoran. Pada tahap inilah dihasilkan prototipe 1 yang melalui dua tahap kegiatan yakni tahap analisis dan tahap desain. (1) Pada tahap analisis dilakukan analisis kurikulum, siswa, dan materi. Kegiatan analisis kurikulum dilakukan untuk memastikan bahwa tujuan yang dibutuhkan dalam pengembangan tes sesuai, sehingga tes dapat mengukur kemampuan berpikir tingkat tinggi dan literasi matematis siswa dengan baik. Berdasarkan hasil wawancara dan pengkajian dokumen perangkat guru diketahui bahwa kurikulum yang diterapkan pada kelas VIII di SMP Negeri 1 Campalagian adalah

Kurikulum 2013. Sementara itu, pada kegiatan analisis siswa dilakukan pengumpulan data tentang karakteristik siswa yang sesuai dengan instrumen yang dikembangkan. Dari hasil wawancara diperoleh bahwa kemampuan siswa baik kemampuan berpikir tingkat tinggi dan literasi matematis bervariasi, namun kebanyakan siswa berkemampuan rendah. Hal ini memungkinkan adanya faktor dari minat yang dimiliki setiap siswa berbeda-beda terhadap pelajaran matematika. Analisis siswa melibatkan 27 siswa kelas VIII A sebagai subjek penelitian, 6 siswa kelas VIII B dan 3 siswa kelas VIII C sebagai subjek uji coba. Kemudian pada analisis materi dilakukan sebagai proses untuk mengidentifikasi konsep-konsep utama yang digunakan dalam menyusun soal. Pada tahap ini dilakukan identifikasi materi dengan mengacu pada buku paket yang digunakan. Dari hasil wawancara diketahui bahwa pada materi bangun ruang sisi datar masih banyak siswa yang mengalami kesulitan dalam mengerjakan soal terutama pada soal HOTS serta berbentuk soal cerita. (2) Pada tahap desain, soal tes disusun dengan mengacu pada materi yang telah dianalisis yakni materi Bangun Ruang Sisi Datar serta indikator pada materi tersebut. Butir soal yang dikembangkan merupakan soal yang mewakili masing-masing jenis bangun ruang berdasarkan indikator materi. Soal-soal tersebut berupa soal uraian yang memiliki karakteristik sebagai soal HOTS berbasis literasi matematis.

**3. Tahap Membuat Prototipe (Prototyping)**

Pada tahap *prototyping*, instrumen tes yang telah dibuat diujicobakan dan direvisi seterusnya hingga menghasilkan *prototype 2*. Dilakukan melalui 3 proses uji coba, yaitu:

- a. *Expert review*, validasi oleh pakar pada uji coba *expert review* diperlukan 2 dosen Pendidikan Matematika Universitas Sulawesi Barat sebagai

validator ahli, 2 guru Matematika dan 1 guru Bahasa Indonesia SMP Negeri 1 Campalagian sebagai praktisi lapangan. Validator diberikan angket serta instrumen berupa kisi-kisi, soal tes, lembar kerja, kriteria jawaban dan pedoman penilaian untuk dianalisis. Berdasarkan tanggapan dan saran oleh pakar kemudian dilakukan revisi untuk selanjutnya diujicoba di proses berikutnya. Saran perbaikan dari validator diantaranya redaksi soal disesuaikan dengan PUEBI, beberapa kunci jawaban perlu

diperbaiki karena tidak akurat dan tidak sesuai dengan yang di lembar soal, gambar yang disajikan harus jelas dalam menggambarkan bentuk bangun ruang sisi datar, penulisan kalimat dan penggunaan tanda baca pada soal yang perlu perbaikan. Dari angket yang diberikan kepada validator akan diukur kelayakan dari aspek materi, konstruk, dan bahasa dari instrumen tes yang telah dikembangkan. Hasil validasi dapat dilihat pada tabel berikut:

**Tabel 7.** Hasil validasi instrumen tes

Validator	Skor	Persentase	Keterangan
Ahli	376	78,33%	Layak
Ahli	388	80,83%	Layak
Praktisi	434	90,42%	Layak
Praktisi	445	92,71%	Layak
Praktisi	376	87,04%	Layak

b. *One-to-one*, selain validasi oleh pakar, instrumen tes juga divalidasi oleh 3 siswa yang sebelumnya telah dipilih pada tahap analisis siswa. Siswa yang dipilih merupakan siswa yang mewakili karakteristik kemampuan berpikir tingkat tinggi dan literasi matematis yang berbeda (1 siswa berkemampuan rendah, 1 siswa berkemampuan sedang, dan 1 siswa berkemampuan tinggi). Ketiga siswa tersebut diminta untuk mengerjakan *prototype* yang telah dibuat dan memberikan tanggapan/ komentar. Dari komentar ketiga siswa hampir seluruh soal sudah bisa dipahami dengan baik, namun ada beberapa redaksi soal baik dari informasi umum maupun pertanyaan soal yang memerlukan perbaikan. Berdasarkan komentar dari tahap

