



Analisis Soal Ujian Matematika Berkategori *Higher Order Thinking Skills* (Hots) Siswa SMP dengan Analisis Rasch Model

Ernawati

Sekolah Menengah Pertama Negeri Satu Atap 10 Pesawaran, Indonesia

Email: ernawati22@admin.smp.belajar.id

Received: 14 Nov, 2022 | Revised: 12 Dec, 2022 | Accepted: 29 Dec, 2022 | Published Online: 31 Dec, 2022

Abstract

Destination of this research is to determine the process and results of developing mathematical questions categorized as Higher Order Thinking Skills (HOTS) for the test. These types of research are Research and Development. The product developed is the Higher Order Thinking Skills (HOTS) Categorized Mathematics Exam for Junior High School Students. The trial sample was taken using a non probability sampling technique, namely accidental sampling. Research Procedure The development that will be carried out in this study is taken from the Suryabrata development research procedure. This study uses Rasch model data analysis techniques. The results obtained 18 items of HOTS mathematics test items consisting of 15 multiple choice questions and 3 valid essay items. Based on the analysis of the level of difficulty of the questions in the field trial trials obtained on multiple choice questions, there are eight easy category items and seven medium category items. In essay questions, there are one easy category items and two medium category items.

Keywords: exam questions; hots; rasch model

Abstrak

Tujuan dalam penelitian ini untuk mengetahui Proses dan hasil pengembangan soal matematika yang berkategori Higher Order Thinking Skills (HOTS) untuk ujian. Jenis penelitian ini adalah Penelitian dan Pengembangan. Produk yang dikembangkan adalah Soal Ujian Matematika Berkategori Higher Order Thinking Skills (HOTS) Siswa SMP. Pengambilan sampel uji coba dilakukan dengan teknik *non probability sampling* yaitu *accidental sampling*. Prosedur Penelitian pengembangan yang akan dilakukan pada penelitian ini diambil dari prosedur penelitian pengembangan Suryabrata. Penelitian ini menggunakan teknik analisis data Rasch model. Hasil penelitian diperoleh 18 butir soal ujian matematika berkategori HOTS yang terdiri dari 15 butir soal pilihan ganda dan 3 butir soal essay yang valid. Berdasarkan analisis tingkat kesukaran soal dalam uji coba uji lapangan diperoleh pada soal pilihan ganda terdapat delapan butir soal kategori mudah dan tujuh butir soal kategori sedang. Pada soal essay terdapat satu butir soal kategori mudah dan dua butir soal kategori sedang.

Kata Kunci: hots; rasch model; soal ujian

PENDAHULUAN

Permasalahan pendidikan muncul dan berkembang bersamaan dengan meningkatnya kemampuan siswa, situasi dan kondisi lingkungan yang ada, pengaruh informasi dan kebudayaan, serta berkembangnya ilmu pengetahuan dan teknologi.

Menurut Kamus Besar Bahasa Indonesia (KBBI), pendidikan ialah proses perubahan sikap dan tata laku seseorang ataupun kelompok dalam upaya mendewasakan manusia melalui sebuah pengajaran maupun pelatihan.

Pendidikan dapat mengubah pola pikir seseorang untuk selalu melakukan inovasi dan perbaikan dalam segala aspek kehidupan ke arah peningkatan kualitas diri. Pada pendidikan formal, penyelenggaraan pendidikan tidak lepas dari tujuan pendidikan yang akan dicapai karena tercapai atau tidaknya tujuan pendidikan merupakan tolak ukur dari keberhasilan penyelenggaraan pendidikan. Ki Hajar Dewantara, ia mengemukakan bahwa pengertian pendidikan ialah tuntunan tumbuh dan berkembangnya anak. Artinya, pendidikan merupakan upaya untuk menuntun kekuatan kodrat pada diri setiap anak agar mereka mampu tumbuh dan berkembang sebagai manusia maupun sebagai anggota masyarakat yang bisa mencapai keselamatan dan kebahagiaan dalam hidup mereka.

Newman serta Wehlage (Hamidah, 2018:75) menyatakan bahwa HOTS dapat memberi peranan yang sangat besar dalam mendukung prestasi akademik siswa, dengan HOTS siswa mampu memecahkan masalah, menyeleksi ide ataupun pendapat, berhipotesis, berpendapat dengan bijak serta sanggup menguasai situasi yang lebih rumit. Salah satu contoh di SMP Negeri 11 Pesawaran, Penilaian Hasil Belajar Ujian pada tahun pelajaran 2019/2020 untuk mata pelajaran matematika masih tergolong rendah. Hal ini disebabkan karena guru-guru juga belum terbiasa memberikan soal-soal matematika yang berkategori HOTS kepada siswa. Pemberian soal matematika berkategori HOTS tersebut diharapkan agar siswa dapat menumbuhkan kemampuan berpikir kritis, logis, sistematis, cermat, efektif dan efisien dalam menyelesaikan masalah tersebut. Diharapkan, hasil belajar siswa pada aspek pengetahuan dapat membekali siswa untuk memiliki keterampilan-keterampilan yang dibutuhkan pada abad ke-21 atau juga dikenal dengan sebutan *21st century skills*. Apa saja keterampilan abad 21 itu? Keterampilan abad 21 dibagi menjadi 3 kategori yaitu *learning skills*, *literacy skills*, dan *life skills*. Namun kenyataannya, siswa lebih sering cukup menghafalkan rumusnya tanpa sering latihan mengerjakan soal-soal.

Permendikbud No. 23 Tahun 2016 tentang Standar Penilaian Pendidikan yang merupakan kriteria mengenai lingkup, tujuan, manfaat, prinsip, mekanisme, prosedur, dan instrumen penilaian hasil belajar peserta didik. Soal adalah salah satu bentuk instrumen untuk mengukur keberhasilan siswa dalam pembelajaran. Soal digunakan untuk mengukur ketrampilan, pengetahuan, intelegensi, dan kemampuan yang dimiliki oleh individu atau kelompok. Soal berkategori HOTS adalah soal yang menuntut kemampuan berpikir khusus, seperti kemampuan berpikir logis, rasional, kritis, imajinatif, dan kreatif, kemampuan berpikir seperti itulah yang menjadi dasar

kemampuan berpikir tingkat tinggi. Untuk itu, soal yang digunakan untuk tujuan evaluasi harus berkualitas baik sehingga menghasilkan hasil pengukuran yang diandalkan. Melalui soal, guru dapat melakukan evaluasi pembelajaran dan mengetahui kemampuan akhir siswa setelah proses pembelajaran. Mahrens & Lehmann (1978 dalam Purwanto, 2013, hlm. 3) menyatakan bahwa evaluasi adalah suatu proses merencanakan, memperoleh dan menyediakan informasi yang sangat diperlukan untuk membuat alternatif-alternatif keputusan.

