

## KUALITAS SOAL HOTS (*HIGH ORDER THINKING SKILL*) PADA PESERTA DIDIK SMP KELAS VII

Koryna Aviory<sup>1\*</sup>, M.M Endang Susetyawati<sup>2</sup>

<sup>1\*,2</sup>Pendidikan Matematika, Universitas PGRI Yogyakarta, Yogyakarta, Indonesia

\*Corresponding author. Jalan PGRI I Sonosewu 117, 55182, Yogyakarta

Email: [koryna@upy.ac.id](mailto:koryna@upy.ac.id)<sup>1\*)</sup>  
[magda.fausta@yahoo.com](mailto:magda.fausta@yahoo.com)<sup>2)</sup>

Received 07 September 2020; Received in revised form 28 February 2021; Accepted 28 June 2021

### ABSTRAK

Tujuan penelitian ini adalah menganalisis kualitas butir soal HOTS pada peserta didik SMP kelas VII. Jenis penelitian yang digunakan adalah penelitian pengembangan. Penelitian pengembangan terdiri dari tahap pendefinisian, perancangan, pengembangan, dan diseminasi. Penelitian ini hanya sampai tahap pengembangan, lebih tepatnya sampai analisis kualitas butir soal HOTS. Analisis soal bertujuan untuk menghasilkan soal valid dan reliabel, serta dapat mengidentifikasi soal. Penilaian hanya terbatas pada aspek keterampilan, sehingga pembuatan instrumen meliputi beberapa kompetensi dasar yang ada pada kelas VII. Instrumen yang telah disusun dilakukan uji validitas. Uji validitas tersebut meliputi validitas isi, validitas konstruk dan validitas empiris. Validitas empiris melalui ujicoba skala terbatas dengan jumlah sampel 80 peserta didik. Hasil penelitian menunjukkan bahwa dari 12 butir soal yang dibuat hanya 5 soal yang dinyatakan valid memenuhi kriteria HOTS, selanjutnya setelah diuji secara empiris hanya ada 4 butir soal yang bisa digunakan, dengan 2 butir soal diantaranya memerlukan perbaikan. Nilai reliabilitasnya berada pada level tinggi, yaitu 0,74.

**Kata kunci:** HOTS (*high order thinking skill*); soal.

### ABSTRACT

The purpose of this study was to analyze the quality of HOTS items in class VII junior high school students. This type of research used is development research. Development research consists of the stages of defining, designing, developing, and disseminating. This research only reached the development stage, more precisely to the quality analysis of HOTS items. Problem analysis aims to produce valid and reliable questions, in addition to identifying problems. Assessment is only limited to the aspect of skills, so the making of the instrument includes some basic competencies that exist in class VII. The instrument that was compiled was tested for validity. The validity test includes content validity, construct validity and empirical validity. Empirical validity through limited scale trials with a sample of 80 students. The results showed that out of the 12 questions made only 5 questions that were declared valid HOTS's criteria, then after empirical testing there were only 4 questions that could be used, with 2 of them requiring improvement. The reliability value is at a high level, which is 0.74.

**Keywords:** HOTS (*high order thinking skill*); item.



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

### PENDAHULUAN

Pembelajaran matematika merupakan proses membangun konsep matematika secara terstruktur melalui pengalaman proses pembelajaran (Tompong & Jailani, 2019). Menurut Heris (2014), matematika sekolah dilihat dari 4 aspek, yaitu penyajian,

pola pikir, semesta pembicaraan, dan tingkat keabstrakan. Semakin meningkat kemampuan peserta didik maka semesta matematikanya juga semakin meningkat. Demikian juga dengan tingkat keabstrakan yang harus sejalan dengan tingkat perkembangan intelektual peserta didik. Inilah

DOI: <https://doi.org/10.24127/ajpm.v10i2.3087>

pentingnya peserta didik selalu dihadapkan dengan soal-soal matematika sekolah yang membutuhkan pemikiran tingkat tinggi.

Berdasarkan hasil PISA (*Programme for International Student Assessment*) tahun 2018, Indonesia memiliki rata-rata skor 379 dalam kategori matematika. Rerata ini masih jauh di bawah rerata skor matematika OECD 489. Kurang lebih hanya sekitar 28% peserta didik dapat menafsirkan situasi kemudian merubahnya dalam model matematika (OECD, 2019). Hal ini dapat diartikan bahwa peserta didik di Indonesia masih lemah dalam memahami makna kalimat terutama dalam konteks matematika. Perkembangan teknologi membuat guru menjadi khawatir terhadap hal tersebut.

