
PENGEMBANGAN INSTRUMEN PENILAIAN BERBASIS *HIGHER ORDER THINKING (HOT)* PELAJARAN MATEMATIKA KELAS IV SEKOLAH DASAR

Andy Rusdianto

Program Magister Pendidikan Dasar Universitas Bengkulu.
andyadel54@gmail.com

Agus Susanta

Program Magister Pendidikan Dasar Universitas Bengkulu.

Abdul Muktadir

Program Magister Pendidikan Dasar Universitas Bengkulu.

Abstract

Penelitian ini bertujuan mendeskripsikan bagaimana pengembangan instrumen penilaian berbasis HOT, serta untuk menghasilkan instrumen penilaian berbasis HOT pelajaran Matematika kelas IV sekolah dasar yang valid dan reliabel. Penelitian dan pengembangan ini diadaptasi dari model Borg dan Gall. Prosedur penelitian mencakup: penelitian dan pengumpulan informasi; perencanaan; pengembangan produk awal; uji coba terbatas; revisi produk awal; dan uji coba lapangan. Hasil penelitian menunjukkan bahwa instrumen penilaian berbasis HOT yang dikembangkan berupa 15 soal pilihan ganda dan 5 soal uraian pada aspek materi, konstruksi dan bahasa dinyatakan valid secara teoritis berdasarkan judgment ahli dan secara empiris instrumen valid dan reliabel berdasarkan hasil analisis ujicoba lapangan. Pada soal pilihan ganda rata-rata validitas butir soal 0,46 “cukup baik”, koefisien reliabilitas sebesar 0,73 “baik”, sedangkan soal uraian rata-rata validitas butir soal 0,76 “baik”, koefisien reliabilitas sebesar 0,81 “baik”. Secara umum instrumen penilaian yang dikembangkan telah memenuhi kriteria tes yang baik, yaitu valid dan reliabel.

Keywords: Kata Kunci : instrumen penilaian, HOT, matematika

Pendahuluan

Matematika adalah salah satu pelajaran yang diberikan kepada semua peserta didik sejak dari sekolah dasar, melalui pelajaran Matematika diharapkan peserta didik memiliki kemampuan berpikir logis, analitis, sistematis, kritis, kreatif, serta kemampuan bekerjasama (PP No.19. 2005). Hal ini sesuai dengan karakteristik skills pebelajar abad ke-21 yang dipublikasikan oleh *Partnership of 21st Century Skill* (2011: 1) yang berfokus pada kemampuan berpikir tingkat tinggi disebut dengan 4 C's, yaitu berpikir kritis (*critical thinking*), kreatif (*creativity*), komunikasi (*communication*), dan (*collaboration*).

Melalui pengintegrasian *Partnership of 21st Century Skill* diharapkan peserta didik memiliki sikap yang baik, dapat bersaing dan meraih kesuksesan di kehidupan mereka di masa yang akan datang. Hal ini sesuai dengan tuntutan Kurikulum 2013 dalam Kemendikbud (2016: 5), yaitu pengembangan kompetensi Matematika di arahkan untuk meningkatkan kecakapan hidup, terutama dalam membangun kreatifitas, kemampuan berpikir kritis, kolaborasi atau bekerjasama, dan keterampilan berkomunikasi.

Kemampuan berpikir tingkat tinggi merupakan kemampuan yang sangat penting dimiliki oleh peserta didik, karena kemampuan berpikir tingkat tinggi akan menjadikan peserta didik terbiasa menghadapi tantangan. Hal ini sejalan dengan pemikiran Ramos, Dolipas, dan Villamor (2013: 49) yaitu “kemampuan berpikir tingkat tinggi diperlukan peserta didik karena diyakini dapat mempersiapkan peserta didik untuk menghadapi tantangan pekerjaan dalam kehidupan sehari-hari”.

Berdasarkan hasil *Trend in International Mathematics and Science Study* atau *TIMMS* tahun 2015 (Puspendik Kemdikbud. 2015), hasil skor pencapaian Indonesia di bidang Matematika kelas 4 adalah 379 poin menduduki peringkat ke 45 dari 50 Negara. Hasil survey *TIMMS* menunjukkan kemampuan siswa Indonesia dalam mengerjakan soal-soal komputasi rutin 57% benar, namun dalam mengerjakan soal-soal dalam domain bernalar masih sangat minim. Hal ini menunjukkan bahwa mayoritas peserta didik Indonesia masih berada pada tataran *LOTS (Lower Order Thinking Skills)* atau kemampuan berpikir tingkat rendah, karena kemampuan berpikir peserta didik masih sekedar cenderung mengingat, menyatakan kembali, dan merujuk tanpa melakukan sesuatu.

Jarangnya guru memberikan soal-soal yang menuntut kemampuan berpikir tingkat tinggi kepada peserta didik, sesuai juga dengan penelitian yang dilakukan oleh Khan dan Inamullah (2011: 151) menyatakan “guru lebih banyak memberikan soal-soal *LOT* kepada peserta didik dari pada soal-soal yang *HOT*”.