Pembelajaran matematika berorientasi pada HOTS siswa bertujuan agar kemampuan berpikir tingkat tinggi siswa menjadi meningkat. Dengan demikian untuk mengetahui tercapai tidaknya tujuan tersebut maka perlu adanya pengukuran HOTS siswa dalam pembelajaran matematika. Allen & Yen (dalam Lazwardi, 2017), mengemukakan bahwa pengukuran merupakan prosedur pemberian angka dengan menggunakan langkah sistematis sebagai sebuah pemaknaan terhadap apa yang telah dicapai. Reynolds, Livingston, & Willson (dalam Arifin & Retnawati, 2018), mendefinisikan bahwa pengukuran sebagai seperangkat aturan untuk menetapkan skor yang mewakili benda, sifat, atribut, atau perilaku. Berdasarkan kedua pendapat tersebut dapat disimpulkan bahwa pengukuran merupakan aturan atau prosedur untuk menetapkan angka atau skor terhadap objek tertentu berdasarkan kriteria-kriteria yang telah ditetapkan.

Beberapa pendapat ahli terkait pengertian HOTS antara lain dikemukakan oleh Thomas & Thorne (dalam Jailani & Sugiman, (2018: 3)), yang menyatakan bahwa berpikir tingkat tinggi adalah berpikir pada level yang lebih tinggi dari pada sekedar mengingat fakta atau menceritakan kembali sesuatu yang didengar kepada orang lain. Senada dengan pendapat tersebut, Lewis & Smith (dalam Jailani & Sugiman, (2018: 3)), menyatakan bahwa berpikir tingkat tinggi terjadi ketika seseorang memperoleh informasi baru dan disimpan dalam memori dan mengaitkan dan atau menata ulang dan memperluas informasi tersebut untuk mencapai tujuan atau menemukan kemungkinan jawaban dalam kondisi yang membingungkan. Dari pendapat kedua ahli tersebut secara ringkas dapat disimpulkan bahwa HOTS menuntut adanya proses berpikir yang lebih kompleks dalam menghadapi situasi atau memecahkan suatu masalah. Newmann (Goethals, 2013) mengatakan bahwa terdapat 3 elemen penting dalam mengefektifkan pengajaran HOTS, yaitu pengetahuan, keterampilan, dan disposisi (sikap terhadap matematika). Karena sudah terbiasa berlatih menggunakan keterampilan berfikir tingkat tinggi atau HOTS, siswa akan terbiasa berfikir kritis, kreatif dalam memecahkan permasalahan dan soal-soal tipe HOTS sesuai dengan tingkat usianya. Pada akhirnya mereka akan siap menghadapi ujian nasional yang berbasis soal HOTS (Fajriyatun & Saliman, (2018)). Fisher (dalam Saputra, 2021), menyatakan bahwa berpikir kritis

secara esensial adalah proses berpikir pada diri sendiri dengan tujuan untuk membuat keputusan-keputusan sesuai dengan akal sehat mengenai sesuatu yang dapat diyakini kebenarannya.

HOTS dalam taksonomi Bloom revisi yang dikemukakan oleh Anderson & Krathwohl (2001), pada dimensi proses kognitif HOTS meliputi proses menganalisis (*analyze*), mengevaluasi (*evaluate*), dan mencipta (*create*) (Liu, 2010), sedangkan pada dimensi pengetahuan HOTS meliputi pengetahuan konseptual (*conceptual knowledge*), pengetahuan prosedural (*procedural knowledge*), dan pengetahuan metakognitif (*metacognitive knowledge*). Adapun klasifikasi HOTS pada masing-masing dimensi dalam taksonomi Bloom revisi dapat dilihat pada Tabel 1.

Tabel 1. HOTS dalam Taksonomi Bloom Revisi

Dimensi Proses Kognitif						
	Mengingat	Memahami	Menerapkan	Menganalisis	Mengevaluasi	Mencipta
Dimensi Pengetahuan						
Faktual						
Konseptual						
Prosedural						Keterampilan Berpikir Tingkat Tinggi
Metakognitif						

Standar Penalaran dalam *National Council of Teachers of mathematics* (NCTM, 2017) merupakan standar penilaian HOTS karena soal-soal yang berlevel penalaran sudah pasti merupakan soal-soal yang berkategori HOTS. Adapun indikator standar penalaran dari prinsip-prinsip dan Standar Matematika Sekolah yaitu: (1) Mengenal pemahaman dan bukti sebagai aspek yang mendasar dalam matematika; (2) Membuat dan menyelidiki dugaan-dugaan matematis; (3) Mengembangkan dan mengevaluasi argument dan bukti matematis; (4) Memilih dan menggunakan berbagai macam pemahaman dan metode pembuktian. Tujuan dalam penelitian ini untuk mengetahui proses dan hasil pengembangan soal matematika yang berkategori *Higher Order Thinking Skills* (HOTS) untuk ujian.

Menurut Kamus Besar Bahasa Indonesia (KBBI), arti kata ujian adalah serangkaian tes yang dilakukan yang memiliki tujuan untuk menentukan kemampuan yang dimiliki oleh seorang pelajar apakah sudah memenuhi kriteria atau bahkan belum sama sekali. Pada jenjang pendidikan sekolah mulai dari yang kelas dasar sampai kelas atas semuanya sudah menggunakan yang namanya ujian. Ujian merupakan kegiatan pengukuran capaian kompetensi peserta didik yang dilakukan Satuan Pendidikan dengan mengacu pada Standar Kompetensi Lulusan untuk memperoleh pengakuan atas

prestasi belajar. Bentuk ujian atau tes yang paling umum dipakai oleh guru dalam menguji siswanya di kelas adalah tes tertulis.

