

Artikel Novia Dwi - 4

by Novia Dwi

Submission date: 09-Feb-2022 03:22PM (UTC+0700)

Submission ID: 1758385864

File name: Author_Novia_TIM_Jurnal_Aksioma_2022.docx (670.31K)

Word count: 3700

Character count: 22860

Novia Dwi Rahmawati^{1*}, Komarudin², Suherman³

^{1*}Universitas Hasyim Asy'ari, Jombang, Indonesia

² Universitas Islam Negeri Raden Fatah, Palembang, Indonesia

³ Universitas of Szeged, Hungary

*Corresponding author. Jl. Irian Jaya No 55 Tebuireng Jombang

E-mail: noviadwirahmawati@unhasy.ac.id ^{*1)}

qhomar8@gmail.com ²⁾

suherman@edu.u-szeged.hu ³⁾

Received dd Month yy; Received in revised form dd Month yy; Accepted dd Month yy (9pt)

Abstrak

Penelitian ini bertujuan untuk mengembangkan instrumen penilaian berupa soal-soal matematika berbasis kemampuan berpikir tingkat tinggi (HOTS). Pengembangan soal mengacu pada model Borg dan Gall (2003). Untuk mengetahui validitas instrumen, soal yang dikembangkan divalidasi oleh ahli (validasi konstruk). Pada tahap uji coba, 15 calon guru SD diberikan 18 soal. Hasil analisis jawaban subjek menunjukkan bahwa semua soal matematika dasar berbasis HOTS yang dikembangkan valid dan reliabel. Hasil analisis daya diskriminatif menunjukkan bahwa semua soal memiliki indeks diskrepansi lebih besar dari 0,3 yang berarti dapat membedakan kemampuan siswa. Selain itu, semua soal yang dikembangkan memiliki indeks tingkat kesukaran antara 0,3-0,7 yang berarti semua butir soal yang dikembangkan memiliki tingkat kesukaran yang baik. Hasil pengembangan soal dalam penelitian ini dapat digunakan sebagai alat penilaian HOTS dalam pendidikan calon guru SD dan menjadi acuan calon guru dalam mengembangkan masalah matematika berbasis HOTS.

Kata kunci: Calon Guru Sekolah Dasar; HOTS; Pengembangan soal

Abstract

This study aims to develop an assessment instrument in the form of math questions based on higher-order thinking skills (HOTS). The development of the questions refers to the model of Borg and Gall (2003). To determine the validity of the instrument, the questions developed were validated by experts (construct validation). In the pilot phase, 15 prospective primary school teachers were given 18 questions. The results of the analysis of the subjects' answers showed that all the HOTS-based basic mathematics questions developed were valid and reliable. The results of the discriminative power analysis show that all the questions have a discrepancy index greater than 0.3, which means that they can distinguish the abilities of the students. In addition, all the questions developed have a difficulty level index between 0.3-0.7, which means that all the items developed have a good level of difficulty. The results of question development in this study can be used as a HOTS assessment tool in the education of future primary school teachers and become a reference for future teachers in developing HOTS-based mathematical problems.

Keywords: Development of tests; HOTS; Preservice elementary teacher



This is an open access article under the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

PENDAHULUAN

Higher order thinking skills (HOTS) adalah salah satu aspek penting dalam pendidikan (Pahdi, Mailizar, & Abidin, 2020; Tanujaya, Mumu, &

Margono, 2017; Zaid et al. 2018; Rahmawati et al. 2020; Mulyatna et al. 2021). Pengembangan kemampuan berpikir tersebut sangat penting untuk mempersiapkan siswa menghadapi

DOI: <https://doi.org/10.24127/ajpm>

tantangan di masa depan dan membantu mereka memecahkan masalah (Yunita et al. 2020). HOTS memiliki beberapa aspek, diantaranya menerapkan, menganalisis, mengevaluasi, dan membuat (Misrom et al. 2020). Selain itu, kemampuan berpikir kritis dan kreatif juga termasuk kemampuan berpikir tingkat tinggi (Husamah, Fatmawati, & Setyawan, 2018; Yusuf et al. 2020) dan pemecahan masalah (Friyatmi, Mardapi, & Haryanto, 2020; Priyaadharshini & Sundaram, 2018; Yang, 2015). Dasar pengkategorian meliputi berpikir kritis, berpikir kreatif, berpikir metakognitif dan mengambil keputusan (Feldhusen & Goh, 1995; Jerome, Lee, & Ting, 2017; Mohamad 2015). Berdasarkan pendapat beberapa ahli di atas, dalam penelitian ini representasi HOTS terdiri dari 3 aspek, yaitu menganalisis, mengevaluasi, dan mencipta. **Matematika merupakan salah satu mata pelajaran yang mempunyai peranan dalam mengembangkan HOTS siswa (Tanujaya, Prahmana, & Mumu, 2017).** Dalam hal ini, diperlukan alat evaluasi yang dapat mengukur capaian HOTS mahasiswa. Penggunaan asesmen HOTS dalam asesmen pendidikan merupakan isu penting yang harus direspon oleh semua pelaku pendidikan (Widana, 2018; Widana et al., 2018). Untuk memperoleh instrumen yang berkualitas, selain analisis teoritis (peninjauan butir soal berdasarkan aspek isi, konstruksi, dan kebahasaan) analisis butir soal empiris juga perlu dilakukan.