Kemampuan pemahaman seseorang dalam berpikir matematika akan berdampak pada kreativitas seseorang. Sejalan dengan hasil penelitian Tambunan & Surya (2019), peserta didik yang memiliki kemampuan berpikir kreatif yang optimal, maka mereka akan mampu memecahkan soal matematika secara cepat, tepat, dan benar.

Berdasarkan hasil penilaian yang dilakukan OECD, maka peserta didik di Indonesia harus lebih terbiasa dengan soal-soal yang menuntut mereka untuk menunjukkan ide kreatifnya. Ide kreatif hanya dapat dihasilkan melalui pemikiran tingkat tinggi, bukan pemikiran tingkat rendah melalui penerapan pengetahuan yang dipelajari dalam kehidupan sehari-hari (Mei et al., 2012). Hasil ini diperkuat dengan adanya penelitian menurut Kurniati, hampir 40% dari jumlah sampel tidak mampu memenuhi kriteria penilaian HOTS (Kurniati et al., 2016)

Hasil wawancara dengan guru di SMPN 3 Gamping menunjukkan bahwa

peserta didik di sekolah tersebut belum terbiasa mengerjakan soal-soal dengan tingkat berpikir yang lebih tinggi, sehingga mereka merasa kesulitan untuk menghasilkan ide-ide kreatif. Sebenarnya mereka tidak menyadari bahwa setiap individu mempunyai kapasitas untuk menghasilkan ide. Ide harus dihasilkan melalui stimulasi.

HOTS didasarkan pada kemampuan berpikir tingkat rendah (*Lower Order Thinking Skill*) seperti: diskriminasi, aplikasi dan analisis sederhana, serta strategi kognitif. Hasil kemampuan berpikir HOTS jauh lebih kompleks termasuk penjelasan, keputusan, kinerja, dan pengambilan keputusan dalam kehidupan nyata sehingga peserta didik mampu mengidentifikasi kemampuan yang mereka miliki (Thnull, 2017). Kemampuan berpikir tingkat tinggi merupakan kemampuan berpikir kritis, kreatif dan analitis (Sa'idah et al., 2018). Hasil penelitian menunjukkan bahwa peserta didik yang diuji sepanjang semester dengan pertanyaan berpikir tingkat tinggi memperoleh pemahaman konseptual yang mendalam tentang materi yang diajarkan (Jensen et al., 2014). Dengan demikian, peserta didik harus terbiasa terhadap tantangan intelektual yang mencakup konsepsi pemikiran pada gagasan non-rutin. Mereka dituntut untuk menyelesaikan masalah yang memerlukan interpretasi, analisis, dan aplikasi pengetahuan sebelumnya.

Skor pemikiran kreatif mulai menurun dan bersifat tetap pada peserta didik kelas VI (Kim, 2011). Sehingga peneliti mengambil subyek penelitian kelas VII pada jenjang SMP untuk memperdalam penelitian yang dilakukan oleh Kim.

Proses pembelajaran juga akan berkesinambungan dengan evaluasi.

DOI: <https://doi.org/10.24127/ajpm.v10i2.3087>

Evaluasi merupakan proses kegiatan yang sistematis, berkelanjutan, dan menyeluruh dalam pengendalian, penetapan kualitas terhadap komponen pembelajaran berdasarkan kriteria tertentu dan tujuan yang telah ditetapkan oleh guru. Hasil wawancara dengan beberapa guru menunjukkan bahwa sebagian besar guru termasuk guru matematika masih mengalami kesulitan dalam menyusun instrumen penilaian yang berkualitas. Guru kadang tidak memperhatikan analisis butir soal yang dibuatnya. Mereka membuat soal bukan berdasarkan analisis yang dilakukan tetapi berdasarkan keahlian di bidangnya (Susanto H, Rinaldi, 2015). Dengan demikian, guru seharusnya mampu membuat instrumen tes yang berkualitas. Instrumen yang berkualitas merupakan instrumen yang mampu mengukur apa yang hendak diukur. Instrumen tersebut harus mampu mengukur pemahaman dan juga pengetahuan peserta didik sehingga mereka dapat meningkatkan kemampuan berpikirnya. Nilai alpha dapat dijadikan untuk menentukan instrumen yang berkualitas (Taber, 2018). Instrumen tersebut akan mengukur kemampuan peserta didik dalam menguasai suatu materi pembelajaran (Amelia, 2016).