*expert review* dan *one-to-one*, dilakukan revisi hingga menghasilkan *prototype* 2 untuk diuji coba ditahap selanjutnya.

c. *Small group*, hampir sama pada uji coba *one-to-one* namun jumlah siswa yang mengerjakan soal sebanyak 6 siswa sebaya (2 siswa berkemampuan rendah, 2 siswa berkemampuan sedang, 2 siswa berkemampuan tinggi). Hasil dari *prototype* 2 yang telah dikerjakan oleh siswa kemudian dianalisis mengenai kepraktisan, efektivitas, reliabilitas, tingkat kesukaran, dan daya pembeda soal tes.

Hasil analisis kepraktisan diperoleh dari data hasil angket respon oleh observer yakni guru matematika SMP Negeri 1 Campalagian. Hasil analisis disajikan pada tabel berikut.

**Tabel 8.** Hasil nilai kepraktisan instrumen tes

Aspek Penilaian	P (%)	Kriteria
Kejelasan bahasa, materi, dan simbol	93,67	Sangat Praktis
Kualitas isi	91,85	Sangat Praktis
Kesesuaian soal dengan HOTS dan literasi matematis	83,92	Sangat Praktis
Tingkat kesulitan butir soal	87,75	Sangat Praktis
<b>Rata-rata</b>	<b>89,30</b>	<b>Sangat Praktis</b>

Berdasarkan tabel diatas, diperoleh persentase rata-rata adalah 89,30 dengan kriteria sangat praktis. Dengan demikian instrumen tes yang telah dikembangkan dapat digunakan.

Sedangkan hasil analisis keefektifan yang diperoleh dari data analisis respon siswa, hasil analisis diuraikan pada tabel berikut.

**Tabel 9.** Hasil analisis keefektifan instrumen tes

Aspek Penilaian	P (%)	Kriteria
Kejelasan bahasa, materi, dan simbol	95,33	Sangat Baik
Kualitas isi	92,25	Sangat Baik
Kesesuaian soal dengan HOTS dan literasi matematis	89,93	Sangat Baik
Tingkat kesulitan butir soal	90,87	Sangat Baik
<b>Rata-rata</b>	<b>92,09</b>	<b>Sangat Baik</b>

Hasil analisis menunjukkan bahwa persentase rata-rata keefektifan instrumen tes adalah 92,09 dengan kriteria sangat baik. Menurut siswa instrumen tes yang telah dikembangkan mudah

dipahami baik dari segi bahasa maupun materi yang disajikan. Soal mampu mengukur kemampuan berpikir tingkat tinggi dan literasi matematis siswa serta materi yang disajikan sesuai dengan materi

pembelajaran. Sehingga dapat disimpulkan bahwa instrumen tes mendapat respon positif dari seluruh siswa.

Selanjutnya untuk nilai reliabilitas soal tes berdasarkan hasil nilai skor jawaban oleh siswa berada pada interpretasi ‘sangat tinggi’, tingkat kesukaran soal tes berada pada interpretasi ‘sedang’ dan ‘sukar’, serta daya pembeda setiap soal berada pada interpretasi ‘sangat baik’. Dengan demikian dapat disimpulkan bahwa tes yang telah dikembangkan dapat digunakan tanpa revisi. Hasil dari tahap ini diperoleh *prototype* 3 dan selanjutnya diujicobakan pada skala yang lebih luas (*field test*).

**4. Tahap Uji Lapangan (Field Test)**

Pada tahap ini *prototype* yang telah divalidasi dan direvisi diujicobakan dengan melibatkan 27 subjek yakni siswa kelas VIII A SMP Negeri 1 Campalagian. Uji coba dilakukan pada hari Senin 13 Mei 2024 dengan jumlah soal sebanyak 12 butir soal

uraian dengan alokasi waktu pengerjaan 2 × 45 menit.

Pada awal kegiatan pelaksanaan tes, siswa diberikan lembar soal dan lembar kerja. Setiap siswa menjawab pertanyaan pada lembar kerja yang tersedia dan dikumpulkan setelah waktu yang ditentukan selesai. Hasil nilai yang diperoleh dari tes yang dikerjakan siswa selanjutnya dianalisis kembali untuk mengetahui kelayakan dari setiap butir soal yang telah dikembangkan. Adapun uji kelayakan yang dilakukan terdiri dari:

a. Uji Validitas

Hasil uji validitas diperoleh dari perbandingan skor r tabel dan r hitung, skor r tabel yang digunakan yaitu 0,396 yang ditentukan berdasarkan tingkat kepercayaan 95% dengan responden sebanyak 27. Untuk rincian skor r hitung yang diperoleh diuraikan pada tabel berikut.

