

PENGEMBANGAN SOAL MATEMATIKA BERBASIS *HIGHER ORDER THINKING SKILLS* PADA LINGKUP MATERI UJIAN NASIONAL UNTUK TINGKATAN SEKOLAH MENENGAH PERTAMA

DEVELOPMENT OF HIGHER ORDER THINKING SKILLS MATHEMATICS TEST FOR NATIONAL EXAMINATION AT THE JUNIOR HIGH SCHOOL LEVEL

Muhamad Syahidul Qirom*, Nyoman Sridana, Sudi Prayitno

Program Studi Pendidikan Matematika FKIP Universitas Mataram, Indonesia

*Email: syahidulqirom36@gmail.com

Diterima: 24 Agustus 2020. Disetujui: 26 November 2020. Dipublikasikan: 26 November 2020

Abstrak: Soal matematika berbasis *higher order thinking skills* dapat membantu meningkatkan keterampilan abad 21 siswa sehingga Kementerian Pendidikan dan Kebudayaan mencanangkan pelaksanaan program penilaian berbasis HOTS di setiap jenjang pendidikan. Faktanya adalah di SMPN 1 Mataram program tersebut belum berjalan sesuai mestinya. Terdapat dua masalah yang menyebabkan hal tersebut terjadi, yaitu kurangnya kemampuan guru dalam membuat soal HOTS matematika dan ketersediaan soal berbasis HOTS pada buku pelajaran matematika yang digunakan di sekolah tersebut masih belum cukup. Penelitian ini bertujuan untuk menghasilkan soal HOTS matematika untuk membantu memenuhi ketersediaan referensi soal HOTS baik bagi guru maupun siswa di SMPN 1 Mataram. Metode yang digunakan dalam penelitian ini adalah model pengembangan Tessmer yang terdiri atas dua tahap yakni *preliminary* dan *formative evaluation*. Hasil dari penelitian ini adalah diperolehnya 20 soal HOTS matematika dengan nilai validitas 75,9375%, kepraktisan 80,41%, keefektifan 79%, kesukaran soal dalam kategori sedang, daya pembeda dalam kategori cukup, soal reliabel, dan menunjukkan hasil sesuai harapan yakni keterampilan berpikir tingkat tinggi siswa berada pada kategori cukup. Kesimpulan dalam penelitian ini adalah telah dihasilkan soal HOTS matematika yang valid, praktis, dan efektif.

Kata Kunci : pengembangan soal, matematika, taksonomi bloom revisi, penalaran

Abstract: Mathematics HOTS question has trusted can improve student's skills in 21st century. Therefore, Indonesian ministry of education and culture launched assessment program based on HOTS in every grade of school. However, the data show that at SMPN 1 Mataram the program is not running as it should be. Two problems that cause this are the lack of the teacher's ability to make mathematics HOTS questions and the availability of HOTS-based questions in mathematics textbooks at the school is still insufficient. This study aims to produce mathematics HOTS questions to help meet the availability of HOTS questions for teachers and students at SMPN 1 Mataram. The Tessmer development model was used in this study, which consists of two stages, namely preliminary and formative evaluation. The results showed that 20 HOTS math questions were obtained with the validity, practicality and effectiveness values of 75.93%, 80.41% and 79%, respectively. The difficulty level of the questions is in the medium category, the distinguishing power is in the sufficient category, the questions are classified as reliable, and it shows that students' high-order thinking skills are in the sufficient category. In conclusion, a valid, practical, and effective HOTS math test for junior high school level has been produced.

Keywords : test development, mathematics, bloom revision taxonomy, reasoning

PENDAHULUAN

Mulainya era industri 4.0 serta berkembangnya dunia IPTEK, tentunya merupakan suatu tantangan bagi masyarakat abad 21 khususnya bangsa Indonesia. Dengan berbagai kemajuan yang ada, persaingan baik di kancah nasional maupun internasional akan semakin pesat. Sehingga setiap orang dituntut untuk menguasai keterampilan abad 21 dalam rangka mempersiapkan diri menghadapi persaingan yang terus meningkat. Keterampilan abad 21 sendiri menekankan pada empat hal utama atau lebih dikenal dengan 4C, yakni *critical thinking, creativity, communication, and collaboration* [2].