Berdasarkan penjelasan di atas, peneliti ingin mengembangkan soal HOTS untuk peserta didik SMP kelas VII. Tujuan pengembangan adalah untuk mengidentifikasi soal serta mengetahui kualitas soal HOTS. Kualitas butir soal akan menjadi pertimbangan dalam mengembangkan instrumen penelitian pada penelitian selanjutnya. Soal yang dihasilkan diharapkan dapat digunakan oleh guru sebagai acuan dan contoh dalam membuat instrumen tes yang berkualitas, sehingga guru-guru akan

terbiasa untuk membuat instrumen soal HOTS sehingga peserta didik akan terbiasa pula mengerjakan soal HOTS.

Kualitas butir soal dapat dilihat dari reliabilitas dan validitas butir soal. Reliabilitas dan validitas skor secara substansial dipengaruhi oleh korelasi. Nilai reliabilitas menunjukkan konsistensi internal yang baik dan kestabilan sebuah fungsi dari waktu ke waktu (Mathias B, 2013).

## METODE PENELITIAN

Penelitian ini dilaksanakan di SMP Negeri 3 Gamping. Subyek uji coba terdiri dari 80 peserta didik. Instrumen yang dikembangkan adalah tes untuk mengukur kemampuan berpikir tingkat tinggi (*Higher Order Thinking*) atau biasa disebut HOTS.

Indikator HOTS dalam penelitian ini meliputi: (1) menganalisis (C4), yaitu kemampuan menghubungkan konsep satu dengan yang lain; (2) mengevaluasi (C5), yaitu kemampuan menetapkan sesuatu berdasarkan patokan tertentu; (3) mencipta (C6), yaitu kemampuan memadukan unsur menjadi sesuatu yang baru (Dinni, 2018).

Penelitian ini merupakan penelitian pengembangan (*research and development*) terdiri atas empat tahap yaitu: tahap pendahuluan, tahap perancangan, tahap pengembangan, dan tahap penyebaran. Penelitian ini hanya sampai tahap pengembangan, lebih tepatnya sampai tahap analisis kualitas butir soal HOTS.

Tahap pendahuluan atau disebut tahap pendefinisian meliputi analisis materi dan perumusan indikator pencapaian kompetensi. Tahap perancangan berarti tahap penyusunan instrumen penelitian. Tahap terakhir adalah pengembangan. Pada tahap ini dilakukan analisis butir soal. Tujuannya

DOI: <https://doi.org/10.24127/ajpm.v10i2.3087>

menghasilkan soal yang valid dan reliabel, serta dapat mengidentifikasi soal berdasarkan daya pembeda dan tingkat kesukaran.

Uji validitas dilakukan dalam tiga tahapan, yaitu: (1) validitas isi, yaitu pengujian validitas menggunakan kisi-kisi dan format penulisan soal; (2) validitas kostruk, yaitu validitas untuk mengukur indikator HOTS; (3) validitas kriteria, yaitu validitas butir soal. Validitas butir soal dilakukan dengan korelasi *product moment Pearson*.

Soal tes dapat dipercaya apabila memberikan hasil pengukuran yang relatif konsisten, hal inilah yang dinamakan reliabilitas. Reliabilitas merupakan akurasi dan presisi yang dihasilkan oleh alat ukur dalam melakukan pengukuran (Waminton, 2015).

Perhitungan reliabilitas menggunakan formula Cronbach's Alpha. Hal ini dikarenakan jumlah butir soal kurang dari 30, sehingga butir soal tidak dibagi kedalam kelompok-kelompok. Soal dikatakan valid ketika koefisien korelasi berada pada interval  $0,4 \leq r_{xy} \leq 1$ . Soal yang valid berarti soal tersebut dapat diterima.

Tingkat kesukaran dapat dihitung dengan membandingkan jumlah nilai peserta tes yang menjawab benar dan jumlah peserta tes secara keseluruhan (Zainal, 2019).