Penelitian tentang HOTS telah dilakukan oleh (Arafah et al., 2021) guna menentukan karakteristik instrumen tes HOTS pada materi mekanika SMA. Mereka mengembangkan instrument tes mengacu pada classical theory berdasarkan model pengembangan

Mardapi. Penelitian serupa dilakukan oleh (Septiani & Paidi, 2021) guna mengukur kemampuan kognitif biologi siswa SMA. Mereka mengembangkan instrumen tes HOTS berbentuk pilihan ganda berdasarkan model pengembangan Oriondo & Dallo-Antonio dan Wilson. Penelitian serupa juga dilakukan oleh (Arip et al., 2018) guna mengembangkan, menguji nilai validitas konten dan keandalan Modul Pelatihan HOTS pada materi Sains SMA. Mereka mengembangkan modul HOTS berdasarkan metode deskriptif menggunakan teknik survei. (Bahtiar et al., 2020) juga melakukan penelitian guna mengembangkan instrument soal HOTS pada materi fisika berdasarkan kearifan lokal pada jenjang SMA. Penelitian tersebut mengembangkan instrument tes berdasarkan empat (4) tahap: mendefinisikan, merancang, mengembangkan, dan mendistribusikan. Dalam penelitian ini, soal HOTS dikembangkan pada materi matematika untuk calon guru sekolah dasar berdasarkan langkah pengembangan Borg & Gall. Temuan ini dapat dijadikan alat evaluasi HOTS dalam pendidikan calon guru SD dan menjadi rujukan calon guru dalam mengembangkan soal matematika berbasis HOTS. Kemampuan menganalisis, mengevaluasi, dan mencipta merupakan salah satu kemampuan dasar yang wajib dikuasai oleh calon guru (mahasiswa) khususnya pada mata kuliah matematika. Selain itu, sebagai upaya pembiasaan dalam mendesain pembelajaran yang berbasis pada HOTS. Penelitian ini bertujuan untuk mengembangkan instrumen soal matematika berbasis high order thinking skills (HOTS) untuk calon guru sekolah dasar.

Berbicara tentang HOTS, kami merujuk pada Dillon & Scott (2002;

DOI: <https://doi.org/10.24127/ajpm>

Miri, David, & Uri (2007); Zohar & Dori, 2003) yang mendefinisikan HOTS sebagai suatu kemampuan menganalisis, mensintesis, mengevaluasi, mengembangkan keterampilan, memperkirakan, generalisasi, dan menciptakan pemikiran, membuat keputusan, mengatur tujuan, berpikir kritis dan sistemik. Berdasarkan definisi, menunjukkan adanya hubungan yang erat antara HOTS dan kemampuan berpikir matematis. Kemampuan ini digunakan ketika seseorang menerima informasi baru dan menyimpannya untuk digunakan atau diatur ulang untuk tujuan pemecahan masalah. Merujuk pendapat Benyamin Bloom (1965) tentang level HOTS dalam taksonomi Bloom terdiri dari menganalisis, mengevaluasi, dan mencipta (Domain 1956; Kantar 2014; Seman, Yusoff, & Embong 2017). Dalam taksonomi ini, siswa dituntut tidak hanya memahami dan menerapkan konsep matematika, tetapi juga mampu menganalisis dan menyelesaikan dengan cara tepat, bahkan tidak menutup kemungkinan menemukan hal atau penyelesaian yang baru. Sehingga membuat siswa terbiasa dengan soal-soal yang membutuhkan kemampuan analisis yang tinggi. Hal ini sejalan dengan pendapat Rahmawati (2020), keterampilan berpikir tingkat tinggi yang melibatkan keterampilan analitis, evaluatif dan kreatif ditetapkan sebagai tujuan pendidikan.

Membahas tentang *classical test theory*, kami merujuk pendapat (Allen & Yen, 1979; Ruzbarsky et al., 2018) yaitu skor yang diamati adalah jumlah skor aktual dan skor kesalahan pengukuran. Secara matematis ditulis sebagai $X = T + E$. Dimana X adalah skor yang diamati, T adalah skor sebenarnya dan E adalah *error score*. *Error* ini merupakan penyimpangan teoritis dari skor observasi yang

diperoleh terhadap skor observasi yang diharapkan. Sedangkan kesalahan pengukuran sistematis tidak dianggap sebagai kesalahan pengukuran.