Dalam menghadapi tantangan tersebut salah satu upaya yang dilakukan pemerintah adalah melalui pengembangan dan inovasi kurikulum 2013. Inovasi

terus dilakukan agar relevan dengan perkembangan zaman [6]. Kurikulum 2013 dirancang sesuai dengan standar pendidikan internasional sehingga telah dilakukan berbagai penyempurnaan. Pada standar isi, dilakukan pengurangan materi yang tidak relevan kemudian memberikan pendalaman dan perluasan materi yang relevan serta diperkaya dengan kebutuhan peserta didik untuk berpikir kritis dan analitis. Pada standar penilaian, diadaptasi secara bertahap model-model penilaian standar internasional. Melalui penilaian tersebut diharapkan dapat membantu peserta didik meningkatkan keterampilan berpikir tingkat tinggi (*Higher Order Thinking Skills/HOTS*), karena hal tersebut dapat mendorong mereka untuk berpikir secara luas dan mendalam tentang suatu materi pelajaran [18].

Matematika sendiri, sudah memiliki peran dalam mengembangkan keterampilan abad 21. Selain itu menurut Hashimoto [5], pembelajaran matematika dengan penyajian masalah *open – ended* memiliki potensi untuk meningkatkan kreativitas peserta didik. Menurut Sumarna [14], pembelajaran matematika dengan pendekatan investigasi dapat membantu meningkatkan kemampuan berpikir kritis siswa. Kemudian menurut Madu [10] pembelajaran matematika berbasis HOTS dapat meningkatkan kemampuan pemecahan masalah, berpikir kritis, kemampuan bertanya, memberikan alasan, dan berkomunikasi.

Soal-soal berbasis HOTS sering sekali digunakan dalam ajang penilaian internasional terhadap kualitas pendidikan di suatu Negara seperti PISA. Berdasarkan hasil PISA 2018 Indonesia merupakan negara dengan skor PISA terendah [12]. Rendahnya nilai PISA Indonesia meunjukkan bahwa siswa belum dibiasakan mengerjakan soal-soal berbasis HOTS.

Tidak hanya hasil menurut penilaian dalam tingkat internasional, bahkan kemampuan peserta didik Indonesia melalui penilaian standar nasional juga masih tergolong rendah. Hal ini dibuktikan dengan hasil ujian nasional tahun 2016 – 2019, khususnya pada pelajaran matematika merupakan mata pelajaran yang menyumbang nilai ujian nasional terendah. Untuk tingkat SMP sederajat rata – rata nilai UN matematika secara nasional masih tergolong kurang karena berdasarkan standar yang ditetapkan pemerintah untuk setiap mata pelajaran memiliki standar ketercapaian minimal yakni rata – rata sama dengan 55 [4].

Tabel 1. Rata-rata UN Matematika Tingkat SMP Sederajat [8]

Tahun	Rata – rata
2016	49,84
2017	50,34
2018	44,05
2019	46,56

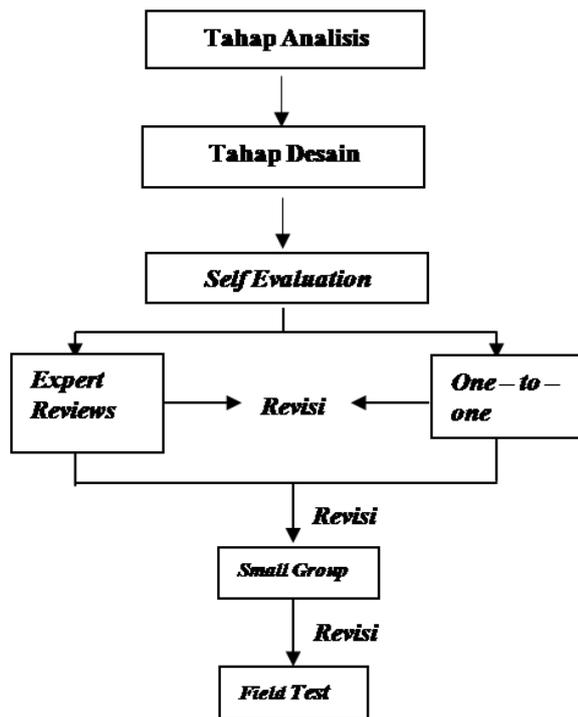
Dari Tabel 1 penurunan rata-rata nilai UN matematika siswa SMP/MTS di Indonesia selama empat tahun berturut-turut selalu berada pada kategori kurang selain itu pada tahun 2018 terjadi penurunan nilai rata-rata yang cukup tinggi [8]. Menurut Mendikbud dilansir dalam tempo.co.net penurunan hasil ujian nasional salah satunya dipicu oleh dimulainya penerapan soal berstandar HOTS pada ujian nasional pada tahun 2018 [16]. Padahal hal tersebut merupakan langkah awal pemerintah dalam membiasakan peserta didik menyelesaikan soal yang tidak hanya membutuhkan pemahaman saja, namun juga membantu membiasakan menyelesaikan soal yang membutuhkan penalaran. Sehingga untuk mengatasi permasalahan tersebut, hendaknya semua sekolah mulai membiasakan peserta didik menyelesaikan soal berstandar HOTS sehingga ketika

menghadapi ujian nasional mereka sudah memiliki persiapan yang matang.