## HASIL DAN PEMBAHASAN

Tahap pertama, yaitu tahap pendefinisian (*define*) meliputi analisis awal-akhir, analisis materi, analisis tugas peserta didik, dan perumusan indikator pencapaian kompetensi. Analisis awal dilakukan dengan cara observasi. Hasil observasi menunjukkan bahwa guru belum terbiasa membuat soal ber kriteria HOTS. Kemudian untuk analisis akhir, dibuatlah kisi-kisi soal

dengan kriteria HOTS. Dalam pembuatan kisi-kisi pemilihan materi juga menjadi bahan pertimbangan. Analisis tugas peserta didik juga menjadi dasar pertimbangan dalam mengembangkan soal. Pertimbangan ini dimaksudkan untuk mengatasi permasalahan terkait penguasaan kompetensi peserta didik kelas VII.

Tahap kedua, perancangan, meliputi pembuatan kisi-kisi soal HOTS, pemilihan format, dan penyusunan indikator soal. Format soal berupa soal uraian. Soal dibuat sebanyak 12 butir yang meliputi beberapa kompetensi dasar kelas VII. Penilaian diutamakan pada aspek keterampilan. Soal dibuat berdasarkan kriteria HOTS, yaitu kemampuan berpikir tingkat tinggi dengan 3 indikator (menganalisis, mengevaluasi, dan mencipta).

Tahap ketiga, pengembangan, soal HOTS dikembangkan berdasarkan kisi-kisi yang telah dibuat pada tahap perancangan. Instrumen yang dikembangkan haruslah berkualitas. Instrumen yang berkualitas ditinjau dari beberapa aspek, baik dari segi isi, kontruksi, kevalidan, daya pembeda maupun tingkat kesukaran.

Instrumen yang sudah dibuat kemudian divalidasi oleh ahli. Validasi dilakukan dengan tiga tahapan melalui validasi isi, konstruk, dan kriteria. Pertama, validasi isi, isi tes sebaiknya sesuai dengan materi yang hendak diujikan. Kedua, validasi konstruk, tes harus memiliki konstruk yang baik. Dalam hal ini kontruksi terkait kompetensi terhadap HOTS. Ketiga, validitas kriteria. Instrumen tes yang dihasilkan berfungsi sebagai alat diagnosis kemampuan berpikir tingkat tinggi, maka validasi dilakukan secara konkuren dan akan diperoleh nilai reliabilitas.

DOI: <https://doi.org/10.24127/ajpm.v10i2.3087>

Soal yang telah dibuat kemudian divalidasi oleh 3 validator (*expert judgement*). Validitas isi yang dilakukan dalam penelitian ini bertujuan untuk melihat apakah soal yang dibuat sudah sesuai dengan kisi-kisi soal dan format penulisan soal. Dari segi format dan bahasa yang digunakan, hasil validasinya sebesar 0,905, artinya format dan bahasa yang digunakan sangat baik. Sedangkan untuk isi, hasil validasinya sebesar 0,924. Hal ini menunjukkan bahwa kedua belas soal sudah sesuai dengan kisi-kisi soal, walaupun ada beberapa kalimat yang harus diperbaiki karena kurang sesuai dengan kaidah bahasa Indonesia.

Setelah validitas isi, kemudian dilanjutkan validitas konstruk. Validitas konstruk bertujuan untuk mengetahui apakah soal yang dibuat sesuai dengan indikator HOTS, yaitu menganalisis, mengevaluasi, dan mencipta. Dari 12 soal yang dibuat, hanya 5 soal yang dikatakan valid secara konstruksi. Soal divalidasi oleh 2 orang validator, dimana validator merupakan seseorang

yang memiliki kompetensi terhadap HOTS. Setelah didapatkan soal yang valid secara isi maupun konstruk kemudian dilanjutkan validitas kriteria. Validitas ini bertujuan untuk melihat soal berdasarkan kriterianya. Soal HOTS yang dibuat apakah sudah mengukur kemampuan berpikir tingkat tinggi peserta didik kelas VII. Soal diujicobakan secara terbatas dengan jumlah sampel 80 peserta didik.