Merujuk pada asumsi-asumsi teori klasik merupakan dasar untuk mengembangkan berbagai rumusan yang berguna dalam mengukur butir-butir soal antara lain daya pembeda, indeks kesukaran, efektivitas distraktor, reliabilitas, dan validitas (Suprpto et al., 2020). Teori tes klasik diambil karena memiliki kelebihan dibandingkan teori respon item, salah satunya terkait dengan analisis yang lebih sederhana dan mudah dipahami.

METODE PENELITIAN

Pengembangan instrumen soal HOTS dalam penelitian ini merujuk pada model Borg & Gall (1983). Model Borg & Gall dipilih karena lebih sistematis dan sesuai dengan tujuan penelitian yang dilakukan, yaitu menghasilkan Soal HOTS matematika untuk calon guru sekolah dasar yang valid, menarik dan efektif untuk pembelajaran. Langkah-langkah model pengembangan Borg & Gall dalam penelitian dan pengembangan ini, yaitu:

Menganalisis kebutuhan (potensi masalah)

Sebelum melakukan pengembangan soal matematika berbasis HOTS, penelitian melakukan analisis kebutuhan. Hal ini bertujuan mengidentifikasi masalah yang akan dicarikan solusi melalui pengembangan sebuah produk. Hasil analisis kebutuhan menunjukkan tantangan besar yang dihadapi guru matematika terkait kemampuan berpikir tingkat tinggi

DOI: <https://doi.org/10.24127/ajpm>

siswa. Hasil PISA (*Programme for International Student Assessment*) diketahui bahwa kemampuan literasi matematika siswa Indonesia termasuk kategori rendah. Hal ini dibuktikan dengan nilai rata-rata kemampuan matematika yang diperoleh siswa Indonesia tahun 2015 sebesar 380 dari 490 poin yang artinya masih di bawah rata-rata internasional (OECD 2017). Hal ini bisa disebabkan oleh beberapa faktor, yaitu kurangnya kemampuan guru dalam mengembangkan soal-soal HOTS (Riswanda 2018) dan calon guru belum dipersiapkan secara maksimal untuk memahami dan menguasai HOTS (Maryani & Martaningsih 2020). Analisis kebutuhan tersebut menunjukkan pentingnya pengembangan soal matematika berbasis HOTS.

Mengumpulkan informasi

Analisis kebutuhan (potensi masalah) yang ditemukan menjadi dasar untuk mendapatkan informasi yang sesuai untuk bahan pengembangan produk. Penelitian ini mengumpulkan informasi melalui tinjauan literatur yang berkaitan dengan kriteria soal HOTS berdasarkan taksonomi Bloom revisi yaitu, C4 (Analisis), C5 (Evaluasi), dan C6 (mencipta). Hal ini dikarenakan taksonomi Bloom revisi dianggap lebih relevan dalam mengukur indikator HOTS. Selain itu, indikator dan tujuan pembelajaran dirumuskan dengan mengacu pada taksonomi Bloom revisi (Arifin & Retnawati 2017).

Indikator Taksonomi Bloom revisi (menganalisis, mengevaluasi dan mencipta) merupakan indikator yang

masih bersifat umum. Soal HOTS yang diajukan tidak hanya terfokus pada isi materi tetapi juga mengukur seluruh kompetensi dasar yang terkandung dalam materi. Selain kriteria menganalisis (C4), mengevaluasi (C5) dan mencipta (C6), dalam penelitian ini juga mengukur keterampilan dasar yang berhubungan dengan sifat-sifat bangun datar.

Mendesain produk

Penulis mendesain instrumen asesmen berupa soal matematika dasar berbasis HOTS yang terdiri dari Kisi-kisi soal HOTS, 18 soal HOTS, kunci jawaban, dan rubrik penskoran. Instrumen evaluasi disusun menurut kaidah-kaidah mengajukan pertanyaan uraian secara umum, yaitu menggunakan rubrik penilaian yang memperhatikan ciri-ciri dan tingkat kesukaran butir-butir soal. (Arifin & Retnawati 2017). Rubrik penskoran ini menerapkan kriteria yang berbeda pada tiap butir soalnya, tergantung pada karakteristik dan tingkat kesukaran yang dimiliki soal. Semakin tinggi tingkat kesukaran soal, maka semakin tinggi juga bobot soal tersebut. Selain mengacu pada kriteria taksonomi Bloom revisi, rubrik penskoran ini juga didasarkan pada langkah-langkah dan konsep yang digunakan untuk memecahkan masalah matematis.