Peserta didik tingkat sekolah dasar seharusnya mulai dilatih berpikir tingkat tinggi sesuai dengan usia mereka. Salah satu cara yang dapat dilakukan guru adalah memberikan soal-soal yang menuntut peserta didik berpikir yang tidak hanya sekedar mengingat, tetapi juga menganalisis, mengevaluasi bahkan mencipta. Hal ini sesuai dengan pendapat yang dikemukakan oleh McNeill, Gosper, dan Xu (2012: 285) yaitu:

“untuk menilai kemampuan *HOT* peserta didik dibutuhkan sebuah instrumen yang melibatkan kemampuan berpikir kritis, pemecahan masalah dan kreatifitas yang dapat menantang peserta didik, sehingga dibutuhkan instrumen penilaian tertentu yang disusun berdasarkan kompetensi yang terkait dalam pembelajaran”.

Pentingnya *HOT* dimiliki peserta didik karena dapat membekali peserta didik agar siap menghadapi kehidupan yang jauh lebih kompleks suatu saat nanti. Hal ini sejalan dengan pemikiran Johnson (2014: 183) bahwa:

“berpikir kreatif dan kritis memungkinkan peserta didik untuk mempelajari masalah secara sistematis, menghadapi berjuta tantangan dengan cara yang terorganisasi, merumuskan pertanyaan inovatif dan merancang solusi orisinal.”

Newman dan Wehlage dalam Winarni (2018: 261) menyatakan bahwa “Dengan *HOT* siswa akan belajar lebih mendalam, siswa akan memahami konsep lebih baik. Dengan *HOT* siswa dapat membedakan ide dengan jelas, berargumen dengan baik, mampu memecahkan masalah, mampu mengkonstruksi penjelasan, mampu berhipotesis, dan memahami hal-hal kompleks menjadi lebih jelas”.

Berdasarkan penjelasan yang telah dipaparkan, maka penelitian ini akan mengembangkan instrumen penilaian berbasis *HOT*, hal ini merujuk pada hasil penelitian Sofiah, Susanto, dan Setiawani (2015: 7) mengatakan bahwa “pengembangan paket tes yang levelnya termasuk dalam katagori berpikir tingkat tinggi masih sangat minim”.

Metode

Penelitian ini merupakan penelitian pengembangan. Adapun produk yang dikembangkan adalah instrumen penilaian berbasis *HOT* pada pelajaran Matematika. Model pengembangan merujuk kepada model Borg dan Gall dalam Sugiyono (2016: 35) yang mengemukakan sepuluh langkah *R & D*. Namun pada penelitian ini diadaptasi enam langkah pengembangan mengingat keterbatasan waktu, biaya dan tenaga. Prosedur penelitian mencakup kegiatan: (1) penelitian dan pengumpulan informasi, (2) perencanaan, (3) pengembangan produk awal, (4) uji coba terbatas, (5) revisi produk awal, dan (6) uji coba lapangan.

Pada penelitian ini, subjek uji coba terbatas adalah siswa kelas IV SD Negeri 5 Megang Sakti. Sedangkan uji coba lapangan dilaksanakan di tiga sekolah, yaitu SD Negeri 1 Megang Sakti, SD Negeri 6 Megang Sakti dan SD Negeri 1 Wonosari Kecamatan Megang Sakti Kabupaten Musi Rawas Provinsi Sumatera Selatan. Objek penelitian adalah instrumen penilaian berbasis *HOT* Matematika, yang terdiri dari 15 soal pilihan ganda dan 5 soal uraian pada materi bangun datar.

1. Data Penelitian

Data pada penelitian ini meliputi data kualitatif dan data kuantitatif. Data kualitatif diperoleh dari lembar validasi ahli produk awal instrumen berbasis *HOT*. Sedangkan data kuantitatif diperoleh dari uji coba produk instrumen *HOT*.

2. Instrumen Penelitian

Instrumen untuk mengukur kevalidan secara teoritis digunakan lembar validasi yang dianalisis secara kualitatif sedangkan instrumen untuk mengukur validitas empiris dan reliabilitas digunakan instrumen penilaian yaitu soal-soal berbasis *HOT* yang di uji cobakan kepada peserta didik. Data-data tersebut bertujuan untuk memberikan gambaran kualitas produk yang akan dikembangkan.

3. Teknik Pengumpulan Data

Teknik pengumpulan data yang digunakan dalam penelitian ini meliputi lembar validasi dan instrumen penilaian berbasis *HOT*.

a. Lembar Validasi

Lembar validasi yang digunakan adalah lembar validasi ahli, untuk mengukur kevalidan instrumen tes termasuk kisi-kisi instrumen tes dengan meminta bantuan beberapa ahli, yaitu ahli Matematika, ahli Bahasa serta dua orang praktisi pendidikan. Untuk memvalidasi instrumen secara teoritis, b. Instrumen penilaian

Instrumen penilaian dalam penelitian ini terdiri dari 15 butir soal pilihan ganda dan 5 soal uraian. Setiap butir soal memuat level menganalisis (C4), mengevaluasi (C5), dan mencipta (C6).