Surapranata (2009) mengemukakan bahwa analisis kualitas soal dilakukan untuk mengetahui berfungsi tidaknya suatu soal maka perlu dianalisis baik secara kualitatif maupun kuantitatif untuk memperoleh informasi mengenai reliabilitas, validitas, daya beda, tingkat kesukaran, dll. Dalam pengukuran suatu tes terdapat banyak alat analisis yang dapat digunakan, salah satunya dengan menggunakan Teori Respon Butir (IRT). Butir yang memiliki tingkat kesulitan lebih tinggi memiliki probabilitas untuk menyelesaikan butir tersebut lebih rendah daripada butir yang lain.” (Rasch, 1960 dalam Bond dan Fox, 2015). Pemodelan Rasch merupakan satu model IRT yang paling populer. Prinsip dasar pemodelan Rasch adalah model *probabilistic* yang didefinisikan sebagai berikut: “individu yang memiliki tingkat abilitas yang lebih besar dibandingkan individu lainnya seharusnya memiliki peluang yang lebih besar untuk menjawab soal dengan benar. Dengan prinsip yang sama, butir yang lebih sulit menyebabkan peluang individu untuk mampu menjawabnya menjadi kecil.”

Di Indonesia, pengujian instrumen riset saat ini masih terbatas pada teori pengukuran klasik. Pada saat yang sama ini tersedia Pemodelan Rasch (Rasch Model) yang dapat menghasilkan instrumen pengukuran yang lebih baik sekaligus akurat. Sejauh ini, hanya Rasch Model yang merupakan alat analisis untuk dapat menguji validitas (kesahihan) dan realibilitas instrumen riset, bahkan menguji kesesuaian *person* dan *item* secara simultan--sesuatu yang belum tersaingi oleh teknik analisis lain. Rasch Model juga memiliki beberapa kelebihan karena memenuhi lima prinsip model pengukuran yaitu: (1) mampu memberikan ukuran yang linier dengan interval yang sama; (2) dapat mengatasi data yang hilang; (3) bisa memberikan estimasi yang lebih tepat; (4) mampu mendeteksi ketidaktepatan model; dan (5) memberikan instrumen pengukuran yang independen dari parameter yang diteliti (Sumintono, B. & Widhiarso, W, 2014).

METODE

Jenis penelitian ini adalah Penelitian dan Pengembangan. Produk yang dikembangkan adalah Soal Ujian Matematika Berkategori *Higher Order Thinking Skills* (HOTS) Siswa SMP. Data penelitian yang diperoleh adalah data hasil penyelesaian soal HOTS yang kemudian akan dianalisis menggunakan analisis Rasch Model. Prosedur penelitian ini terdiri atas beberapa tahapan yaitu: persiapan pelaksanaan penelitian, analisis data dan pembuatan laporan.

HOTS dapat diukur melalui tugas dan tes yang disusun berdasarkan aspek-aspek dan indikator yang terdapat pada HOTS. Pengukuran HOTS melalui tugas dapat

menggunakan rubrik yang dikembangkan secara lokal untuk tujuan mengevaluasi kemampuan berpikir. Sedangkan untuk pengukuran HOTS melalui tes dapat menggunakan soal pilihan ganda (*multiple choice*), soal uraian (*essay*), maupun bentuk soal lainnya.

Seluruh peserta didik pada lingkungan SMP se-Kabupaten Pesawaran merupakan subjek pada tahap ini. Dalam hal ini akan diwakili oleh 3 sekolah berdasarkan kriteria status akreditasinya, yaitu:

- a. SMP Negeri 19 Pesawaran Berakreditasi A
- b. SMP Negeri 11 Pesawaran Berakreditasi B
- c. SMP Negeri Satu Atap 10 Pesawaran Berakreditasi C

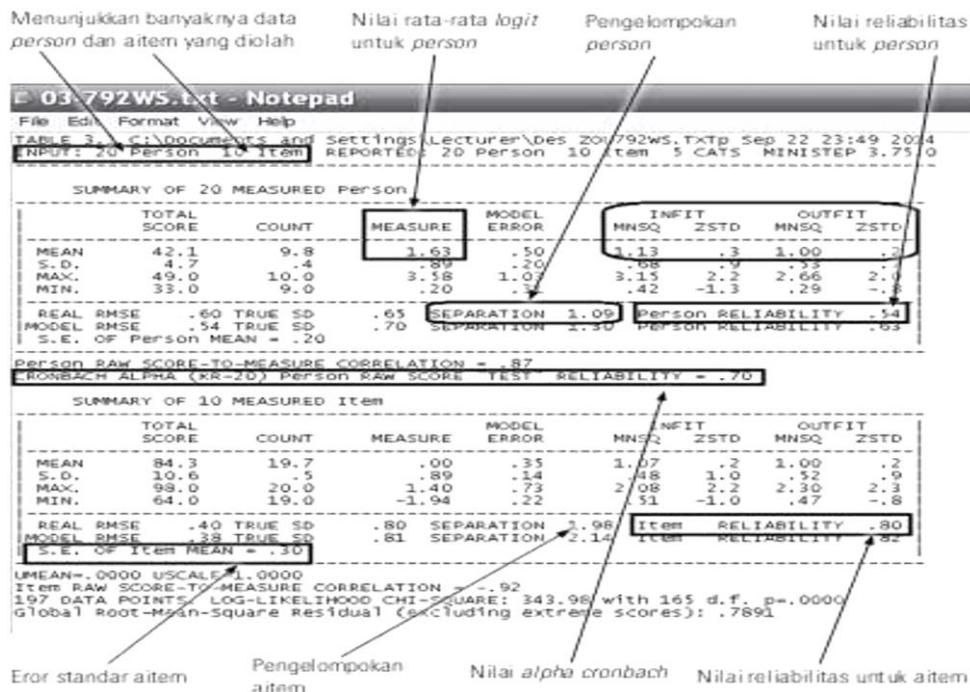
Pengambilan sampel uji coba dilakukan dengan teknik *non probability sampling* yaitu *accidental sampling*. Dalam *non probability sampling*, setiap unsur tidak memiliki kesempatan atau peluang yang sama untuk dipilih sebagai sample. Pemilihan unit sampling didasarkan pada pertimbangan atau penilaian subjektif dan tidak menggunakan teori probabilitas. Dalam teknik ini, sampel paling sedikit harus empat atau lima kali dari jumlah item pertanyaan. Dalam penelitian ini adalah 414 sampel, melebihi batas minimal yang seharusnya 160 sampel (15 soal pilihan ganda dan 3 soal essay).