Memvalidasi desain

Pada tahap ini, penelitian menyusun instrumen validasi produk berupa angket, yaitu daftar pertanyaan yang diberikan kepada ahli untuk ditanggapi sesuai permintaan peneliti

DOI: <https://doi.org/10.24127/ajpm>

yang telah dimodifikasi sesuai kebutuhan khususnya sesuai dengan kriteria pertanyaan HOTS. Suatu instrumen dikatakan valid jika dapat mengukur apa yang hendak diukur (Oosterhaven, Ofenloch, & Schuttelaar 2020; Selaras et al. 2019). Tujuan validasi desain dalam penelitian ini adalah agar peneliti mengetahui kekurangan dan kelebihan produk yang telah dihasilkan, serta melalui kritik dan saran dari para ahli. Rancangan produk yang telah dibuat, kemudian divalidasi oleh validator ahli (validasi konstruk). Data hasil validasi berupa data kualitatif dan kuantitatif. Data kualitatif berupa komentar atau masukan dari validator ahli yang berjumlah 3 orang. Sedangkan data kuantitatif berupa skor penilaian akan dianalisis menggunakan persentase nilai. Kriteria validitas instrumen dapat dilihat pada Tabel 1 (Aniq Rif'atun, Najihah; Vina, Serevina; Mutia 2018).

Tabel 1. Kriteria validitas instrumen

Skala skor validitas	Kriteria validitas	Keputusan
85-100	Sangat valid	Tanpa revisi
70-84	Valid	Tanpa revisi
55-69	Cukup valid	Tanpa revisi
50-54	Kurang valid	Revisi sebagian
0-49	Tidak valid	Dihapuskan

Merevisi desain

Soal HOTS yang telah divalidasi, kemudian direvisi oleh penulis. Ahli assessmen memberikan kritik dan saran terkait rubrik penilaian untuk memberikan contoh kesalahan yang dapat diberi skor 1 atau 2. Sedangkan ahli materi

memberikan kritik dan saran terkait notasi penulisan geometri, satuan volume yang digunakan, gambar posisi bola dalam tabung pada soal no 8 belum menunjukkan representasi dari soal.

Contoh draf instrumen soal yang telah direvisi ahli validasi adalah sebagai berikut.

Soal HOTS untuk menganalisis sebelum validasi

“Radhika dan Mahendra akan berkemah dengan menggunakan tenda dengan alasnya berbentuk persegi panjang dengan panjang 10 m, lebar 6 m dan tingginya 0,5 m. Luas permukaan tenda 143 m² dengan volume tenda 150 m³. Bagian manakah dari tenda itu yang belum diketahui dan tentukan ukurannya?”

Soal di atas mengukur indikator HOTS yaitu analisis, terkhusus pada sub-indikator mengatribusi. Sedangkan pengukuran pada dimensi pengetahuan mencakup dimensi pengetahuan konseptual, dimensi prosedural, dan dimensi metakognitif.

Soal HOTS untuk menganalisis sesudah validasi

“Radhika dan Mahendra akan berkemah dengan menggunakan tenda yang memiliki alas berbentuk persegi panjang dengan panjang 10 m, lebar 6 m dan tingginya 0,5 m. Luas permukaan tenda 143 m² dengan volume tenda 150 m³. Bagian manakah dari tenda

DOI: <https://doi.org/10.24127/ajpm>

itu yang belum diketahui dan tentukan ukurannya?"

Menguji coba produk

Pada tahap ini, soal yang dikembangkan kepada 15 mahasiswa. Uji coba dilaksanakan secara mandiri dengan pengawasan penulis. Hasil uji coba dianalisis secara empirik dengan pendekatan teori tes klasik. Untuk menentukan kualitas tes dengan menggunakan teori tes klasik yaitu validitas, reliabilitas, indeks kesukaran, dan daya pembeda (Mahjabeen et al. 2018; Sharma 2021) menggunakan bantuan SPSS versi 20.

❖ Uji Validitas

Instrumen dikatakan valid, jika dapat mengukur kemampuan siswa secara akurat. Validitas ini dapat dihitung dengan koefisien korelasi menggunakan *product moment* sebagai berikut:

$$r_{xy} = \frac{n \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \cdot \sum_{i=1}^n Y_i}{\sqrt{[n \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2][n \sum_{i=1}^n Y_i^2 - (\sum_{i=1}^n Y_i)^2]}}$$

Kemudian dicari *corrected item-total correlation coefficient* dengan rumus berikut:

$$r_{x(y-1)} = \frac{r_{xy} S_y - S_x}{\sqrt{S_y^2 + S_x^2 - 2r_{xy}(S_y)(S_x)}}$$

Keterangan:

- x_i : skor siswa pada butir soal ke- i
- y_i : skor total siswa ke- i
- r_{xy} : skor koefisien korelasi pada butir soal ke- i
- S_y : standar deviasi total
- S_x : standar deviasi butir soal ke- i
- $r_{x(y-1)}$: *corrected item-total correlation coefficient*

Nilai $r_{x(y-1)}$ akan dibandingkan dengan koefisien korelasi tabel $r_{tabel} = r_{(\alpha, n-2)}$. Jika $r_{x(y-1)} \geq r_{tabel}$, maka instrumen valid.