Hasil validasi instrumen untuk soal nomor 1, 2, 3, 4, dan 5 berturut-turut 0,779; 0,430; 0,825; 0,644; 0,709. Kelima soal berada pada koefisien korelasi yang valid. Soal yang valid artinya soal tersebut menghasilkan skor yang valid untuk mengukur kemampuan berpikir tingkat tinggi peserta didik kelas VII.

Selain dilihat dari validitasnya, sebagai bahan pertimbangan, apakah kelima soal bisa digunakan semua atau tidak, yaitu dengan menghitung tingkat kesukaran dan daya pembeda. Hasil perhitungan tingkat kesukaran dan daya pembeda dapat dilihat pada Tabel 1.

Tabel 1. Tingkat kesukaran dan daya pembeda

No. Soal	Tingkat Kesukaran	Tafsiran	Daya Beda	Tafsiran	Status Soal
1	0,39	Soal Sedang	0,56	Baik	Soal Diterima
2	0,53	Soal Sedang	0,17	Jelek	Soal Dibuang
3	0,39	Soal Sedang	0,60	Baik	Soal Diterima
4	0,18	Soal Sulit	0,26	Cukup	Soal Diterima tapi Diperbaiki
5	0,14	Soal Sulit	0,31	Cukup	Soal Diterima tapi Diperbaiki

Hasil penelitian menunjukkan, ada 3 soal tingkat kesukaran pada tafsiran sedang, sedangkan 2 soal lainnya pada tingkat kesukaran tafsiran sulit.

Butir soal yang baik adalah butir soal yang berada pada rentang 0,4 sampai 0,7. Daya pembeda dikatakan

jelek ketika daya pembeda berada pada rentang 0 – 0,2, daya pembeda cukup berada pada rentang 0,21 – 0,40, daya pembeda baik berada pada rentang 0,41 – 0,7, sedangkan daya pembeda baik sekali berada pada rentang 0,71 – 1.

Untuk daya pembeda, ada 2 soal yang memiliki daya pembeda baik, ada

DOI: <https://doi.org/10.24127/ajpm.v10i2.3087>

2 soal yang memiliki daya pembeda cukup, dan hanya ada satu soal yang memiliki daya pembeda yang jelek. Daya pembeda yang baik berarti peserta didik yang sudah menguasai materi mampu mengerjakan soal dengan baik dibandingkan peserta didik yang belum menguasai materi. Berbeda halnya dengan daya pembeda yang jelek, peserta didik yang menguasai materi atau yang belum menguasai materi sama-sama tidak bisa mengerjakan soal dengan benar.

Setelah soal HOTS dianalisis melalui perhitungan validitas kriteria, daya pembeda dan tingkat kesukaran, diperoleh 4 soal valid yang mampu mengukur kemampuan berpikir tingkat tinggi peserta didik kelas VII. Soal tersebut adalah soal nomor 1, 3, 4, dan 5. Soal HOTS akan dipakai tanpa revisi jika tingkat kesukaran soal berada pada kriteria minimal sedang, sedangkan daya pembeda soal berada pada kriteria baik. Soal HOTS akan direvisi jika tingkat kesukaran soal berada pada kriteria minimal sedang tetapi daya pembeda berkriteria cukup.

Soal nomor 2 menghasilkan tingkat kesukaran sesuai dengan kriteria, yaitu 0,53 dalam kategori sedang, tetapi daya pembeda dalam kategori jelek. Daya pembeda ini di bawah kriteria soal yang direvisi, maka soal tersebut harus dibuang.

Soal nomor 1 dan 3 dapat digunakan tanpa melalui perbaikan. Soal tersebut memenuhi kriteria soal HOTS yang baik, karena memiliki tingkat kesukaran (0,39) sedang dan daya pembeda (0,56) yang baik. Sedangkan soal nomor 4 dan 5 bisa digunakan tetapi melalui perbaikan. Tingkat kesukaran kedua soal sudah memenuhi kriteria minimal sedang tetapi daya pembeda masih berada pada kriteria cukup, yaitu 0,26 dan 0,31.

Sehingga, kedua soal tersebut memenuhi kriteria soal HOTS dengan perbaikan.