Teknik analisis data yang digunakan ada dua yaitu analisis kualitatif dan analisis kuantitatif. Analisis kualitatif instrumen penilaian berbasis *HOT* diperoleh dari hasil lembar validasi ahli. Data berupa nilai tiap butir soal hasil penilaian ahli dianalisis menggunakan Aiken V. Menurut Retnawati (2018: 18) indeks Aiken V merupakan indeks kesepakatan ahli terhadap kesesuaian butir dengan indikator yang ingin diukur.

Selanjutnya menurut Azwar (2012: 113) jika saat melakukan analisis butir soal dengan indeks Aiken V diperoleh angka pada rentang 0,4 – 0,8, dengan

interpretasi sedang, maka item tersebut memiliki validitas yang baik dan sudah mendukung validitas tes secara keseluruhan.

Adapun konsistensi antar validator untuk melihat kekonsistenan validator dalam memberikan penilaian terhadap rancangan produk yang dikembangkan. Interpretasi konsistensi antar validator yang dikemukakan oleh McHugh (2012: 279) dapat dilihat pada Tabel 1.1.

Tabel 1.1. Interpretasi of *Cohen's Kappa*

| Nilai Kappa | Level Kesepakatan | % Data yang Reliabel |
|----------------|-------------------|----------------------|
| 0 – .02 | Tidak Ada | 0 – 4% |
| .21 – .39 | Kurang | 4 – 15% |
| .40 – .59 | Lemah | 15 – 35% |
| .60 – .79 | Sedang | 35 – 63% |
| .80 – .90 | Kuat | 64 – 81% |
| Lebih dari .90 | Sangat Kuat | 82 – 100% |

Sumber: McHugh. *Biochemia Medica* (2012: 276)

Berdasarkan Tabel 1.1, jika diperoleh nilai kappa kurang dari 0,60, maka konsistensi antar validator dalam memberikan penilaian berbeda. Jika nilai kappa antara 0,60 – 0,79, maka konsistensi antar validator dalam memberikan penilaian cukup baik. Sedangkan jika nilai kappa di atas 0,80, maka konsistensi antar validator dalam memberikan penilaian sangat baik.

Analisis kuantitatif instrumen penilaian berbasis *HOT* dianalisis menggunakan bantuan *software SPSS* versi 24 dan *microsoft excel*. Analisis butir soal digunakan untuk mengetahui karakteristik butir soal yang dikembangkan, untuk soal pilihan ganda meliputi validitas, reliabilitas, tingkat kesukaran, daya pembeda dan efektifitas option pengecoh. Sedangkan untuk soal uraian akan diperoleh validitas, reliabilitas, tingkat kesukaran, dan daya pembeda.

Hasil

1. Hasil Penelitian dan Pengembangan

a. Hasil Penelitian Tahap Penelitian dan Pengumpulan Informasi

Tahap penelitian dan pengumpulan informasi meliputi studi literatur dan penelitian lapangan. Studi literatur bertujuan untuk memperoleh data sebagai landasan teoritis untuk memperkuat argumentasi produk yang hendak dikembangkan. Sedangkan penelitian lapangan bertujuan untuk melakukan analisis kebutuhan dan mendapatkan data yang berkaitan dengan pengembangan instrumen penilaian berbasis *HOT* pelajaran Matematika.

1) Hasil Studi Literatur

Berdasarkan studi literatur atau kajian teoritik yang dilakukan, diperoleh informasi bahwa *HOT* penting dimiliki oleh peserta didik, dan kurikulum 2013 juga menuntut soal-soal tes yang mengarah pada keterampilan berpikir tingkat tinggi.

2) Hasil Penelitian Lapangan

Berdasarkan data angket identifikasi kebutuhan dari guru SD di Kabupaten Musi Rawas, mayoritas guru mengatakan bahwa mereka belum begitu memahami bagaimana merancang instrumen penilaian berbasis keterampilan berpikir tingkat tinggi.

Berdasarkan data dari angket peserta didik, diperoleh informasi: Peserta didik sangat jarang mengerjakan soal-soal keterampilan berpikir tingkat tinggi; dan peserta didik setuju dikembangkan soal-soal keterampilan berpikir tingkat tinggi.

b. Hasil Penelitian Tahap Perencanaan

Pada tahap perencanaan kegiatan yang peneliti lakukan meliputi: menganalisis kompetensi dasar, dan merancang instrumen penilaian.