**Tabel 10.** Hasil uji validitas butir soal

Nomor Soal	Nilai R Hitung	Keterangan
1	0,832	Valid
2	0,938	Valid
3	0,825	Valid
4	0,878	Valid
5	0,810	Valid
6	0,888	Valid
7	0,911	Valid
8	0,838	Valid
9	0,807	Valid
10	0,792	Valid
11	0,690	Valid
12	0,712	Valid

Berdasarkan hasil uji validitas pada tabel 6 diatas, skor r hitung yang diperoleh dari tiap butir soal lebih besar dari r hitung. Maka dapat disimpulkan bahwa semua butir soal dinyatakan valid.

b. Uji Reliabilitas

Hasil uji reliabilitas instrumen tes dinyatakan reliabel berdasarkan hasil analisis reliabilitas yang diperoleh adalah 0,950 dengan interpretasi reliabilitas sangat tinggi. Hal ini dilihat dari kriteria kualitas instrumen tes yang

layak apabila memiliki derajat reliabilitas lebih dari 0,60 dan kurang dari 1. Berdasarkan teori yang dikemukakan oleh Basuki bahwa ciri soal yang memiliki kualitas layak adalah minimal valid dan reliable [13].

c. Uji Tingkat Kesukaran

Soal tes dapat dikatakan memiliki kualitas yang baik jika setiap butir soal memiliki tingkat kesukaran berada pada interval  $0,00 < P \leq 0,70$ , rincian indeks kesukaran tiap butir soal diuraikan pada tabel berikut.

**Tabel 11.** Hasil uji tingkat kesukaran butir soal

Nomor Soal	Indeks Kesukaran	Keterangan
1	0,63	Sedang
2	0,43	Sedang
3	0,56	Sedang
4	0,29	Sukar
5	0,61	Sedang
6	0,36	Sedang
7	0,22	Sukar
8	0,55	Sedang
9	0,22	Sukar
10	0,25	Sukar
11	0,52	Sedang
12	0,64	Sedang

Berdasarkan hasil uji tingkat kesukaran dari 12 soal yang telah dikembangkan terdapat 23% soal termasuk kategori sukar dan 77% soal termasuk kategori sedang. Jika dilihat pada kriteria ideal, proporsi tingkat kesukaran soal yang dikembangkan belum tersebar secara ideal. Hal ini dikarenakan proporsi tingkat kesukaran yang ideal berdasarkan teori yang dikemukakan oleh Sudjana yaitu menggunakan perbandingan 3-4-3, yaitu 30% soal mudah, 40% soal sedang, dan 30% soal sukar. Selain itu dapat juga

menggunakan perbandingan 3-5-2, yaitu 30% soal mudah, 50% soal sedang, dan 20% soal sukar[14].

d. Uji Daya Pembeda

Soal tes dapat dinyatakan memiliki kualitas yang baik jika butir soal tersebut memiliki nilai daya pembeda minimal 0,20 dalam artian daya pembeda berada pada kategori cukup. Adapun rincian hasil perhitungan daya pembeda tiap butir soal diuraikan pada tabel berikut.

**Tabel 12.** Hasil uji daya pembeda butir soal

Nomor Soal	Daya Pembeda	Keterangan
1	0,794	Sangat Baik
2	0,924	Sangat Baik
3	0,790	Sangat Baik
4	0,850	Sangat Baik
5	0,775	Sangat Baik
6	0,866	Sangat Baik
7	0,892	Sangat Baik
8	0,797	Sangat Baik
9	0,772	Sangat Baik
10	0,756	Sangat Baik
11	0,625	Baik
12	0,642	Baik

Berdasarkan hasil uji daya pembeda instrumen tes, terdapat 2 soal yang termasuk dalam kategori "baik" yang menunjukkan nomor soal mampu membedakan siswa berkemampuan tinggi dan rendah, dan 10 soal yang masuk dalam kategori "sangat baik" yang menunjukkan nomor soal sangat baik dalam membedakan siswa berkemampuan literasi matematis dan berpikir tingkat tinggi dan rendah. Berdasarkan kriteria kualitas daya pembeda butir soal yang baik menurut Arikunto, butir soal yang mampu membedakan siswa berkemampuan tinggi dan rendah masuk dalam kategori sangat baik dan baik [15].

menghasilkan instrumen tes *Higher Order Thinking Skill* berbasis literasi matematis pada materi Bangun Ruang Sisi Datar.