Penelitian ini menggunakan Teknik Pengumpulan Data, yaitu sebagai berikut: (1) Data Validasi Ahli. Data Validasi Ahli dianalisis secara deskriptif dengan menelaah hasil penilaian dari para ahli terhadap soal ujian matematika berkategori HOTS. Hasil telaah digunakan sebagai masukan untuk merevisi/menyempurnakan soal ujian matematika yang digunakan; (2) Data tes soal berkategori HOTS. Tes soal ujian matematika berkategori HOTS dalam penelitian ini digunakan untuk memperoleh data kuantitatif setelah pembelajaran. Data tersebut akan dianalisis dengan tujuan yang spesifik yaitu untuk melihat komentar siswa terhadap kepraktisan dari perangkat soal. Tes juga dimaksudkan untuk melihat pencapaian siswa dalam mengerjakan soal untuk diuji validitas butir soalnya secara kuantitatif. Selain itu, fokus tes juga untuk mengetahui efek potensial siswa berupa kemampuan berpikir tingkat tinggi siswa dalam menyelesaikan soal-soal ujian matematika berkategori HOTS.

Penelitian ini menggunakan teknik analisis data Rasch model. Teknik pengolahan data menggunakan pemodelan Rasch model ini menggunakan Perangkat Lunak (software) Ministep yaitu program komputer khusus untuk analisis pemodelan Rasch yang bisa bekerja di bawah sistem Microsoft Windows yang dibuat oleh John Linacre. Paket program ini bersifat *freeware* (gratis dan tidak perlu meminta ijin pemakaian). Program ministep ini dapat diunduh di laman <http://www.winsteps.com/ministep.htm>.

Ministep memerlukan pengolahan data dengan menggunakan jenis berkas/file tertentu, yang isinya hanya berbentuk data mentah saja. Terdapat dua jenis data, yaitu data responden (demografi) dan data jawaban siswa. Pemodelan Rasch dengan program Ministep dapat melakukan semua analisis secara cepat dan hasilnya adalah kualitas informasi pengukuran yang baik dan informatif.

Berikut contoh hasil pengolahan data berdasarkan pemodelan Rasch dengan menggunakan program komputer *Ministep (Winstep Rasch)*.



Gambar 1. Output Summary Statistic Winstep

Tabel *Summary Statistics* memberikan info secara keseluruhan tentang kualitas responden secara keseluruhan, kualitas instrumen yang digunakan maupun interaksi antara *person* dan *item*.

Keterangan:

- a. *Person measure* menunjukkan rata-rata nilai responden dalam instrumen
 Nilai rata-rata yang lebih dari logit 0,0 menunjukkan kecenderungan responden yang lebih banyak menjawab benar di berbagai item.
- b. Nilai *alpha Cronbach* (mengukur reliabilitas, yaitu interaksi antara *person* dan *item* secara keseluruhan)

Tabel 2. Kriteria Reliabilitas Berdasarkan Nilai Alpha Cronbach

Nilai <i>alpha Cronbach</i> (<i>n</i>)	Keterangan
$n < 0,5$	Buruk
$0,5 \leq n < 0,6$	Jelek
$0,6 \leq n < 0,7$	Cukup
$0,7 \leq n < 0,8$	Bagus
$n \geq 0,8$	Bagus sekali

c. Nilai *Person Reliability* dan *Item Reliability*

Nilai *Person Reliability* digunakan untuk melihat konsistensi jawaban dari responden, sedangkan *Item Reliability* untuk mengetahui kualitas itemitem dalam instrumen.

Tabel 3. Kriteria Nilai *Person Reliability* dan *Item Reliability*

Nilai <i>Person Reliability</i> dan <i>Item Reliability</i> (<i>P</i>)	Keterangan
$P < 0,67$	Lemah
$0,67 \leq P < 0,80$	Cukup
$0,81 \leq P < 0,90$	Bagus
$0,91 \leq P < 0,94$	Bagus sekali
$P \geq 0,94$	Istimewa

d. Pengelompokkan *person* dapat diketahui dari nilai *separation*. Makin besar nilai *separation* maka kualitas instrumen dalam hal keseluruhan responden dan aitem makin bagus, karena bisa mengidentifikasi kelompok responden dan kelompok aitem. Persamaan lain yang digunakan yang melihat pengelompokkan secara lebih teliti disebut pemisahan strata:

$$H = \frac{[4 \times SEPARATIO N + 1]}{3}$$

e. Untuk memeriksa butir soal yang tidak sesuai (*outliers* atau *misfits*) dengan meninjau hal-hal berikut ini.

- 1) Nilai *Oufit Mean Square (MNSQ)* yang diterima : $0,5 < MNSQ < 1,5$
- 2) Nilai *Oufit Z-Standard (ZSTD)* yang diterima : $-2,0 < ZSTD < +2,0$
- 3) Nilai *Point Measure Coorelation (Pt Mean Corr)* : bernilai positif.

Soal yang *valid/fit* minimal memenuhi salah satu kriteria diatas (Sumintono & Widiharso, 2015: 98).

f. Tingkat kesulitan butir soal akan ditunjukkan pada Item map yang akan menampilkan hasil berupa grafik vertikal. Untuk menentukan kategori mudah, sedang, atau sulit sebuah soal dengan memperhatikan:

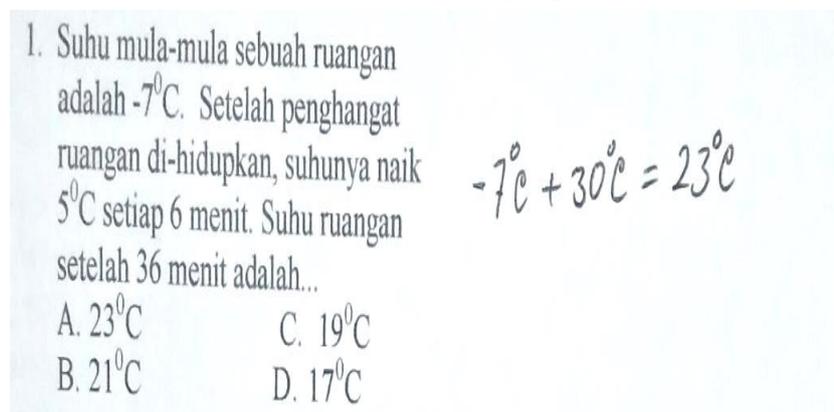
- 1) Kelompok butir soal mudah, dapat diketahui dari nilai batas bawah yang diperoleh dari rata-rata item logit dari Tabel item measure.
- 2) Kelompok butir soal sulit, dapat diketahui dari nilai batas atas yang diperoleh dari rata-rata person logit dari Tabel person measure.
- 3) Kelompok butir soal sedang berada diantara batas atas dan batas bawah.