1) Hasil Analisis Kompetensi Dasar

Hasil analisis kompetensi dasar yang dilakukan di semester 2, pada penelitian ini dipilih kompetensi dasar 3.9 Menjelaskan dan menentukan keliling dan luas persegi, persegipanjang, dan segitiga serta hubungan pangkat dua dengan akar pangkat dua. Berdasarkan hasil analisis kompetensi dasar, maka ditetapkan materi yang akan diteliti yaitu materi bangun datar. Pemilihan materi bangun datar berkaitan dengan tuntutan kompetensi dasar yang menghendaki peserta didik agar dapat berpikir *HOT*.

2) Hasil Perencanaan Instrumen Penilaian Berbasis *HOT*

Hasil dari perancangan instrumen penilaian berbasis *HOT*, yaitu 25 butir soal pilihan ganda dan 10 soal uraian. Di mana soal-soal yang dibuat berpedoman pada langkah-langkah penyusunan soal dan kaidah penulisan soal secara umum dengan menambahkan karakteristik soal *HOT*.

c. Hasil Penelitian Tahap Pengembangan Produk Awal

1) Hasil Pengujian Internal

Tahap selanjutnya pada penelitian ini adalah pengujian internal, yaitu validasi ahli secara teoritis meliputi aspek materi, konstruksi dan Bahasa. Pengujian internal dilakukan dengan cara memberikan lembar validasi dan rancangan instrumen (kisi-kisi soal, instrumen soal, pedoman penskoran) kepada masing-masing ahli. Validasi ahli bertujuan untuk mendapatkan masukan, saran dan sekaligus

Pengembangan Instrumen Penilaian berbasis Higher Order Thinking (Hot) Pelajaran Matematika Kelas IV Sekolah Dasar

memberikan penilaian pada masing-masing item butir soal menggunakan *skala likert*, yaitu memberikan angka 1 – 5.

Hasil dari penilaian yang diberikan validator dikumpulkan, untuk selanjutnya dilakukan analisis dengan menggunakan formula Aiken V.

(a) Hasil Validasi Ahli Soal Pilihan Ganda

Hasil analisis validasi teoritis dari lima orang ahli sebagai validator pada soal pilihan ganda dapat dilihat pada Tabel 1.2.

Tabel 1.2. Hasil Analisis Validasi Ahli Soal Pilihan Ganda

| Aspek yang Dinilai | Interpretasi Validitas Butir Soal | Jumlah item Soal | Rata-rata validitas Butir Soal |
|---------------------|--|------------------------|---|
| 1. Aspek Materi | Sedang soal nomor: 1, 2, 3, 4, 7, 8, 9, 10, 11, 12, 14, 17, 19, 20, 21, 22, 23, 25 Baik soal nomor: 5, 6, 13, 15, 16, 18, 24 | Sedang: 18 Baik: 7 | 0,74 Tertinggi 0,85 Terendah 0,50 |
| 1. Aspek Konstruksi | Sedang soal nomor: 1, 2, 3, 5, 7, 8, 10, 11, 13, 17, 18, 20, 21, 22, 23. Baik soal nomor: 4, 6, 9, 12, 14, 15, 16, 19, 24, 25 | Sedang: 15 Baik: 10 | 0,75 Tertinggi 0,85 Terendah 0,50 |
| 2. Aspek Bahasa | Sedang Nomor Soal: 1, 2, 3, 4, 6, 7, 11, 12, 15, 17, 18, 19, 20, 21, 22, 23. Baik nomor soal: 5, 8, 9, 10, 13, 14, 16, 24, 25 | Sedang: 16 Baik: 9 | 0,75 Tertinggi 0,90 Terendah 0,50 |
| Rata-rata | | | 0,75 |

Berdasarkan Tabel 1.2, pada aspek materi rata-rata validitas yaitu 0,74 kategori “sedang”. Pada aspek konstruksi rata-rata validitas yaitu 0,75 kategori “sedang”. Sedangkan pada aspek bahasa rata-rata validitas yaitu 0,75 dengan kategori “sedang”.

Hasil perhitungan konsistensi antar validator dapat dilihat pada Tabel 1.3.

Tabel. 1.3. Konsistensi Antar Validator
Soal Pilihan Ganda

| Aspek Yang dinilai | Reliabilitas Antar Validator Butir Soal | Jumlah Butir Soal | Rata-rata reliabilitas Antar Validator |
|--------------------|---|--|---|
| Aspek Materi | Sangat baik: 10, 19 Baik: 1, 4, 5, 9, 11, 12, 14, 17, 20, 22, 23, 25, Sedang: 2, 6, 7, 13, 15, 16, 18, 21, 24 Rendah: 3, 8 | Sangat Baik: 2 Baik: 12 Sedang: 9 Rendah: 2 | 0,71 Tertinggi: 1,00 Terendah: 0,40 |