Penilaian berbasis HOTS hingga saat ini pun tampaknya masih belum maksimal dilakukan di semua sekolah. Salah satu sekolah yang mengalami kendala dalam melakukan penilaian berbasis HOTS adalah SMPN 1 Mataram. Hal ini disebabkan karena beberapa faktor seperti kurangnya kemampuan guru membuat soal, siswa yang jarang dibiasakan soal HOTS, hingga pada ketersediaan sumber soal-soal berbasis HOTS pada buku pelajaran di sekolah tersebut. Oleh karena itu, dalam rangka membantu ketersediaan sumber soal HOTS baik bagi guru maupun siswa di SMPN 1 Mataram maka peneliti tertarik untuk melakukan penelitian tentang Pengembangan Soal *Higher Order Thinking Skills* Matematika Lingkup Materi Ujian Nasional SMP/MTs di SMPN 1 Mataram Tahun Ajaran 2019/2020. Tujuan dari penelitian ini adalah untuk menghasilkan satu set soal *higher order thinking skills* matematika pada lingkup materi ujian nasional tingkat SMP/MTs yang valid, praktis, dan efektif.

METODE PENELITIAN

Jenis penelitian yang digunakan adalah *Research and Development* (R&D) dengan model *Formative Evaluation* yang dikembangkan oleh Tessmer [17]. Dalam penelitian ini terdiri atas dua tahap utama, yakni tahap *preliminary* dan *Formative Evaluation*. Pada tahap *preliminary* terdiri atas tahap analisis dan tahap desain. Pada tahap *Formative Evaluation* terdiri atas *self evaluation*, *expert review*, *one – to – one*, *small group*, dan *field test*. Subjek pada penelitian ini adalah siswa kelas IX – A di SMPN 1 Mataram tahun ajaran 2019/2020 dengan objek penelitian 20 soal HOTS matematika berbentuk pilihan ganda pada lingkup materi UN tingkat SMP/MTs berdasarkan Taksonomi Bloom Revisi. Berikut rancangan beserta prosedur dalam penelitian ini:.



Gambar 2 Desain Formative Evaluation Tessmer [21]

1. Tahap Preliminary

a. Tahap Analisis

Pada tahap ini terdapat tiga hal yang dianalisis, yakni analisis siswa, materi, dan kurikulum. Analisis siswa bertujuan untuk mengetahui tingkat kemampuan matematis siswa. Analisis kurikulum bertujuan untuk mengetahui kurikulum yang digunakan di sekolah tempat penelitian dilaksanakan. Analisis materi bertujuan untuk mengetahui lingkup materi yang diujikan dalam ujian nasional.

b. Tahap Desain

Pada tahap ini dikembangkan soal HOTS berdasarkan Taksonomi Bloom Revisi sebanyak 20 buah. Soal HOTS tersebut disesuaikan dengan lingkup materi yang diujikan dalam ujian nasional matematika tingkat SMP sederajat yakni materi bilangan, aljabar, geometri dan pengukuran, serta statistika dan peluang.

2. Tahap Formative Evaluation

c. Self Evaluation

Pada tahap ini dilakukan evaluasi secara mandiri terhadap soal HOTS yang telah didesain, bertujuan untuk mengevaluasi kembali apakah soal yang telah dibuat sesuai dengan indikator HOTS. Soal yang telah didesain pada tahap ini disebut *prototype I*. Terdapat tiga karakteristik yang menjadi fokus prototipe yang akan dikembangkan yakni materi, konstruksi, dan bahasa.

d. Expert Review

Pada tahap ini dilakukan evaluasi atau penilaian oleh ahli mengenai soal HOTS yang telah dibuat sekaligus penilaian terhadap keefektifan soal. Hasil komentar dan saran dari ahli tersebut akan digunakan sebagai acuan dalam merevisi soal HOTS.