❖ Uji Reliabilitas

Instrumen dikatakan reliabel, jika mengukur apa yang akan diukur dengan hasil konsisten. Uji reliabilitas soal menggunakan rumus *Cronbach Alpha* berikut.

$$r_{11} = \left[\frac{k}{k-1} \right] \left[1 - \frac{\sum_{i=1}^n S_i^2}{S_t^2} \right]$$

Keterangan:

- r_{11} : Koefisien reliabilitas
- k : Jumlah soal
- $\sum_{i=1}^n S_i^2$: jumlah varians skor masing-masing soal
- S_t^2 : Varians total

Koefisien reliabilitas (r_{11}) selanjutnya di konsultasikan terhadap $r_{tabel} = r_{(\alpha, n-2)}$. Jika $r_{11} > r_{tabel}$, maka instrumen reliabel.

❖ Uji Tingkat Kesukaran Soal

Uji tingkat kesulitan digunakan untuk mengetahui tingkat kesulitan suatu butir soal. Uji tingkat kesukaran soal digunakan rumus berikut.

$$I = \frac{B}{J}$$

Keterangan:

- I : indeks kesukaran butir soal
- B : banyak siswa yang menjawab benar soal ke- i .
- J : banyak siswa yang menjawab soal ke- i .

Semakin kecil indeks kesukaran yang didapat, makin sukar soal tersebut dan sebaliknya. Kriteria tingkat kesukaran disajikan pada Tabel 2.

Tabel 2. Tingkat kesukaran butir soal

Besar indeks	Kriteria
$0,00 \leq I \leq 0,30$	Sukar

DOI: <https://doi.org/10.24127/ajpm>

$0,30 < I \leq 0,70$	Sedang
$0,70 < I \leq 1,00$	Mudah

❖ Uji Daya Beda

Uji daya pembeda digunakan untuk mengetahui sejauh mana setiap butir soal mampu membedakan siswa dengan kemampuan tinggi dan rendah. Uji daya pembeda digunakan rumus berikut.

$$DB = PT - PR$$

Keterangan:

DB : daya pembeda

PT : proporsi kelompok tinggi

PR : proporsi kelompok rendah

Hasil dari perhitungan daya pembeda dibandingkan dengan kriteria indeks daya pembeda. Klasifikasi yang digunakan untuk menginterpretasikan indeks daya pembeda disajikan pada Tabel 3.

Tabel 3. Klasifikasi daya pembeda

Nilai	Interpretasi Daya Pembeda
0,00	Sangat buruk
$0,00 < DB \leq 0,20$	Buruk
$0,20 < DB \leq 0,40$	Cukup
$0,40 < DB \leq 0,70$	Baik
$0,70 < DB \leq 1,00$	Sangat baik

Merevisi Produk

Berdasarkan hasil uji coba produk diketahui bahwa beberapa soal perlu adanya revisi karena dianggap kurang komunikatif/sukar dipahami oleh siswa, seperti berikut ini:

“Sebuah tangki berbentuk tabung dengan luas daerah bidang lengkung $200\pi \text{ cm}^2$ dan luas seluruh permukaan tabung $250\pi \text{ cm}^2$ dengan tinggi 20 cm. Bagian manakah dari tangki itu yang belum diketahui dan tentukan ukurannya?.”

Berikut adalah hasil revisi berdasarkan saran validator ahli:

“Rizki memiliki sebuah tangki berbentuk tabung dengan luas daerah bidang lengkung $200\pi \text{ cm}^2$ dan luas seluruh permukaan tabung $250\pi \text{ cm}^2$ dengan tinggi 20 cm. Bagian manakah dari tangki itu yang belum diketahui dan tentukan ukurannya?.”

HASIL DAN PEMBAHASAN

Hasil analisis uji coba instrument tes yang terdiri dari uji validasi konstruk, uji validasi, uji reliabilitas, uji daya beda dan uji tingkat kesukaran sebagai berikut.

Hasil validasi konstruk

Hasil validasi konstruk mendapatkan nilai kuantitatif 92. Apabila mengacu pada kriteria kevalidan Arikunto (2010) maka dikatakan soal yang dikembangkan sangat valid. Hal ini berarti hasil pengembangan soal dalam penelitian ini dapat digunakan sebagai alat evaluasi HOTS dalam pendidikan calon guru SD.