| | | | |
|------------------|---|--|---|
| Aspek Konstruksi | Sangat baik: 10, 11, 18 Baik: 3, 5, 7, 8, 17, 21, 22 Sedang: 1, 2, 4, 6, 9, 12, 13, 14, 15, 16, 19, 20, 23, 24, 25 Rendah: - | Sangat Baik: 3 Baik: 7 Sedang: 15 Rendah: - | 0,70 Tertinggi: 1,00 Terendah: 0,60 |
| Aspek Bahasa | Sangat baik: 7, 19, 23 Baik: 3, 4, 17, 22 Sedang: 1, 2, 5, 6, 8, 9, 10, 11, 12, 13, 14, 16, 18, 20, 24, 25. Rendah: 3, 8 | Sangat Baik: 3 Baik: 4 Sedang: 16 Rendah: 2 | 0,66 Tertinggi: 1,00 Terendah: 0,40 |
| Rata-rata | | | 0,69 |

Berdasarkan Tabel 1.3, diperoleh rata-rata konsistensi antar validator dari aspek materi, aspek konstruksi, dan aspek bahasa adalah 0,69 dengan interpretasi “sedang”. Artinya seluruh validator cukup konsisten memberikan penilaian terhadap butir soal pilihan ganda pada ketiga aspek tersebut.

(b) Hasil Validasi ahli soal Uraian

Hasil analisis validasi teoritis dari lima orang ahli sebagai validator pada soal uraian dapat dilihat pada Tabel 1.4.

Tabel 1.4. Hasil Analisis Validasi Ahli Soal Uraian

| Aspek yang dinilai | Validitas Butir Soal | Jumlah Validitas Butir Soal | Rata-rata Validitas Butir Soal |
|---------------------|---|-----------------------------|---|
| 1. Aspek Materi | Sedang, nomor soal: 1, 2, 3, 6, 7, 10 Baik, nomor soal: 4, 5, 8, 9 | Sedang: 6 Baik: 4 | 0,73 Tertinggi: 0.90 Terendah: 0,55 |
| 2. Aspek konstruksi | Sedang nomor soal: 1, 2, 3, 6, 7, 10 Baik nomor soal: 4, 5, 8, 9 | Sedang: 6 Baik: 4 | 0,74 Tertinggi: 0.90 Terendah: 0,60 |
| 3. Aspek bahasa | Sedang, nomor soal: 1, 2, 3, 6, 7, 10 Baik, nomor soal: 4, 5, 8, 9 | Sedang: 6 Baik: 4 | 0,75 Tertinggi: 0.90 Terendah: 0,55 |
| Rata-rata | | | 0,74 |

Berdasarkan Tabel 1.4, rata-rata validitas teoritis soal uraian pada aspek materi, aspek konstruksi dan aspek bahasa adalah 0,74 dengan interpretasi “sedang”. Konsistensi antar validator pada butir soal uraian dapat dilihat pada Tabel 1.5.

Tabel. 1.5. Konsistensi Antar Validator
Soal Uraian

| Aspek yang dinilai | Reliabilitas Interrater Nomor Butir Soal | Jumlah Butir Soal | Rata-rata reliabilitas Antar Validator |
|--------------------|---|----------------------|---|
| Aspek Materi | Baik: 1, 8 Sedang: 2, 3, 4, 5, 6, 7, 9, 10 | Baik: 2 Sedang: 8 | 0,64 Tertinggi: 0,80 Terendah: 0,60 |
| Aspek Konstruksi | Baik: 6 Sedang: 1, 2, 4, 5, 7, 8, 9, 10 | Baik: 1 Sedang: 9 | 0,62 Tertinggi: 0,80 Terendah: 0,60 |
| Aspek Bahasa | Baik: 10 Sedang: 1, 2, 3, 4, 5, 6, 7, 8, 9 | Baik: 1 Sedang: 9 | 0,62 Tertinggi: 0,80 Terendah: 0,60 |
| Rata-rata | | | 0,63 |

Berdasarkan Tabel 1.5, diperoleh rata-rata konsistensi antar validator pada aspek materi, konstruksi, dan bahasa adalah 0,62 “sedang”. Artinya seluruh validator cukup konsisten memberikan penilaian terhadap soal uraian pada ketiga aspek tersebut.

2) Hasil Penelitian Tahap Revisi Desain

Adapun hasil revisi desain yang disarankan oleh validator, karena kesalahan penulisan, yaitu perbaikan penulisan yang kurang tepat dan kesalahan penulisan angka.

3) Hasil Penelitian Tahap Rancangan Desain Pengembangan Produk Awal

Hasil pengembangan produk awal berupa instrumen soal pilihan ganda dan soal uraian sebagai berikut: instrumen penilaian berbasis *HOT* soal pilihan ganda meliputi: instrumen soal sebanyak 15 butir soal, dan soal uraian sebanyak 5 butir soal, pedoman penskoran dan penilaian.

d. Hasil Penelitian Tahap Uji Coba Terbatas

Uji coba terbatas dilaksanakan di kelas IV SD Negeri 5 Megang Sakti pada tanggal 18 April 2018 dengan jumlah peserta tes sebanyak 20 orang terdiri dari 12 laki-laki dan 8 perempuan. Karakteristik butir soal instrumen penilaian berbasis *HOT* untuk soal pilihan ganda dapat dilihat pada Tabel 1.6.