e. One – To – one

Pada tahap ini dilakukan parallel dengan *Expert Review*. *Prototype I* diuji coba kepada tiga orang siswa (bukan subjek penelitian) yang memiliki kemampuan berbeda, satu siswa dengan kemampuan tinggi, satu siswa dengan kemampuan sedang, dan satu siswa dengan kemampuan rendah. Respon siswa melalui angket, serta hasil komentar, dan saran dari ahli pada tahap *Expert Review* akan digunakan sebagai acuan dalam merevisi *prototype I*. Hasil revisi pada tahap tersebut disebut *prototype II*.

f. Small Group

Pada tahap ini dilakukan uji coba terhadap *prototype II* kepada enam orang siswa dengan kemampuan berbeda (bukan subjek penelitian), dua siswa dengan kemampuan tinggi, dua siswa dengan kemampuan sedang, dan dua siswa dengan kemampuan tinggi. Respon siswa melalui angket akan digunakan sebagai acuan dalam merevisi soal yang akan digunakan pada tahap field test. Hasil revisi pada tahap ini disebut *prototype III*.

g. Field Test

Pada tahap ini *prototype III* diujikan pada siswa kelas IX – A di SMPN 1 Mataram yang merupakan subjek dari penelitian. Pada tahap ini soal HOTS yang telah dikembangkan dihitung nilai reliabilitas, daya pembeda, dan indeks kesukarannya. Tujuan dari tahap ini adalah untuk mendapatkan produk akhir berupa soal matematika pada lingkup materi ujian nasional SMP yang sesuai kriteria HOTS menurut Taksonomi Bloom Revisi.

Instrumen Penelitian dan Teknik Pengumpulan Data

Instrumen penelitian yang digunakan pada penelitian ini adalah lembar angket dan soal tes. Sedangkan teknik pengumpulan data yang digunakan adalah teknik angket dan tes.

Teknik Analisis Data

a. Analisis Validitas

Dalam penelitian ini uji validitas yang digunakan adalah Validitas versi Aiken. Menurut Novaliah dan Kartowagiran [11], uji ini didasarkan pada hasil penilaian dari validator terhadap suatu item dari segi sejauh mana item tersebut mewakili sesuatu yang akan diukur. Untuk menghitung validitas berdasarkan Validitas Aiken digunakan rumus berikut:

$$V = \frac{\sum s}{[n(c-1)]} \dots(1)$$

Keterangan:

V = Indeks validitas Aiken
 $s = r - l_0$
 r = Angka yang diberikan validator
 l_0 = Angka Penilaian Validitas Terendah
 c = Angka Penilaian Validitas Tertinggi
 n = Banyak Validator

b. Analisis Angket Respon Siswa (Kepraktisan)

Untuk masing-masing aspek pernyataan ditentukan persentase jawaban siswa untuk masing-masing butir pernyataan dalam angket menggunakan rumus berikut [19]:

$$P = \frac{f}{n} \times 100\% \quad \dots(2)$$

Keterangan:

P = persentase jawaban
 f = Skor yang diperoleh
 n = Skor ideal (skor maksimum)
 Persentase yang diperoleh pada masing-masing butir pertanyaan, kemudian ditafsirkan berdasarkan kriterianya. Butir pernyataan diterima apabila memperoleh kriteria minimum baik.

c. Analisis Butir Soal Indeks Kesukaran

Untuk mengukur tingkat atau indeks kesukaran untuk masing – masing item pilihan ganda, digunakan rumus sebagai berikut [1]:

$$IK = \frac{B}{JS} \quad \dots(3)$$

Keterangan

IK= Indeks Kesukaran
 B = Banyak siswa yang menjawab benar
 JS = Jumlah seluruh siswa peserta tes

Daya Pembeda

Untuk masing – masing item soal pilihan ganda, digunakan rumus sebagai berikut [1]:

$$D = \frac{JAB}{BA} - \frac{JBB}{BB} \quad \dots(4)$$

Keterangan:

D= Daya pembeda
 BA = Banyak siswa pada kelompok atas (50% siswa dengan nilai tertinggi)
 BB = Banyak siswa pada kelompok bawah (50% siswa dengan nilai terendah)
 JAB = Banyak siswa pada kelompok atas yang menjawab benar
 JBB = Banyak siswa pada kelompok bawah yang menjawab benar

Penyebaran Pilihan Jawaban

Selanjutnya Penyebaran pilihan jawaban dihitung menggunakan rumus [20]:

$$PPJ = \frac{BJ}{N} \quad \dots(5)$$

Keterangan:

PPJ : Penyebaran pilihan jawaban
 BJ : Banyak siswa yang memilih jawaban tertentu
 N : Jumlah seluruh siswa

d. Reliabilitas Soal

Untuk mengukur reliabilitas soal digunakan rumus K-R. 20. Menurut Arikunto [1], rumus realibilitas dengan K-R 20 sebagai berikut:

$$r_{11} = \left(\frac{n}{n-1} \right) \left(\frac{S^2 - \sum pq}{S^2} \right) \quad \dots(6)$$

Keterangan:

r_{11} : reliabilitas soal secara keseluruhan
 p : proporsi subjek yang menjawab item dengan benar
 q : proporsi subjek yang menjawab item dengan salah ($q = 1 - p$)
 $\sum pq$: jumlah hasil perkalian antara p dan q
 n : banyaknya item
 S : standar deviasi

Setelah diperoleh indeks reliabilitas (r_{hitung}), selanjutnya mengkonsultasikan indeks tersebut dengan tabel r *product moment* dengan jumlah n yang sama pada taraf signifikansi 5%. Apabila $r_{hitung} \geq r_{tabel}$ maka instrumen dianggap reliabel. Sebaliknya apabila $r_{hitung} < r_{tabel}$ maka instrumen dianggap tidak reliabel [19].

e. Analisis Keefektifan

Soal HOTS dikatakan efektif dalam penelitian ini apabila:

- a) Menurut penilaian ahli soal tersebut efektif untuk mengukur keterampilan berpikir tingkat tinggi siswa.
- b) Menunjukkan hasil yang sesuai dengan yang diharapkan. Pada penelitian ini diharapkan soal dapat menunjukkan rata-rata hasil tes keterampilan berpikir tingkat tinggi siswa pada kriteria maksimal cukup.

Berikut teknik analisis data untuk menentukan keefektifan soal sesuai kriteria di atas:

- a) Keefektifan Menurut Pendapat Ahli
 Skor yang diberikan masing-masing ahli dijumlahkan kemudian dianalisis secara kuantitatif dengan menggunakan rumus [19]:

$$PK = \frac{f}{n} \times 100\% \quad \dots(7)$$

Keterangan:

- PK = Persentase Keefektifan
 f = Skor yang diperoleh
 n = Skor ideal (skor maksimum)
- b) Menunjukkan Hasil Sesuai dengan yang Diharapkan

Dalam penelitian ini keterampilan berpikir tingkat tinggi siswa akan dilihat dari hasil rata-rata skor siswa. Selanjutnya rata-rata skor yang diperoleh dikonsultasikan menurut tabel berikut:

Tabel 2 Keterampilan Berpikir Tingkat Tinggi Siswa [9]

Skor	Kategori
$15 \leq \text{skor} \leq 20$	Sangat Baik
$10 \leq \text{skor} < 15$	Baik
$5 \leq \text{skor} < 10$	Cukup
$0 \leq \text{skor} < 5$	Kurang

HASIL DAN PEMBAHASAN

Kevalidan Instrumen Soal

Untuk menguji validitas soal digunakan Uji Validitas Aiken. Validitas soal dilakukan melalui dua tahapan, yakni (i) validitas soal secara umum untuk mengetahui kaidah penulisan dan kesesuaian soal dengan kompetensi dasar yang akan diuji, aturan pembuatan soal berbentuk pilihan ganda, serta tata bahasa yang digunakan; dan (ii) validitas soal HOTS untuk mengetahui tingkat kemunculan indikator HOTS pada soal yang dikembangkan. Masing-masing tahapan dinilai atau divalidasi oleh dua orang dosen Pendidikan Matematika FKIP UNRAM. Secara umum berikut hasil validasi soal HOTS yang dikembangkan:

Tabel 3. Hasil Validasi Instrumen Soal

Jenis Validasi	Rata-Rata Validitas Aiken (%)
Umum	78,75
HOTS	73,125
Total	151,875
Rata-rata	75,9375