Sumber: <http://www.rasch.org/rmt/rmt211m.html>

HASIL DAN PEMBAHASAN

Penelitian ini merupakan analisis Soal Ujian Matematika Berkategori *Higher Order Thinking Skills* (HOTS) yang dikembangkan berdasarkan spesifikasi butir soal yang disusun sesuai dengan Kompetensi Inti, Kompetensi Dasar, dan Indikator dari materi yang akan dibuat soal. Butir soal yang akan dianalisis terdapat dalam kisi-kisi penulisan soal tes yang terlebih dahulu ditelaah secara kualitatif, selanjutnya dijadikan pedoman penulisan soal untuk menghasilkan produk berupa soal tes. Produk yang dihasilkan kemudian diujikan kepada 414 peserta didik kelas IX SMP di Pesawaran. Soal tes sebanyak 15 butir soal pilihan ganda dan 3 butir soal essay. Soal tes dikerjakan oleh peserta tes yang ditetapkan sebagai *anchor (common-person)*.

Berdasarkan Taksonomi Bloom Revisi dan Standar Penalaran dalam *National Council of Teachers of mathematics* (NCTM, 2017), berikut merupakan analisis jawaban peserta didik dalam menjawab soal-soal yang diberikan :



Gambar 2. Hasil Pekerjaan Siswa dengan Jawaban Benar tetapi Penyelesaian Kurang

Berdasarkan hasil kerja siswa di atas, terlihat bahwa siswa menjawab soal tersebut sudah menggunakan nalar, ini terlihat dari jawaban yang diberikan sudah benar. Tetapi penjelasan dari jawaban tersebut masih kurang, seharusnya hasil dari jawaban tersebut dijabarkan lebih jelas lagi sehingga siswa lain bisa memahami dengan cepat.

2. Penjual mempunyai 8 karung beras. Beras yang ada pada setiap karung sebanyak 0,73 kuintal dan sebanyak $\frac{5}{8}$ dari seluruh beras sudah terjual. Jika sisanya akan dibagikan kepada keenam adiknya sama rata, berapa banyak beras yang diterima masing-masing?

A. 36,5 kg C. 39,5 kg
B. 38 kg D. 42 kg

$$\frac{5}{8} \times 0,73 = 5 \times 0,09125$$

$$= 0,45625$$

$$= 45$$

$$= 42$$

Gambar 3. Hasil Pekerjaan Siswa dengan Penyelesaian Kurang Tepat

Berdasarkan hasil kerja siswa di atas, terlihat bahwa siswa menggunakan strategi yang kurang tepat yaitu siswa salah dalam melakukan perhitungan. Hal ini dikarenakan siswa kurang memahami strategi apa yang harus digunakan untuk menyelesaikan soal di atas.

3. Perbandingan uang Ani dan Ina 3 : 5. Jumlah uang mereka Rp400.000,-. Selisih uang keduanya adalah...

A. Rp80.000,- C. Rp150.000,-
B. Rp100.000,- D. Rp200.000,-

Ani: Ina = 3 : 5

$$3x + 5x = 400.000$$

$$8x = 400.000$$

$$x = \frac{400.000}{8} = 50.000$$

$$\text{Ina} - \text{Ani} = 5 \cdot 50.000 - 3 \cdot 50.000$$

$$= 250.000 - 150.000$$

$$= 100.000$$

Siswa mempunyai pemahaman konsep yang baik terkait materi perbandingan

Gambar 4. Hasil Pekerjaan Siswa dengan Penyelesaian Tepat Soal Pilihan Ganda no 3

Berdasarkan hasil kerja siswa di atas menunjukkan bahwa siswa tersebut mampu memberikan argumen bagaimana menyusun strategi untuk menghitung selisih uang dengan menggunakan perbandingan. Hal ini menunjukkan bahwa siswa tersebut mempunyai kemampuan dalam penalaran yang bagus.

4. Sebuah mesin di suatu pabrik minuman mampu memasang tutup botol untuk 14 botol dalam waktu 84 detik. Banyak botol yang dapat di tutup oleh mesin dalam waktu 2 menit adalah ...

A. 16 botol C. 28 botol
B. 20 botol D. 35 botol

84 detik = $84:60$ menit = 1,4 menit

$1,4 \times 2$ menit = 2,8

= 28 Botol

Siswa melakukan penghitungan tetapi masih salah

Gambar 5. Hasil Pekerjaan Siswa dengan Penyelesaian Kurang Tepat Pilihan Ganda no 4

Berdasarkan hasil kerja siswa di atas, terlihat bahwa siswa menggunakan strategi yang kurang tepat yaitu siswa salah dalam melakukan perhitungan. Hal ini dikarenakan siswa tersebut belum menggunakan penalaran dalam menyelesaikan masalah di atas.

5. Terdapat empat toko olahraga yang menjual bola basket dan bola sepak dengan daftar dan diskon seperti pada tabel berikut:

Barang	Harga	Diskon			
		Toko A	Toko B	Toko C	Toko D
Bola Basket	Rp 400.000	30%	25%	40%	15%
Bola Sepak	Rp 300.000	30%	35%	20%	45%

Desta akan membeli 1 bola basket dan 1 bola sepak. Di toko manakah Desta berbelanja agar mendapatkan harga yang paling murah?