Perbaikan soal dapat dilakukan dengan merevisi bahasa atau menambah ilustrasi pada soal agar soal mudah dipahami oleh peserta didik. Dalam penelitian ini, perbaikan soal pada nomor 4 dan 5 dilakukan dengan tidak mengubah isi soal HOTS tetapi hanya menambahkan ilustrasi gambar dan merubah bahasa soal. Perbaikan ini berdasarkan pertimbangan hasil pekerjaan peserta didik dan hasil wawancara dari beberapa peserta didik. Lebih dari 55% peserta didik kurang mampu untuk merubah soal ke dalam kalimat matematika. Mereka kurang mampu memahami maksud dari soal, kurang mampu untuk mengemukakan ide kreasinya dalam menyelesaikan masalah matematika.

Soal tidak dilengkapi dengan ilustrasi gambar sehingga peserta didik tidak tertarik untuk memahami soal matematika yang disajikan. Perubahan soal dapat dilihat pada Tabel 2.

Tabel 2. Soal revisi

---

#### Soal nomor 4 sebelum revisi

---

Seekor ayam melihat elang sedang terbang. Jika sudut pandang ayam terhadap elang tersebut membentuk sudut  $30^\circ$  serta jarak antara elang dan ayam adalah 36m, elang akan menggapai ayam dengan kecepatan terbang 10 m/s. Maka berapakah jarak elang dari permukaan tanah dan waktu yang diperlukan Elang untuk menggapai ayam!

---

#### Soal nomor 4 setelah revisi

---

Perhatikan gambar berikut!



Seekor Ayam melihat Elang sedang terbang. Jika sudut pandang ayam terhadap elang tersebut membentuk sudut  $30^\circ$  serta jarak antara elang dan ayam adalah 36m, jika elang akan menggapai ayam dengan kecepatan terbang 10 m/s. Maka berapakah jarak elang dari permukaan tanah dan waktu yang diperlukan Elang untuk menggapai ayam!

---

DOI: <https://doi.org/10.24127/ajpm.v10i2.3087>

### Soal nomor 5 sebelum revisi

Daffa pergi ke sebuah toko masker untuk membeli beberapa masker. Di toko tersebut terdapat berbagai jenis masker yang dijual dengan harga per box. Harga setiap masker berbeda-beda, untuk maser JPS Mask Rp. 24.000, Alkindo Mask Rp. 26.000, Disposable Mask Rp. 20.000, Ramada Mask Rp. 22.000 dan Orlee Mask Rp. 25.000. Kemudian pemilik toko memberikan promosi dengan minimal membeli 3 box masing-masing jenis masker, dengan diskon 10%. Jika Daffa membeli sebanyak 2 masker jenis Orlee Mask dan 3 masker jenis Alkindo Mask serta 1 masker jenis Ramada Mask. Jika Daffa membayar dengan uang limaribuan 2 lembar dan selembaer uang duaribuan maka sisa uang Daffa adalah...

### Soal nomor 5 setelah revisi

Daffa pergi ke sebuah toko masker untuk membeli beberapa masker. Di toko tersebut terdapat berbagai jenis masker yang dijual dengan harga per box sebagai berikut.



Harga setiap masker berbeda-beda untuk maser JPS Mask Rp. 24.000, Alkindo Mask Rp. 26.000, Disposable Mask Rp. 20.000, Ramada Mask Rp. 22.000 dan Orlee Mask Rp. 25.000. Kemudian pemilik toko memberikan promosi sebesar 10% dengan minimal membeli 3 box masing-masing jenis masker. Jika Daffa membeli sebanyak 2 masker jenis Orlee Mask dan 3 masker jenis Alkindo Mask serta 1 masker jenis Ramada Mask. Jika Daffa membayar dengan uang limaribuan 2 lembar dan selembaer uang duaribuan maka sisa uang Daffa adalah...

Perubahan bahasa dan ilustrasi pada soal nomor 4 dan 5 diharapkan membuat peserta didik menjadi lebih mengerti dan mampu menerjemahkan soal ke dalam konteks matematika. Soal yang sudah direvisi dapat digunakan dalam pengukuran skala luas. Jadi, soal yang dapat digunakan sebanyak 4 soal.