Berdasarkan Tabel 3 diperoleh bahwa instrumen soal yang dikembangkan memperoleh rata-rata validitas sebesar 75,9375% dengan kategori valid. Hasil ini diperoleh dari penilaian yang diberikan oleh ahli dengan memberikan skor dari 1 (Tidak Baik) sampai 5 (Sangat Baik). Selanjutnya skor tersebut dikalkulasikan untuk kemudian dianalisis lebih lanjut dengan menggunakan Uji Validitas Aiken untuk mengetahui tingkat validitas soal. Berdasarkan Tabel 3 juga dapat dilihat bahwa Validasi soal secara umum memperoleh skor 78,75% dengan kategori valid [1] artinya dalam penyusunan soal secara teknis dari aspek isi, konstruksi, dan bahasa sudah sesuai dengan pedoman yang seharusnya. Untuk Validasi soal HOTS memperoleh skor 73,125% dengan kategori valid [1] juga, hal ini menunjukkan soal yang disusun sudah sesuai dengan karakteristik soal HOTS, yakni menggunakan stimulus yang menarik dan kontekstual, soal mengungkap sifat kebaruan, jawaban tersirat pada stimulus, dan soal mengukur level kognitif penalaran [18].

Kepraktisan Soal

Kepraktisan soal diukur dari hasil angket respon siswa yang diberikan setelah siswa menjawab soal yang diberikan. Siswa berdasarkan

pengalamannya akan memberikan pendapat sangat setuju, setuju, tidak setuju, dan sangat setuju terhadap pernyataan yang diberikan pada angket. Pemberian angket diberikan pada tahap *one-to-one* dan *small group*. Bertujuan untuk mengetahui (i) karakteristik HOTS berupa kebaruan dan kemenarikan soal (menurut pengalaman siswa); dan (ii) kepraktisan soal berupa kebahasaan, alokasi waktu, serta kegrafikan.

Tabel 4. Hasil Kepraktisan Soal

Tahap	Kepraktisan (%)
<i>one-to-one</i>	80
<i>small group</i>	80,417

Dari Tabel 4 menunjukkan tingkat kepraktisan soal sudah praktis. Hal ini dilihat dari hasil baik pada tahap *one-to-one* maupun *small group* menunjukkan skor ≥ 80 . Kepraktisan soal juga menunjukkan bagaimana reaksi siswa terhadap soal. Berdasarkan skor yang diperoleh soal yang telah dikembangkan dapat membantu meningkatkan ketertarikan siswa dalam mengerjakan soal matematika. Hal ini sejalan dengan pendapat Widana dkk., bahwa soal HOTS dapat membantu meningkatkan motivasi siswa dalam belajar matematika [18]. Selain itu, berdasarkan angket respon tersebut soal yang disajikan merupakan permasalahan yang baru dan jarang mereka temukan dalam pembelajaran di kelas, hal tersebut tentunya mendukung salah satu kriteria HOTS yakni mengungkap kebaruan [18].

Keefektifan Soal

Berdasarkan pengalamannya ahli/validator akan menilai (memberikan tanggapan) apakah soal yang dikembangkan efektif dalam mengukur keterampilan berpikir tingkat tinggi siswa. Diperoleh skor untuk keefektifan soal menurut ahli sebesar 79% dengan kategori efektif. Selanjutnya, berdasarkan hasil analisis keterampilan berpikir tingkat tinggi diperoleh bahwa nilai rata-rata tingkat keterampilan berpikir tingkat tinggi siswa sebesar 6,6 berada dalam kategori cukup. Hal ini menunjukkan bahwa soal yang dikembangkan sudah efektif karena menunjukkan hasil yang sesuai dengan yang diharapkan. Sesuai dengan pendapat Sagala dan Andriani [13] menyatakan bahwa suatu soal dikatakan efektif apabila menurut ahli soal tersebut efektif dan soal tersebut memberikan hasil sesuai yang diharapkan. Selain itu dari hasil keterampilan berpikir tingkat tinggi tersebut mengindikasikan bahwa adanya efek potensial soal untuk membantu melatih keterampilan berpikir tingkat tinggi siswa.

Analisis Butir Soal Efektifitas Option dan Omit

Berdasarkan hasil analisis butir soal diperoleh bahwa semua pengecoh (*distractor*) berfungsi dengan baik artinya tidak ada pilihan

jawaban yang mengarahkan siswa pada jawaban yang benar atau menyesatkan dan pengecoh tersebut memiliki daya tarik yang besar bagi siswa yang kurang memahami konsep [1]. Selain itu, semua tingkatan omit pada soal dapat diterima. Hal ini menunjukkan bahwa soal dapat dipahami dan dikerjakan oleh siswa dengan baik.