Hasil validasi

Berdasarkan hasil analisis data uji validitas terhadap 18 butir soal yang dikembangkan disajikan pada tabel 4.

Tabel 4. Hasil uji validitas butir soal

Item-Total Statistics

	Scale Mean if Deleted	Scale Variance if Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Deleted

DOI: <https://doi.org/10.24127/ajpm>

Butir1	49.50	428.111	.380	.724
Butir2	49.54	436.776	.113	.730
Butir3	49.43	436.032	.120	.730
Butir4	49.57	435.143	.160	.729
Butir5	49.75	431.009	.245	.727
Butir6	49.57	418.847	.562	.717
Butir7	49.68	431.856	.198	.728
Butir8	49.61	418.396	.616	.717
Butir9	49.68	419.485	.500	.718
Butir10	49.54	403.591	.782	.706
Butir11	49.64	405.646	.713	.708
Butir12	49.82	414.374	.607	.714
Butir13	49.39	403.803	.780	.706
Butir14	49.21	394.471	.865	.698
Butir15	49.07	400.884	.724	.704
Butir16	48.86	405.090	.705	.707
Butir17	48.71	409.989	.615	.711
Butir18	48.32	420.300	.513	.719
Skor	24.61	110.470	.931	.884

Berdasarkan kriteria validitas butir soal yaitu jika $r_{x(y-1)} \geq r_{tabel}$, maka butir soal dikatakan valid dan jika $r_{x(y-1)} < r_{tabel}$, maka butir soal dikatakan tidak valid. Dengan $\alpha = 5\%$ dan $n = 18$, diperoleh $r_{tabel} = 0,444$. Berdasarkan kriteria uji validitas tersebut terlihat bahwa $r_{x(y-1)} \geq 0,444$, atau dengan kata lain 18 butir yang dikembangkan berkategori valid atau dapat digunakan untuk melakukan mengukur.

Hasil Uji Reliabilitas

Berdasarkan hasil analisis uji reliabilitas terhadap 18 butir soal dengan menggunakan SPSS versi 20 diperoleh koefisien reliabilitas (r_{11})

sebesar 0,730. Dengan $\alpha = 5\%$, diperoleh $r_{tabel} = 0,468$. Berdasarkan kriteria uji reliabilitas terlihat bahwa $r_{11} > 0,468$, yang artinya soal yang dikembangkan reliabel.

Daya pembeda

Berdasarkan hasil analisis data uji daya beda terhadap 18 butir soal yang dikembangkan akan disajikan pada tabel 5.

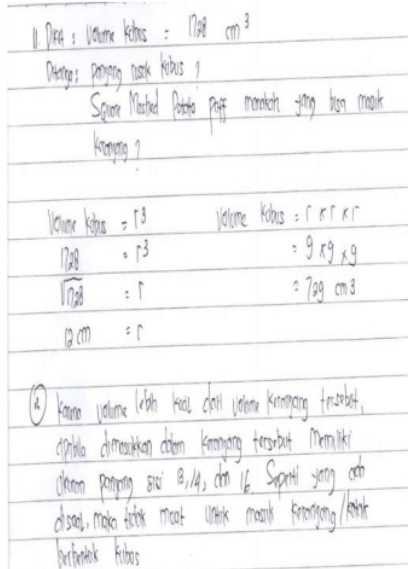
Tabel 5. Hasil uji daya beda

No soal	Skor uji daya beda	Kriteria daya beda
1	0,88	Baik
2	0,58	Baik
3	0,46	Baik
4	0,58	Baik
5	0,44	Baik
6	0,58	Baik
7	0,76	Baik
8	0,80	Baik
9	0,64	Baik
10	0,58	Baik
11	0,50	Baik
12	0,44	Baik
13	0,80	Baik
14	0,80	Baik
15	0,58	Baik
16	0,58	Baik
17	0,76	Baik
18	0,58	Baik

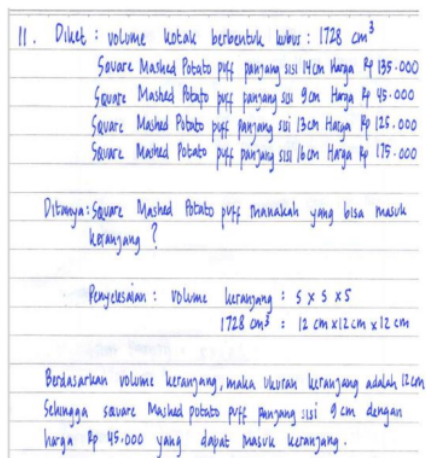
Tabel 5 menunjukkan bahwa semua butir soal memiliki kriteria daya beda baik, yang artinya semua butir soal dapat membedakan siswa yang kemampuan tinggi dan siswa yang kemampuannya rendah. Hasil analisis daya beda di atas didukung oleh jawaban siswa saat mengerjakan instrumen assessmen berupa soal matematika dasar berbasis high order thinking skills (HOTS). Enam dari 15 subjek siswa calon guru sekolah dasar