Keempat soal HOTS yang dikembangkan ditinjau dari daya pembeda (baik dan cukup) serta tingkat kesukaran (sulit dan sedang), masing-masing memiliki proporsi yang sama, yaitu 50%. Soal tersebut diharapkan mampu menjadi instrumen penilaian peserta didik jenjang SMP kelas VII pada skala yang lebih luas. Keterbatasan waktu menjadi kendala peneliti untuk melakukan penelitian lebih lanjut. Sehingga penelitian ini hanya terbatas pada uji coba skala terbatas. Walaupun demikian, proporsi soal ditinjau dari daya pembeda, tingkat kesukaran dapat menjadi pertimbangan dalam membuat soal. Dalam membuat soal, sebaiknya ada soal yang mudah, sukar, dan sedang. Tujuannya untuk membedakan antara peserta tes yang mampu dan tidak mampu mengerjakan soal.

Sedangkan pada daya pembeda, jika daya pembeda jelek maka soal tersebut tidak dapat mengukur sesuatu atau soal tersebut tidak bisa membedakan antara peserta didik yang berkemampuan tinggi dan rendah. Sehingga daya pembeda haruslah minimal pada kategori cukup. Selain itu, jumlah soal yang dibuat sebaiknya lebih dari 15 soal, untuk mengantisipasi hasil analisis butir soal.

Kualitas soal dapat dilihat pada nilai korelasinya. Terdapat hubungan antara reliabilitas dan koefisien korelasi. Reliabilitas sebanding dengan nilai korelasi (Wu et al., 2016). Koefisien korelasi dikatakan baik jika nilainya minimal 0,7 (Kartowagiran, 2012). Dengan menggunakan rumus Cronbach, koefisien reliabilitas dapat ditentukan. Koefisien reliabilitas yang dihasilkan sebesar 0,74. Hal ini berarti kualitas soal yang dihasilkan memiliki kategori baik. Koefisien ini merupakan indikator konsistensi butir soal dalam menjalankan fungsi ukurnya (mengukur kemampuan berpikir tingkat tinggi peserta didik kelas VII) secara bersama-sama. Hasil perhitungan koefisien reliabilitas berada pada level tinggi. Hal ini berarti soal HOTS ajeg, artinya mampu mengukur kemampuan berpikir tingkat tinggi peserta didik kelas VII dalam keadaan bagaimanapun. Selain itu, penilaian HOTS dalam pembelajaran efektif untuk mengukur keterampilan berpikir seseorang (Kusuma et al., 2017).

Soal yang dihasilkan dalam penelitian ini dapat dijadikan contoh dan acuan guru dalam membuat instrumen tes. Selain itu, dapat dijadikan kumpulan soal yang siap digunakan untuk mengevaluasi peserta didik kelas VII yang diharapkan akan terbiasa mengerjakan soal HOTS, sehingga kemampuan berpikirnya meningkat.

DOI: <https://doi.org/10.24127/ajpm.v10i2.3087>

## KESIMPULAN DAN SARAN

Berdasarkan hasil dan pembahasan dapat disimpulkan, dari 12 butir soal yang dibuat hanya 4 butir soal yang bisa digunakan, dengan 2 butir soal diantaranya memerlukan perbaikan. Nilai reliabilitasnya berada pada level tinggi, yaitu 0,74. Sedangkan untuk kualitas soalnya berada pada kategori baik. Sehingga keempat soal tersebut dapat digunakan untuk mengevaluasi peserta didik kelas VII.

Penelaahan soal HOTS perlu dilakukan untuk penelitian selanjutnya, agar pemahaman soal HOTS menjadi semakin mudah. Sebaiknya jumlah soal yang dikembangkan lebih dari 15. Hal ini untuk mengantisipasi hasil analisis jika terdapat butir soal yang tidak bisa dipakai.

## DAFTAR PUSTAKA

Amelia, M. A. (2016). Analisis Soal Tes Hasil Belajar High Order Thinking Skills (HOTS) Matematika Materi Pecahan Untuk Kelas 5 Sekolah Dasar. *Jurnal Penelitian*, 20(2), 123–131.

Dinni, H. N. (2018). HOTS ( High Order Thinking Skills ) dan Kaitannya dengan Kemampuan Literasi Matematika. *Prisma*, 1, 170–176.

Jensen, J. L., McDaniel, M. A., Woodard, S. M., & Kummer, T. A. (2014). Teaching to the Test...or Testing to Teach: Exams Requiring Higher Order Thinking Skills Encourage Greater Conceptual Understanding. *Educational Psychology Review*, 26(2), 307–329. <https://doi.org/10.1007/s10648-013-9248-9>

Kartowagiran, B. (2012). Penulisan Butir Soal. *Jurnal UNIVERSITAS NEGERI YOGYAKARTA*, 1–33.