Menurut Jakwerth [7] omit atau keadaan dimana siswa tidak merespon atau menjawab soal diakibatkan karena beberapa hal (i) siswa tidak memahami kalimat dalam soal baik kalimat pernyataan atau pertanyaan, (ii) soal dilewatkan oleh siswa tanpa sengaja, (iii) motivasi dalam menjawab soal yang kurang, (iv) waktu menjawab soal, (v) siswa kekurangan pengetahuan untuk menjawab soal tersebut. Jika diperhatikan kembali hasil angket respon siswa, jelas bahwa dari segi kalimat, gambar/grafik, waktu yang diberikan, dan ketertarikan/motivasi siswa dalam menjawab soal mendapat respon baik. Sehingga hasil semua omit diterima dan efektifitas pengecoh yang berfungsi sesuai dengan respon siswa yang baik pula terhadap soal.

Indeks Kesukaran dan Daya Pembeda

Dari segi indeks/tingkat kesukaran diperoleh soal dengan kategori mudah 0%, sedang 70%, dan sukar 30%. Hal ini menunjukkan bahwa sebagian besar soal berada pada kategori sedang yakni soal tidak terlalu mudah atau sulit. Selain itu soal yang baik adalah soal dengan indeks kesukaran dengan kategori sedang (0.30 – 0.70) [1]. Sedangkan jika ditentukan rata-ratanya akan diperoleh skor tingkat kesukaran soal secara umum adalah 0.33 dengan kategori sedang mendekati sulit.

Pada penelitian ini, daya pembeda soal dengan kategori sangat baik 0%, baik 20%, cukup 55%, jelek 10%, dan sangat jelek 15%. Artinya bahwa sebagian besar soal cukup dapat membedakan siswa dengan kemampuan tinggi dan rendah. Sedangkan jika ditentukan rata-ratanya akan diperoleh skor daya pembeda soal secara umum adalah 0.24 dengan kategori cukup mendekati jelek.

Hasil tingkat kesukaran soal yang menuju kearah sulit serta daya pembeda yang menuju kearah jelek disebabkan karena soal yang diberikan kepada siswa adalah soal dengan kategori HOTS semua. Apalagi siswa-siswa di sekolah tersebut belum terbiasa mengerjakan soal dengan level tersebut. Oleh karena itu, apabila guru akan membuat soal evaluasi maka soal tersebut sebaiknya tidak hanya soal yang HOTS (C4-C6) namun juga ada soal yang berbasis LOTS (C1-C3). Menurut Brookhart [3] persentase soal LOTS pada suatu tes adalah 55% sedangkan HOTS 45%. Selain itu, sangat penting untuk menentukan proporsi tingkat kesukaran soal, menurut Sudjana [15] perbandingan tingkat kesukaran soal dengan kategori mudah, sedang, dan sulit berturut-turut adalah 3:4:3.

Reliabilitas

Melalui skor siswa pada tahap uji lapangan (*Field Test*) diperoleh reliabilitas soal HOTS yang dikembangkan memiliki nilai 0,62. Jika mengkonsultasikan pada *r*-tabel diperoleh *r*-hitung > *r*-tabel sehingga instrumen soal sudah reliabel. Tingkat reliabilitas soal menunjukkan daya ketetapan soal, artinya apabila siswa dites kembali dengan soal tersebut akan memberikan hasil yang tetap. Kata “tetap” tidak harus sama, namun lebih menggambarkan kondisi jika hasil tes tersebut berubah-ubah hal tersebut dapat dikatakan tidak berarti [1]. Misalkan pada tes pertama skor A lebih tinggi dari B, maka apabila dites kembali skor A akan “tetap” lebih tinggi dari B bahkan jika skor A dan B berubah (tidak sama dengan skor pada tes awal).

KESIMPULAN

Berdasarkan analisis hasil penelitian dan pembahasan pada penelitian pengembangan soal HOTS matematika lingkup materi UN SMP/MTs di SMPN 1 Mataram tahun ajaran 2019/2020 dihasilkan perangkat soal yang terdiri atas kisi-kisi soal, kunci jawaban dan pembahasan, serta 20 soal HOTS matematika lingkup materi UN SMP/MTs berbentuk pilihan ganda yang valid, praktis, dan efektif. Perangkat soal berdasarkan pada penilaian ahli/validator memperoleh rata-rata skor sebesar 75,9375% yang menunjukkan kategori valid. Perangkat soal berdasarkan analisis angket respon siswa memperoleh rata-rata skor sebesar 80,417% dengan kategori praktis. Perangkat soal berdasarkan penilaian ahli/validator dengan memperoleh rata-rata skor sebesar 79% dengan kategori efektif dan menunjukkan hasil yang sesuai dengan yang diharapkan yakni rata-rata keterampilan berpikir tingkat tinggi siswa berada dalam kategori cukup dengan skor 6,6.