Tabel 1.6. Validitas dan Reliabilitas Butir Soal Pilihan Ganda Produk Awal

| Kriteria Validitas | Nomor Butir Soal | Jumlah | Persentase (%) |
|---------------------------|--|--------|----------------|
| $0,70 \leq r_{xy} < 0,90$ | - | 0 | 0 |
| $0,40 \leq r_{xy} < 0,70$ | 1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15 | 14 | 93,33 |
| $0,20 \leq r_{xy} < 0,40$ | - | 0 | 0 |
| $r_{xy} < 0,20$ | 8 | 1 | 6,67 |

Berdasarkan Tabel 1.6, dapat diketahui bahwa persentase validitas butir soal yaitu 93,33%, dan satu soal, yaitu nomor 8 yang tidak valid dengan persentase 6,67%, serta daya pembeda soal 0,69 dengan kategori baik. Hasil analisis tingkat kesukaran butir soal pilihan ganda produk awal dapat dilihat pada Tabel 1.7.

Tabel 1.7. Tingkat Kesukaran Soal
Pilihan Ganda Produk Awal

| Kriteria | Nomor Butir Soal | Jumlah Soal | Persentase (%) |
|--------------------------------------|--------------------------|-------------|----------------|
| $TK < 0,30$ (sukar) | 3, 4, 5, 7, 8, 9, 11, 14 | 8 | 53,33 |
| $0,30 \leq TK \leq 0,70$ (sedang) | 1, 2, 6, 10, 12, 13, 15 | 7 | 46,67 |

Berdasarkan Tabel 1.7, dapat diketahui bahwa tingkat kesukaran butir soal berkisar pada butir soal sukar sebanyak 8 butir soal (53,33%), butir soal kategori sedang sebanyak 7 butir soal (46,675). Hasil analisis daya pembeda butir soal pilihan ganda produk awal dapat dilihat pada Tabel 1.8.

Tabel 1.8. Daya Pembeda Soal Pilihan
Ganda Produk Awal Instrumen

| Kriteria | Nomor Butir Soal | Jumlah Soal | Persentase (%) |
|---|------------------------|-------------|----------------|
| $DP \geq 0,40$ (baik) | 1, 4, 6, 7, 10, 12, 15 | 7 | 46,67 |
| $0,30 \leq DP < 0,40$ (diterima tanpa revisi) | 5, 9, 11, 13 | 4 | 26,67 |
| $0,20 \leq DP < 0,30$ (diterima dengan revisi) | 2, 3, 14 | 3 | 20 |
| $DP < 0,20$ (diganti/benar-benar revisi) | 8 | 1 | 6,67 |

Berdasarkan Tabel 1.8, dapat diketahui bahwa daya pembeda butir soal berkisar pada kategori baik sebanyak 7 soal (46,67%), diterima tanpa revisi sebanyak 4 butir soal (26,67%), diterima dengan revisi sebanyak 3 butir soal (20%), dan diganti 1 butir soal (6,67%). Penyebaran option pengecoh butir soal pilihan ganda dapat dilihat pada Tabel 1.9.

Tabel 1.9. Hasil Analisis Efektifitas Distraktor Soal Pilihan Ganda
Uji Coba Terbatas

| Kategori | Nomor Butir Soal | Jumlah | Persentase (%) |
|---|--------------------------------------|--------|----------------|
| $\bar{X}_{dis} < \bar{X}$ (distraktor baik) | 2, 6, 10, 15 | 4 | 26,7 |
| $\bar{X}_{dis} \geq \bar{X}$ (distraktor buruk) | 1, 3, 4, 5, 7, 8, 9, 11, 12, 13, 14, | 11 | 73,3 |

Pada Tabel 1.9, terlihat bahwa distraktor soal pilihan ganda berfungsi kurang baik sebanyak 11 butir soal (73,3%), dan berfungsi baik 4 butir soal (26,7%).

Hasil analisis instrumen penilaian berbasis *HOT* uji coba terbatas soal uraian dapat dilihat pada Tabel 1.10.

Tabel 1.10. Validitas Butir Soal
Uraian Produk Awal

| Kriteria Validitas | Nomor Butir Soal | Jumlah | Interpretasi | Persentase (%) |
|---------------------------|------------------|--------|--------------|----------------|
| $0,70 \leq r_{xy} < 0,90$ | 1, 2, 5 | 3 | Baik | 60,00 |
| $0,40 \leq r_{xy} < 0,70$ | 3, 4 | 2 | Sedang | 40,00 |
| $0,20 \leq r_{xy} < 0,40$ | - | 0 | 0 | 0 |
| $r_{xy} < 0,20$ | - | - | 0 | 0 |

Berdasarkan Tabel 1.10, dapat diketahui bahwa persentase validitas butir soal baik yaitu 60% sebanyak 3 butir soal, dan cukup baik 40% sebanyak 2 butir soal, serta reliabilitas butir soal uraian uji coba terbatas adalah 0,89 dengan interpretasi baik.