<http://staffnew.uny.ac.id/upload/130693812/penelitian/penulisan-dan-analisis-butir-soal.pdf>

- Kim, K. H. (2011). The Creativity Crisis: The Decrease in Creative Thinking Scores on the Torrance Tests of Creative Thinking. *Creativity Research Journal*, 23(4), 285–295. <https://doi.org/10.1080/10400419.2011.627805>
- Kurniati, D., Harimukti, R., & Jamil, N. A. (2016). Kemampuan berpikir tingkat tinggi peserta didik SMP di Kabupaten Jember dalam menyelesaikan soal berstandar PISA. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 20(2), 142–155. <https://doi.org/10.21831/pep.v20i2.8058>
- Kusuma, M. D., Rosidin, U., Abdurrahman, A., & Suyatna, A. (2017). The Development of Higher Order Thinking Skill (Hots) Instrument Assessment In Physics Study. *IOSR Journal of Research & Method in Education (IOSRJRME)*, 07(01), 26–32. <https://doi.org/10.9790/7388-0701052632>
- Mathias B, C. et all. (2013). Assessment of Divergent Thinking by Means of the Subjective Top-Scoring Method: Effects of the Number of Top-Ideas and Time-on-Task on ReHabiHty and Vahdity Mathias. *Psychology of Aesthetics, Creativity, and the Arts*, 7(4), 341–349. <https://doi.org/10.1037/a0033544>
- Mei, Y., Yunos, J., Othman, W., Hassan, R., & Tze, T. (2012). *The needs analysis of learning higher order thinking skills for generating ideas*. 59, 197–203. <https://doi.org/10.1016/j.sbspro.2012.09.265>



DOI: <https://doi.org/10.24127/ajpm.v10i2.3087>

- OECD. (2019). Programme for international student assessment (PISA) results from PISA 2018. *Oecd*, 1–10. [https://www.oecd-ilibrary.org/education/pisa-2018-results-volume-iii\\_bd69f805-en%0Ahttps://www.oecd-ilibrary.org/sites/bd69f805-en/index.html?itemId=/content/component/bd69f805-en#fig86](https://www.oecd-ilibrary.org/education/pisa-2018-results-volume-iii_bd69f805-en%0Ahttps://www.oecd-ilibrary.org/sites/bd69f805-en/index.html?itemId=/content/component/bd69f805-en#fig86)
- Sa'idah, N., Yulistianti, H. D., & Megawati, E. (2018). Analisis Instrumen Tes Higher Order Thinking Matematika Smp. *Jurnal Pendidikan Matematika*, 13(1), 41–54. <https://doi.org/10.22342/jpm.13.1.6619.41-54>
- Susanto H, Rinaldi, N. (2015). Analisis Validitas Reliabilitas Tingkat Kesukaran dan Daya Pembeda Pada Butir Soal Ujian Akhir Semester Ganjil Mata Pelajaran Matematika. *The Journal of the Japan Society for Respiratory Endoscopy*, 37(3), 343. [https://doi.org/10.18907/jjsre.37.3\\_343\\_4](https://doi.org/10.18907/jjsre.37.3_343_4)
- Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, 48(6), 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Tambunan, E. T., & Surya, E. (2019). Pengaruh Kemampuan Berpikir Kreatif (Creative Thinking) Peserta didik dalam Menyelesaikan Masalah Matematika. *Jurnal Universitas Negeri Medan, December*, 1–8.
- Thnull, N. (2017). *Influence of Explicit Higher-Order Thinking Skills Instruction on Students' Learning of Linguistics*. <https://doi.org/10.1016/j.tsc.2017.10.004>
- Tompong, B. N. K. J., & Jailani, J. (2019). An evaluation of mathematics learning program at primary education using Countenance Stake Evaluation model. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 23(2), 156–169. <https://doi.org/10.21831/pep.v23i2.16473>
- Wu, M., Tam, H. P., & Jen, T.-H. (2016). Educational Measurement for Applied Researchers. In *Educational Measurement for Applied Researchers*. <https://doi.org/10.1007/978-981-10-3302-5>