**DAFTAR PUSTAKA**

[1] K. Zaqiyah, L. Lutfiyah, and D. N. Sulisawati, "Pengembangan Modul Berbasis Realistic Mathematics Education untuk Pembelajaran Bangun Ruang Sisi Lengkung," *Laplace J. Pendidik. Mat.*, vol. 3, no. 2, pp. 151-162, Oct. 2020, doi: 10.31537/laplace.v3i2.381.

[2] R. Badjeber and J. P. Purwaningrum, "Pengembangan Higher Order Thinking Skills Dalam Pembelajaran Matematika di SMP," vol. 1, no. 1, pp. 36-43, 2018.

[3] O.: Umiyati, A. Basis, A. C. Manongko, and G. M. Tiwow, "Model Pengembangan Pembelajaran Higher Order Thingking Skill (HOTS) Pada Siswa di SMA Kristen 1 Tomohon."

[4] H. Saputra, *Pengembangan mutu pendidikan menuju era global: Penguatan mutu pembelajaran dengan penerapan hots (high order thinking skills)*. Smile's, 2016.

[5] H. Setiawan, N. Diah Sri Lestari, and P. Studi Pendidikan Matematika FKIP Universitas Jember, "Prosiding Seminar Nasional Matematika," 2014.

[6] A. R. As'ari, M. Ali, H. Basri, D. Kurniati, and S. Maharani, "Mengembangkan HOTS (higher order thinking skills) melalui matematika," *Univ. Negeri Malang*, 2019.

[7] P. Astuti, "Kemampuan Literasi Matematika dan Kemampuan Berpikir Tingkat Tinggi." [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>

[8] R. H. N. Sari, "Literasi matematika: apa, mengapa

**KESIMPULAN**

Instrumen tes *Higher Order Thinking Skill* berbasis literasi matematis pada materi Bangun Ruang Sisi Datar dapat dianggap valid, praktis, dan efektif berdasarkan hasil analisis oleh ahli yang menunjukkan bahwa tes valid untuk digunakan. Analisis kepraktisan instrumen tes menghasilkan skor rata-rata 89,30 yang termasuk dalam kategori "sangat praktis". Analisis efektivitas instrumen tes menghasilkan skor rata-rata 92,09 yang termasuk dalam kategori "sangat baik". Selain itu, uji validitas butir soal, reliabilitas, tingkat kesukaran, dan daya pembeda digunakan untuk menentukan hasil uji empiris oleh siswa. Sebanyak 12 butir soal dengan kriteria valid diperoleh dalam hasil uji validitas. Dengan interpretasi reliabilitas yang sangat tinggi, hasil uji reliabilitas dianggap dapat diandalkan. Tingkat kesukaran butir soal memperoleh 4 soal dalam kategori "sulit" dan 8 soal dalam kategori "sedang". Sebanyak 10 soal termasuk kategori "sangat baik" dan 2 soal termasuk kategori "baik" diperoleh dari analisis daya pembeda soal. Dengan demikian, penelitian ini

- dan bagaimana,” in *Seminar Nasional matematika dan pendidikan matematika UNY*, 2015, pp. 713–720.
- [9] R. Faisal, “Pengembangan paket tes kemampuan berpikir tingkat tinggi (higher order thinking skill) berdasarkan taksonomi Bloom pada materi matematika kelas VII SMP,” 2015.
- [10] Z. Arifin and H. Retnawati, “Pengembangan instrumen pengukur higher order thinking skills matematika siswa SMA kelas X,” *PYTHAGORAS J. Pendidik. Mat.*, vol. 12, no. 1, p. 98, Jun. 2017, doi: 10.21831/pg.v12i1.14058.
- [11] N. Rahma Baddu, N. Arsyad, and M. ' Rup, “Infinity: Jurnal Matematika dan Aplikasinya (IJMA) PENGEMBANGAN INSTRUMEN TES HIGHER ORDER THINKING SKILL (HOTS) UNTUK MATERI STATISTIKA DI KELAS IX SMP,” vol. 3, no. 1, pp. 19–28, 2022.
- [12] M. Tessmer, *Planning and conducting formative evaluations*. Routledge, 2013.
- [13] I. Basuki and A. P. Hariyanto, “Bandung: PT,” *Remaja Rosdakarya*, 2014.
- [14] S. Nana, “Penilaian Hasil Belajar Proses Belajar Mengajar: Bandung: PT,” *Remaja Rosdakarya*, 2006.
- [15] S. Arikunto, *Dasar-dasar evaluasi pendidikan edisi 3*. Bumi aksara, 2021.