Hasil analisis daya pembeda butir soal pilihan ganda produk awal dapat dilihat pada Tabel 1.11.

Tabel 1.11. Tingkat Kesukaran
Soal Uraian Produk Awal

| Kriteria | Nomor Butir Soal | Jumlah Soal | Persentase (%) |
|-------------------|------------------|-------------|----------------|
| TK < 0,30 (sukar) | - | 0 | |

| | | | |
|--------------------------------------|------------------|---|-----|
| $0,30 \leq TK \leq 0,70$ (sedang) | 1, 2, 3, 4, 5 | 5 | 100 |
| $TK > 0,70$ (mudah) | - | 0 | |

Berdasarkan Tabel 1.11, dapat diketahui bahwa tingkat kesukaran butir soal berkisar pada butir soal sedang sebanyak 5 butir soal dengan persentase (100%). Hasil analisis daya pembeda butir soal uraian dapat dilihat pada Tabel 1.12.

Tabel 1.12. Daya Pembeda Soal Uraian
Produk Awal

| Kriteria | Nomor Butir Soal | Jumlah Soal | Persentase (%) |
|--|------------------|-------------|----------------|
| $DP \geq 0,40$ (baik) | | | |
| $0,30 \leq DP < 0,40$ (diterima tanpa revisi) | 1 | 1 | 20 |
| $0,20 \leq DP < 0,30$ (diterima dengan revisi) | 2, 3, 4, 5 | 4 | 80 |
| $DP < 0,20$ (diganti/benar-benar revisi) | - | 0 | 0 |

Berdasarkan Tabel 1.12, dapat diketahui bahwa daya pembeda butir soal berkisar pada kategori diterima tanpa revisi sebanyak 1 butir soal (20%), diterima dengan revisi sebanyak 4 butir soal (80%).

e. Hasil Penelitian Tahap Revisi Produk Awal

Berdasarkan hasil analisis uji coba terbatas yang dilaksanakan di SD Negeri 5 Megang Sakti, maka perlu dilakukan revisi produk awal sebelum dilaksanakan uji coba lapangan. Pada soal pilihan ganda dari 15 soal yang di ujikan terdapat 14 soal yang dapat dipakai dan 1 soal yang perlu diganti, yaitu soal nomor 8, dan beberapa soal diperbaiki redaksinya. Sementara untuk soal uraian semua soal dapat langsung digunakan.

f. Hasil Penelitian Tahap Uji Coba Lapangan

1) Hasil Uji Coba Lapangan Instrumen Berbasis *HOT* Pilihan Ganda

Hasil analisis jawaban peserta tes uji coba lapangan di tiga SD untuk menentukan validitas empiris, reliabilitas, tingkat kesukaran, daya pembeda, dan efektifitas option pengecoh soal pilihan ganda. Adapun validitas dan reliabilitas butir soal dapat dilihat pada Tabel 1.13.

Tabel 1.13. Validitas dan Reliabilitas Butir Soal Pilihan Ganda Uji Coba Lapangan

| Kriteria Validitas | Validitas Butir Soal | Jumlah | (%) |
|---------------------------|---|--------|-----|
| $0,70 \leq r_{xy} < 0,90$ | - | 0 | 0 |
| $0,40 \leq r_{xy} < 0,70$ | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, | 15 | 100 |
| $0,20 \leq r_{xy} < 0,40$ | - | 0 | 0 |
| $r_{xy} < 0,20$ | - | - | 0 |
| Reliabilitas: 0,73 (baik) | | | |

Berdasarkan Tabel 1.13, dapat diketahui bahwa persentase validitas butir soal cukup baik sebanyak 15 butir soal (100%), dan reliabilitas soal 0,73 dengan kategori baik. Hasil analisis tingkat kesukaran butir soal pilihan ganda produk hasil revisi dapat dilihat pada Tabel 1.14.

Tabel 1.14. Tingkat Kesukaran Soal Pilihan Ganda Produk Hasil Revisi

| Kriteria | Nomor Butir Soal | Jumlah Soal | Persentase (%) |
|--------------------------------------|--|-------------|----------------|
| TK < 0,30 (sukar) | 9, 11, 13 | 3 | 20 |
| $0,30 \leq TK \leq 0,70$ (sedang) | 1, 2, 3, 4, 5, 6, 7, 8, 10, 12, 14, 15 | 12 | 80 |
| TK > 0,70 (mudah) | - | 0 | 0 |

Berdasarkan Tabel 1.14, dapat diketahui bahwa tingkat kesukaran butir soal berkisar pada butir soal sukar sebanyak 3 butir soal (20%) dan tingkat kesukaran kategori sedang sebanyak 12 butir soal (80%). Artinya secara umum rata-rata tingkat kesukaran butir soal berada pada kategori sedang. Hasil analisis daya pembeda butir soal pilihan ganda dapat dilihat pada tabel 1.15.