DAFTAR PUSTAKA

1. Arikunto, S. (2016). *Dasar - Dasar Evaluasi Pendidikan Edisi 2*. Jakarta: Bumi Aksara.
2. Bialik, M., & Fadel, C. (2015). *Mathematics for 21st Century: What Should Students Learn*. Boston: Central Curriculum Redesign.
3. Brookhart, S. M. (2010). *How To Asses Higher Order Thinking Skills in Your Class*. Virginia: ASCD.
4. BSNP. (2018). *Buku Saku Ujian Nasional 2019*. Jakarta: BSNP.
5. Hashimoto, Y. (1997). The Methods of Fostering Creativity Through Mathematical Problem Solving. *Zentralblatt fur Didaktik der Mathematik*, 29(3): 86 - 87.
6. Kunandar. (2013). *Penilaian Autentik Kurikulum 2013*. Jakarta: PT Grafindo Persada.
7. Jakwerth, P. M., Stancavage, F. B., Reed, E. D., & Research, A. I. (1999). *An Investigation of Why Students Do Not Respond to Questions*. Washington, D.C.: NAEP Validity Studies Panel.

8. Kemdikbud. (2019). *Laporan Hasil Ujian Nasional*. Retrieved September 19, 2019, from Pusat Penilaian Pendidikan: <https://hasilun.puspendik.kemdikbud.go.id>.
9. Lewy, Zulkardi, & Aisyah, N. (2009). Pengembangan Soal untuk Mengukur Kemampuan Berpikir Tingkat Tinggi Pokok Bahasan Barisan dan deret Bilangan di Kelas IX Akselerasi SMP Xaverius Maria Palembang. *Jurnal Pendidikan Matematika*, 3(2): 14-28.
10. Madu, A. (2016). Higher Order Thinking Skills (HOTS) in Math Learning. *IOSR Journal of Mathematic*, 13(5): 70 - 75.
11. Novaliah, & Kartowagiran, B. (2018). The Content Validity of Instrument of Character Education in Chemistry Learning. *Proceedings of 5th ICRiems*, Yogyakarta: 7-18 May 2018. 63 - 69.
12. Schleicher, *Andreas*. "PISA 2018: Insights and Interpretations." *OECD Publishing* (2019).
13. Sagala, P. N., & Andriani, A. (2019, March). Development of Higher-Order Thinking Skills (HOTS) Questions of Probability Theory Subject Based on Bloom's Taxonomy. In *Journal of Physics: Conference Series* (Vol. 1188, No. 1, p. 012025). IOP Publishing.
14. Sumarna, N. W., & T., H. (2017). The Increase of Critical Thinking Skills Through Mathematical Investigation Approach. *Journal of Physics: Conference Series*, 812(1): 1 - 8.
15. Sudjana, N. (2017). *Penilaian Hasil Proses Belajar Mengajar*. Bandung: Remaja Rosdakarya.
16. Susanto, E. (2018). *Mendikbud Jelaskan Kenapa Nilai Rata - Rata UN SMP 2018 Turun*. Retrieved Oktober 27, 2019, from [tempo.co: https://nasional.tempo.co/read/1092395/mendikbud-jelaskan-kenapa-nilai-rata-rata-hasil-un-smp-2018-turun](https://nasional.tempo.co/read/1092395/mendikbud-jelaskan-kenapa-nilai-rata-rata-hasil-un-smp-2018-turun).
17. Tessmer, M. (1994). Formative Evaluation Alternatives. *Performance Improvement Quarterly*, 7(1): 3-18.
18. Widana, I. W., Adi, S., Herdiyanto, Abdi, J., Marsito, & Istiqomah. (2019). *Penyusunan Soal Keterampilan Tingkat Tinggi Matematika*. Jakarta: Ditjen Pendidikan Dasar dan Menengah.
19. Widoyoko, E. P. (2012). *Teknik Penyusunan Instrumen Penelitian*. Yogyakarta: Pustaka Belajar.
20. Zulaiha, R. (2008). *Analisis Soal Secara Manual*. Jakarta: PUSPENDIK.
21. Zulkardi. (2006). *Formative Evaluation: What, why, when, and how*. Retrived Desember 4, 2019, from <https://www.oocities.org/zulkardi/boks.html>