A. Toko A C. Toko C
B. Toko B D. Toko D

Siswa mampu mengevaluasi atau menentukan pilihan

Karena diskon ke 4 toko sama maka kita harus cari toko yang memberi harga dan diskon yang tinggi sehingga harganya murah. Jadi toko yang paling murah adalah toko C

Siswa mampu menuliskan argumen berdasarkan hasil temuannya

Gambar 6. Hasil Pekerjaan Siswa dengan Penyelesaian Tepat Pada Pilihan Ganda no 5

Berdasarkan hasil kerja siswa di atas menunjukkan bahwa siswa tersebut mampu mengevaluasi dan memberikan argumen bagaimana menyusun strategi untuk mengetahui toko mana yang memberi harga lebih murah. Hal ini menunjukkan bahwa siswa tersebut mempunyai kemampuan dalam penalaran yang bagus.

1. Perbandingan harga bawang putih dan harga bawang merah adalah 5 : 7. Sementara itu, perbandingan harga bawang putih dan harga bawang bombay adalah 4 : 6. Selisih harga bawang merah dengan harga bawang bombay per kg Rp4.200,- Berapa jumlah harga bawang merah dan harga bawang bombay?

Putih	Merah	Bombai	
20x	28x	30x	
	Bombai - Merah =	4.200	
	30x - 28x =	4.200	
	2x =	4.200	
	x =	4.200	
		2	
	x =	2.100	
Jumlah =	Merah + Bombai		
	= 28x + 30x		
	= 28(2.100) + 30(2.100)		
	= 58.800 + 63.000		
	= 121.800		

Siswa menggunakan strategi yang kurang tepat dalam menentukan harga bawang merah dan bawang bombay

Gambar 7. Hasil Pekerjaan Siswa dengan Penyelesaian Kurang Tepat Pada Soal Essay no 1

Berdasarkan hasil kerja siswa di atas, terlihat bahwa siswa menggunakan strategi yang kurang tepat yaitu siswa salah dalam menggunakan rumus untuk perhitungan. Hal ini dikarenakan siswa belum bisa mencipta dari suatu permasalahan.

2. Zaki akan membuat 50 teralis yang akan di pasang pada pagar rumahnya. Teralis tersebut berbentuk juring lingkaran yang terbuat dari besi. Jika panjang jari-jari lingkaran 27 cm dan besar sudut pusat 40° . Hitunglah panjang besi minimum yang diperlukan untuk membuat teralis tersebut? ($\pi = 3,14$).

$$\frac{40^\circ}{360^\circ} \times \pi \times 27^2$$

$$\frac{1}{9} \times 3,14 \times 729$$

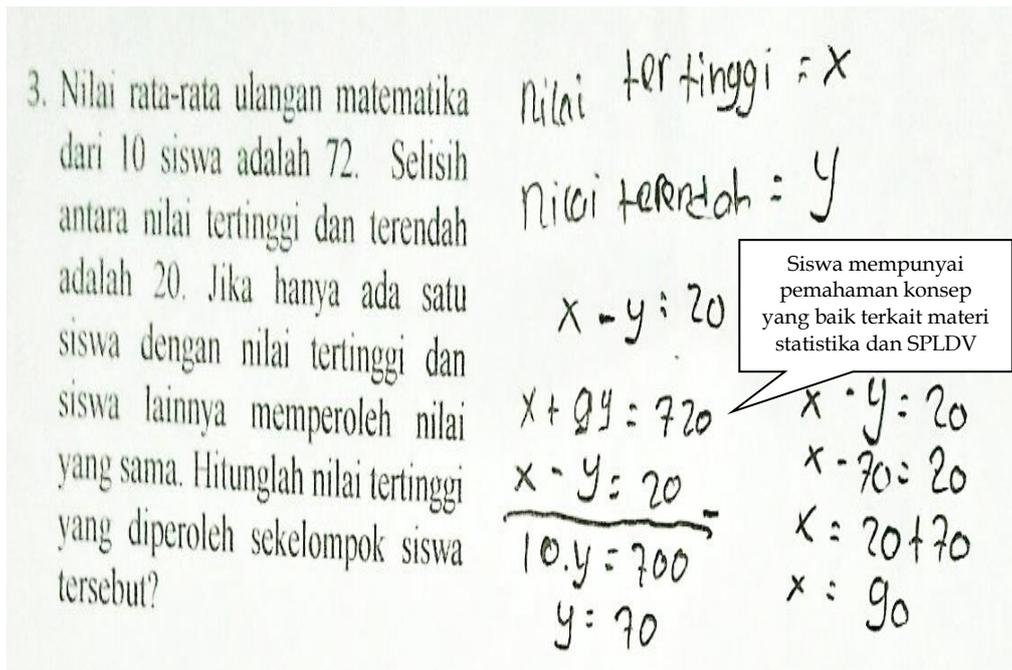
$$= 254,34 \times 50$$

$$= 12717$$

Siswa menggunakan strategi yang kurang tepat dalam menentukan panjang besi minimum untuk membuat teralis

Gambar 8. Hasil Pekerjaan Siswa dengan Penyelesaian Kurang Tepat Soal Essay no 2

Berdasarkan hasil kerja siswa di atas, terlihat bahwa siswa menggunakan strategi yang kurang tepat yaitu siswa salah dalam melakukan perhitungan. Hal ini dikarenakan siswa belum bisa mencipta untuk merumuskan hipotesis dan memproduksi dalam menyelesaikan permasalahan tersebut.



Gambar 9. Hasil Pekerjaan Siswa dengan Penyelesaian Tepat Pada Soal Essay no 3

Berdasarkan hasil kerja siswa di atas menunjukkan bahwa siswa tersebut mempunyai pemahaman konsep dua materi sekaligus untuk menghitung nilai tertinggi yang diperoleh sekelompok siswa yaitu materi statistika dan SPLDV. Hal ini menunjukkan bahwa siswa tersebut mempunyai kemampuan untuk menganalisis permasalahan tersebut.

Secara keseluruhan perangkat soal yang dikembangkan memiliki efek potensial yaitu mampu mengukur kemampuan berpikir tingkat tinggi siswa. Soal yang dikembangkan menunjukkan hasil kemampuan berpikir tingkat tinggi siswa dalam kategori baik dalam artian sebagian besar siswa mampu menyelesaikan soal-soal yang diberikan dengan baik. Hal ini selaras dengan temuan dalam penelitian yang dilakukan Lewy (2009). Dalam penelitian tersebut soal-soal yang mengukur kemampuan berpikir tingkat tinggi memiliki efek potensial terhadap siswa yaitu hasil kemampuan siswa dalam kategori baik.