DOI: <https://doi.org/10.24127/ajpm>

memiliki kemampuan matematika rendah. Dikarenakan soal memiliki daya beda yang baik, ke-6 subjek tersebut menjawab tidak sama (salah) jika dibandingkan dengan 9 subjek yang lain



Gambar 1. Jawaban soal no. 11 oleh subjek dengan kemampuan rendah



Gambar 2. Jawaban soal no. 11 oleh subjek dengan kemampuan tinggi

Gambar 1 dan 2 menunjukkan perbedaan jawaban yang dihasilkan oleh siswa. Gambar 2 memperlihatkan bahwa siswa menuliskan apa yang diketahui dengan detail, begitu pula dengan cara menjawab dan menyimpulkannya sudah benar. Namun berbeda pada gambar 1 bahwa siswa tidak secara detail menuliskannya dan menyimpulkannya masih salah. Hal ini terjadi perbedaan pola menguraikan suatu permasalahan dan menyelesaikannya sehingga dapat ditarik kesimpulan bahwa daya pembeda soal berfungsi dengan baik, sehingga dapat membedakan siswa yang memiliki kemampuan tinggi dan rendah.

Tingkat kesukaran

Berdasarkan hasil analisis data uji tingkat kesukaran terhadap 18 butir soal yang dikembangkan akan disajikan pada Tabel 6.

Tabel 6. Hasil uji tingkat kesukaran

No soal	Skor uji	Kriteria
1	0,34	Sedang
2	0,34	Sedang
3	0,30	Sedang
4	0,31	Sedang
5	0,39	Sedang
6	0,34	Sedang
7	0,41	Sedang
8	0,34	Sedang
9	0,34	Sedang
10	0,50	Sedang
11	0,51	Sedang
12	0,54	Sedang
13	0,34	Sedang
14	0,58	Sedang
15	0,39	Sedang
16	0,56	Sedang
17	0,58	Sedang
18	0,39	Sedang

Tabel 7 menunjukkan bahwa semua butir soal memiliki indeks tingkat


DOI: <https://doi.org/10.24127/ajpm>

kesukaran sedang atau skor berada diinterval $0,30 < I \leq 0,70$, artinya instrumen tes yang dikembangkan memiliki tingkat kesukaran yang baik, yaitu tidak terlalu sukar dan tidak terlalu mudah. Hal ini sejalan dengan pendapat Suharsimi (2013) yang menyatakan bahwa Soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar. Berikut adalah beberapa contoh jawaban siswa dalam menyelesaikan soal yang diberikan.


6. Diket: Jari-jari tabung = 3,5 dm
Tinggi tabung = 10 dm
 $\pi = \frac{22}{7}$
Panjang Balok = 2 dm
Lebar Balok = 2 dm

Ditanya: Berapa buah tangki yang dibutuhkan Narendra?

Dijawab:
Volume tabung = luas alas x tinggi
 $= \pi \cdot r^2 \cdot t$
 $= \frac{22}{7} \times 3,5 \times 3,5 \times 10$
 $= 38,5 \times 10$
 $= 385 \text{ dm}^3$



Volume Balok = $p \times l \times t$
 $= 2 \times 2 \times 3,5$
 $= 14 \text{ dm}^3$



Banyaknya tangki yang dibutuhkan Narendra = $385 : 14$
ada 27 tangki //

Gambar 3. Jawaban soal no. 6 oleh subjek dengan kemampuan rendah

6. Diket: Jari-jari tabung = 3,5 dm
tinggi tabung = 10 dm
Balok berbentuk balok dengan panjang dan lebarnya 2 dm tinggi 3,5 dm
 $\pi = \frac{22}{7}$

Ditanya: Berapa buah tangki yg dibutuhkan Narendra untuk menyimpan premium?

Penyelesaian:
Banyaknya tangki premium = Volume drum yg berbentuk tabung
 $= \frac{22}{7} \times 3,5^2 \times 10$
 $= 385 \text{ dm}^3$

Volume tangki = Volume balok.
 $= p \times l \times t$
 $= 2 \times 2 \times 3,5$
 $= 14 \text{ dm}^3$

Banyaknya tangki yang harus disediakan Narendra adalah $385 : 14 = 27,5$ buah tangki, jadi Narendra harus menyediakan 28 buah tangki berbentuk balok.