Setelah dilakukan tes uji coba terhadap siswa, selanjutnya dilakukan rekapitulasi skor yang diperoleh tiap siswa. Skor yang diperoleh merupakan skor mentah yang harus dianalisis lebih lanjut. Selanjutnya skor mentah tersebut diolah dengan menggunakan perangkat lunak Winsteps.

Pada Tabel 4 menampilkan data statistik dari analisis model Rasch dari butir soal pilihan ganda, data yang diperoleh dari 75 responden/siswa yang menjawab 15 item/soal yang diberikan yang merupakan hasil pengolahan dari perangkat lunak Winsteps. Pada tabel tersebut terlihat nilai dari *person reliability* adalah 0,22. Hal ini

menunjukkan konsistensi jawaban dari siswa lemah, namun nilai dari *item reliability* adalah 0,60 yang menyimpulkan bahwa kualitas item/soal dalam instrumen secara keseluruhan dalam kualitas cukup. Untuk mengetahui reliabilitas secara keseluruhan dapat dilihat berdasarkan nilai alpha cronbach (KR-20) yaitu 0,66. Hal ini menunjukkan nilai reliabilitas yaitu interaksi antara *person* dan *item* secara keseluruhan dalam kategori cukup sehingga analisis lebih lanjut layak untuk dilakukan.

Tabel 4. Ringkasan Statistik Butir Soal Pilihan Ganda

SUMMARY OF 75 MEASURED Person									
	TOTAL			MODEL	INFIT		OUTFIT		
	SCORE	COUNT	MEASURE	S.E.	MNSQ	ZSTD	MNSQ	ZSTD	
MEAN	12.8	15.0	2.45	1.17					
SEM	.3	.0	.17	.07					
P.SD	2.2	.0	1.49	.60					
S.SD	2.2	.0	1.50	.61					
MAX.	15.0	15.0	4.01	1.81					
MIN.	10.0	15.0	.76	.57					
REAL RMSE	1.32	TRUE SD	.69	SEPARATION	.52	Person	RELIABILITY	.22	
MODEL RMSE	1.32	TRUE SD	.70	SEPARATION	.53	Person	RELIABILITY	.22	
S.E. OF Person MEAN = .17									
Person RAW SCORE-TO-MEASURE CORRELATION = .99									
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .66									
SUMMARY OF 15 MEASURED Item									
	TOTAL			MODEL	INFIT		OUTFIT		
	SCORE	COUNT	MEASURE	S.E.	MNSQ	ZSTD	MNSQ	ZSTD	
MEAN	63.9	75.0	.00	.38	1.00	-.12	1.04	-.02	
SEM	1.1	.0	.17	.02	.02	.23	.05	.26	
P.SD	4.2	.0	.62	.07	.07	.88	.17	.96	
S.SD	4.4	.0	.64	.07	.07	.91	.18	.99	
MAX.	72.0	75.0	.87	.60	1.13	1.06	1.57	1.33	
MIN.	57.0	75.0	-1.50	.32	.82	-2.98	.81	-2.84	
REAL RMSE	.39	TRUE SD	.48	SEPARATION	1.21	Item	RELIABILITY	.60	
MODEL RMSE	.39	TRUE SD	.48	SEPARATION	1.24	Item	RELIABILITY	.61	
S.E. OF Item MEAN = .17									

Pada Tabel 5 menampilkan data statistik dari analisis model Rasch dari butir soal essay, data yang diperoleh dari 75 responden/siswa yang menjawab 3 item/ soal yang diberikan yang merupakan hasil pengolahan dari perangkat lunak Winsteps. Pada tabel tersebut terlihat nilai dari *person reliability* adalah 0,66. Hal ini menunjukkan konsistensi jawaban dari siswa cukup, sedangkan nilai dari *item reliability* adalah 0,78 yang menyimpulkan bahwa kualitas item/soal dalam instrumen secara keseluruhan juga dalam kualitas cukup. Untuk mengetahui reliabilitas secara keseluruhan dapat dilihat berdasarkan nilai alpha cronbach (KR-20) yaitu 0,65. Hal ini menunjukkan nilai reliabilitas yaitu interaksi antara *person* dan *item* secara keseluruhan dalam kategori cukup sehingga analisis lebih lanjut layak untuk dilakukan.

Tabel 5. Ringkasan Statistik Butir Soal Essay

SUMMARY OF 75 MEASURED Person										
	TOTAL		MEASURE		MODEL		INFIT		OUTFIT	
	SCORE	COUNT			S.E.	MNSQ	ZSTD	MNSQ	ZSTD	
MEAN	8.9	3.0	1.64		1.13	1.00	-.02	1.00	-.02	
SEM	.2	.0	.26		.01	.08	.10	.08	.10	
P.SD	1.8	.0	2.20		.11	.70	.85	.72	.86	
S.SD	1.9	.0	2.22		.11	.70	.86	.72	.86	
MAX.	11.0	3.0	4.30		1.24	2.91	1.53	2.95	1.55	
MIN.	4.0	3.0	-3.80	.94	.05	-2.01	.05	-2.01		
REAL RMSE	1.28	TRUE SD	1.79	SEPARATION	1.40	Person RELIABILITY	.66			
MODEL RMSE	1.14	TRUE SD	1.89	SEPARATION	1.66	Person RELIABILITY	.73			
S.E. OF Person MEAN = .26										
Person RAW SCORE-TO-MEASURE CORRELATION = 1.00										
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .65										
SUMMARY OF 3 MEASURED Item										
	TOTAL		MEASURE		MODEL		INFIT		OUTFIT	
	SCORE	COUNT			S.E.	MNSQ	ZSTD	MNSQ	ZSTD	
MEAN	223.3	75.0	.00		.22	.99	-.17	1.00	-.19	
SEM	7.1	.0	.35		.00	.13	.99	.16	1.17	
P.SD	10.1	.0	.50		.00	.19	1.40	.22	1.66	
S.SD	12.3	.0	.61		.00	.23	1.72	.27	2.03	
MAX.	237.0	75.0	.51		.22	1.21	1.32	1.20	1.28	
MIN.	213.0	75.0	-.67		.22	.74	-2.05	.69	-2.50	
REAL RMSE	.23	TRUE SD	.44	SEPARATION	1.89	Item RELIABILITY	.78			
MODEL RMSE	.22	TRUE SD	.44	SEPARATION	1.98	Item RELIABILITY	.80			
S.E. OF Item MEAN = .35										