Gambar 4. Jawaban soal no. 6 oleh subjek dengan kemampuan tinggi

Gambar 5 dan 6 menunjukkan jawaban subjek yang memiliki kemampuan rendah dan tinggi menjawab soal dengan benar. Hal ini dikarenakan soal termasuk baik yaitu soal tidak terlalu mudah atau tidak terlalu sukar.

Pengembangan instrumen assessmen berupa soal matematika dasar berbasis HOTS yang terdiri dari 18 butir soal diatas merupakan instrumen dengan kualitas baik. Hal itu sesuai dengan pernyataan (Hernawati & Suparman 2020) bahwa untuk mendapatkan instrumen berkualitas tinggi perlu dilakukan analisis validasi kontruk dan analisis secara empirik. Selain itu bahwa soal yang baik adalah soal yang dapat memperhatikan kemampuan dalam berpikir siswa (Sa'idah, Yulistianti, & Megawati 2018). Hasil pengembangan soal tersebut diharapkan sebagai alternatif dalam memperkaya kreativitas calon guru sekolah dasar dalam mengembangkan soal-soal matematika berbasis HOTS. Penelitian ini terletak bahwa secara kuantitas jumlah subjek uji coba dalam penelitian ini terbatas dan sedikit, oleh karena itu diperlukan pengembangannya untuk

DOI: <https://doi.org/10.24127/ajpm>

dapat meningkatkan kemampuan siswa (Suherman et al. 2020).

KESIMPULAN DAN SARAN

Berdasarkan pembahasan di atas, maka dapat disimpulkan bahwa butir soal matematika dasar berbasis HOTS yang dikembangkan memiliki kriteria valid ($r_{x(y-1)} \geq 0,444$) yang berarti seluruh butir soal dapat dipakai untuk melakukan mengukur, reliabel ($r_{11} > 0,468$), indeks daya beda yang baik ($DB > 0,40$) yang berarti dapat membedakan kemampuan siswa dan indeks tingkat kesukaran sedang ($0,30 < I \leq 0,70$) yang berarti butir soal yang dikembangkan memiliki tidak kesukaran yang baik. Sehingga telah memenuhi semua kriteria uji instrument tes baik, yaitu valid, reliabel, daya beda yang baik dan tingkat kesukaran sedang, atau dengan kata lain soal matematika pada materi geometri dapat digunakan untuk mengukur HOTS siswa.

DAFTAR PUSTAKA

- Allen, M. J., & Yen, W. M. (1979). Introduction to measurement theory, Monterey, CA: Brooks/Cole, 1979. *Google Scholar*.
- Arafah, K., Amin, B. D., Sari, S. S., & Hakim, A. (2021). The Development of Higher Order-Thinking Skills (HOTS) Instrument Assessment in Physics Study. *Journal of Physics: Conference Series*, 1899(1), 012140. <https://doi.org/10.1088/1742-6596/1899/1/012140>
- Arip, M. A. S. M., Derum, A., Ahmad, A., & Jais, S. M. (2018). Development and testing of high

order thinking skills (HOTS) training module for sciences subjects among secondary school students in Malaysia. *Indian Journal of Public Health Research and Development*, 9(5). <https://doi.org/10.5958/0976-5506.2018.00409.6>

- Bahtiar, B., Sukmawati, A., & Setiawan, T. (2020). Developing assessment of higher order thinking skills in physics learning based on local wisdom. *Journal of Physics: Conference Series*, 1567(4). <https://doi.org/10.1088/1742-6596/1567/4/042045>
- Ruzbarsky, J. J., Marom, N., & Marx, R. G. (2018). Measuring Quality and Outcomes in Sports Medicine. In *Clinics in Sports Medicine* (Vol. 37, Issue 3). <https://doi.org/10.1016/j.csm.2018.03.001>
- Septiani, V., & Paidi. (2021). Development of A Test Instrument to Measure Hots and Cognitive Knowledge in Biology. *Journal of Physics: Conference Series*, 1788(1), 012028. <https://doi.org/10.1088/1742-6596/1788/1/012028>
- Suprpto, E., Saryanto, S., Sumiharsono, R., & Ramadhan, S. (2020). The Analysis of Instrument Quality to Measure the Students' Higher Order Thinking Skill in Physics Learning. *Journal of Turkish Science Education*, 17(4). <https://doi.org/10.36681/tused.2020.42>

AKSIOMA: Jurnal Program Studi Pendidikan Matematika
Volume 0, No. 0, 20xx, 00-00

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: <https://doi.org/10.24127/ajpm>

Artikel Novia Dwi - 4

ORIGINALITY REPORT

21 %

SIMILARITY INDEX

19 %

INTERNET SOURCES

15 %

PUBLICATIONS

16 %

STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

17%

★ repository.radenintan.ac.id

Internet Source

Exclude quotes On

Exclude bibliography On

Exclude matches < 1%