Kelebihan soal ujian yang peneliti kembangkan dibandingkan dengan hasil pengembangan soal yang serupa adalah soal-soal yang dikembangkan terfokus pada pengukuran berpikir tingkat tinggi, yaitu selain kemampuan menganalisis, ke-mampuan mengevaluasi, dan kemampuan mencipta, juga ditambah dengan kemampuan penalaran. Hal ini dimaksudkan untuk memberikan gambaran soal-soal ujian yang setara dengan soal kemampuan berpikir tingkat tinggi yang hanya sebagian kecil siswa mampu menyelesaikannya. Hasil pengembangan soal ini dapat dijadikan sebagai bahan pendidik untuk melatih dan meningkatkan kemampuan berpikir tingkat tinggi siswa.

Selain kelebihan dari penelitian pengembangan ini juga terdapat keterbatasan atau kelemahan. Adapun kelemahan pengembangan soal ini adalah soal-soal ujian matematika berkategori HOTS hasil pengembangan jika diterapkan kepada siswa dalam proses pembelajaran memerlukan waktu yang tidak sedikit serta dibutuhkan kesabaran pendidik untuk melatih kemampuan berpikir tingkat tinggi siswa. Uji coba hasil pengembangan soal dilakukan ditiga sekolah SMP Negeri di Pesawaran yang mungkin dapat diterapkan di sekolah lain namun dengan kondisi siswa, guru, dan lingkungan belajar yang tidak jauh berbeda dengan ketiga SMP Negeri di Pesawaran tersebut.

Penerapan hasil pengembangan pada siswa sekolah lain perlu pengkajian dan penyesuaian lebih lanjut.

SIMPULAN

Hasil analisis soal ujian matematika berkategori HOTS diperoleh perangkat soal ujian matematika berkategori HOTS sebanyak 18 butir soal yang terdiri dari 15 butir soal pilihan ganda dan 3 butir soal essay yang valid. Berdasarkan analisis tingkat kesukaran soal dalam uji coba uji lapangan diperoleh pada soal pilihan ganda terdapat delapan butir soal kategori mudah dan tujuh butir soal kategori sedang. Pada soal essay terdapat satu butir soal kategori mudah dan dua butir soal kategori sedang.

Berdasarkan hasil kesimpulan, saran yang dapat dikemukakan sebagai berikut:

1. Bagi guru yang hendak menggunakan instrument soal ini hendaknya dapat melatih kemampuan berpikir tingkat tinggi siswa dalam proses pembelajaran serta dapat mengembangkan atau mendesain soal-soal ujian matematika berkategori HOTS.
2. Bagi peneliti lain, hendaknya dapat mengembangkan instrument soal yang dapat dipergunakan sebagai masukan untuk mengembangkan atau mendesain soal-soal ujian matematika berkategori *Higher Order Thinking Skills* (HOTS) yang lebih luas cakupan materinya.

REFERENSI

- Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for learning, teaching, and assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York, NY: Addison Wesley Longman.
- Arifin, Z., & Retnawati, H. (2018). Pengembangan Instrumen Pengukur Higher Order Thinking Skills Matematika Siswa SMA Kelas X. *Pythagoras: Jurnal Matematika dan Pendidikan Matematika*, 12(1) <https://doi.org/10.21831/pg.v12i1.14058>.
- Bloom, B. S. (1956). *Taxonomy of Education Objectives: The Classification of Educational Goals*. London: Ann Arbor, MI: Penayang Addison Wesley.
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch Model Fundamental Measurement in the Human Sciences, Third Edition*. New York: Routledge
- Fajriyatun., & Saliman. (2018). Media Stupa Merah Putih Bermagnet untuk Meningkatkan HOTS dan Hasil Belajar dalam Penguatan Pendidikan Karakter. *Jurnal Didaktika Pendidikan Dasar*, 2(2), 319-340.
- Goethals, P. L. (2013). *The pursuit of higher-order thinking in the mathematics classroom: A review*. West Point, NY: Centre for Faculty Excellence, United States Military Academy.
- Hamidah, Luluk. 2018. *Higher Order Thinking Skills (Seni Melatih Kemampuan Berpikir Tingkat Tinggi)*. Yogyakarta: Hijaz Pustaka Mandiri.

- Jailani., & Sugiman. (2018). *Desain Pembelajaran Matematika Untuk Melatih-kan Higher Order Thinking Skills*. UNY PRESS. Yogyakarta. 3.
- Lazwardi. (2017). Implementasi Evaluasi Program Pendidikan di Tingkat Sekolah Dasar dan Menengah. *Jurnal Kependidikan Islam*, 7(1).
<https://doi.org/10.24042/alidarah.v7i1.1112>
- Lewy. (2009). Pengembangan Soal Untuk Mengukur Kemampuan Berpikir Tingkat Tinggi. *Jurnal Pendidikan Matematika*, 3(2).
<https://doi.org/10.22342/jpm.3.2.326>
- Liu, X. (2010). *Essentials of sciences classroom assessment*. Los Angeles, LA: SAGE Publication.
- NCTM. (2017). *Principles and standard for school mathematics*. Reston, VA: The National Council of Teacher Mathematics, Inc.
- Purwanto. 2013. *Evaluasi hasil belajar*. Yogyakarta: Pustaka Pelajar.
- Saputra, I. G. E. (2021). Pengaruh Game Edukasi Adventure Berbantuan Online HOTS Test Terhadap Keterampilan Berpikir Kritis Siswa. *Jurnal Didaktika Pendidikan Dasar*, 5(3). <https://doi.org/10.26811/didaktika.v5i3.301.715-736>.
- Surapranata, S. (2009), *Analisis, Validitas, Reliabilitas dan Interpretasi Hasil Tes*. Bandung: PT. Remaja Rosdakarya.