Tabel 1.15. Daya Pembeda Soal Pilihan Ganda Uji Coba Lapangan

| Kriteria | Nomor Butir Soal | Jumlah Soal | Persentase (%) |
|--|--|-------------|----------------|
| DP \geq 0,40 (baik) | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 | 15 | 100 |
| $0,30 \leq DP < 0,40$ (diterima tanpa revisi) | - | 0 | 0 |
| $0,20 \leq DP < 0,30$ (diterima dengan revisi) | - | 0 | 0 |
| DP < 0,20 (diganti/benar-benar revisi) | - | 0 | 0 |

Berdasarkan Tabel 1.15, dapat diketahui bahwa daya pembeda butir soal berkisar pada kriteria baik sebanyak 15 butir soal (100%). Hasil analisis penyebaran option pengecoh butir soal pilihan ganda dapat dilihat pada Tabel 1.16.

Tabel 1.16. Hasil Analisis Efektifitas Distraktor Soal Pilihan Ganda Uji Coba Lapangan

| Kategori | Nomor Butir Soal | Jumlah | Persentase (%) |
|---|------------------------------|--------|----------------|
| $\bar{X}_{dis} < \bar{X}$ (distraktor baik) | 1, 2, 4, 5, 6, 7, 11, 12, 15 | 9 | 60 |
| $\bar{X}_{dis} \geq \bar{X}$ (distraktor buruk) | 3, 8, 9, 10, 13, 14 | 6 | 40 |

Pada tabel 1.16, terlihat bahwa distraktor soal pilihan ganda berfungsi baik 9 butir soal (60%) dan berfungsi kurang baik sebanyak 6 butir soal (40%). Hasil analisis instrumen penilaian berbasis *HOT* uji coba lapangan soal uraian dapat dilihat pada Tabel 1.17.

Tabel 1.17. Validitas dan Reliabilitas Butir Soal Uraian Uji Coba Lapangan

| Kriteria Validitas | Validitas Butir Soal | Jumlah | Interpretasi | Persentase (%) |
|--------------------------------------|----------------------|--------|--------------|----------------|
| $0,70 \leq r_{xy} < 0,90$ | 1, 2, 3, 4, 5 | 5 | Baik | 100 |
| $0,40 \leq r_{xy} < 0,70$ | - | 0 | 0 | 0 |
| $0,20 \leq r_{xy} < 0,40$ | - | 0 | 0 | 0 |
| $r_{xy} < 0,20$ | - | - | 0 | 0 |
| Reliabilitas butir soal 0,81 (Baik) | | | | |

Berdasarkan Tabel 1.17, dapat diketahui bahwa persentase validitas butir soal baik yaitu 100%, artinya seluruh soal uraian memiliki validitas butir soal baik, dan reliabilitas soal 0,81 dengan kriteria baik.

Hasil analisis tingkat kesukaran butir soal uraian produk hasil revisi dapat dilihat pada Tabel 1.18.

Tabel 1.18. Tingkat Kesukaran Soal Uraian Produk Hasil Revisi

| Kriteria | Nomor Butir Soal | Jumlah Soal | Persentase (%) |
|-----------------------------------|------------------|-------------|----------------|
| $TK < 0,30$ (sukar) | 2 | 1 | 20 |
| $0,30 \leq TK \leq 0,70$ (sedang) | 1, 3, 4, 5 | 4 | 80 |
| $TK > 0,70$ (mudah) | - | 0 | 0 |

Pengembangan Instrumen Penilaian berbasis Higher Order Thinking (Hot) Pelajaran Matematika Kelas IV Sekolah Dasar

Berdasarkan Tabel 1.18, dapat diketahui bahwa tingkat kesukaran butir soal berkisar pada butir soal sedang sebanyak 4 butir soal (80%), dan tingkat kesukaran soal kriteria sukar 1 soal (20%). Artinya rata-rata tingkat kesukaran butir soal berada pada kategori “sedang”.

Daya pembeda butir soal uraian uji coba lapangan dapat dilihat pada Tabel 1.19.

Tabel 1.19. Daya Pembeda Soal Uraian Produk Hasil Revisi

| Kriteria | Nomor Butir Soal | Jumlah Soal | Persentase (%) |
|--|------------------|-------------|----------------|
| $DP \geq 0,40$ (baik) | - | 0 | 0 |
| $0,30 \leq DP < 0,40$ (diterima tanpa revisi) | 1, 2, 3, 4, 5 | 5 | 100 |
| $0,20 \leq DP < 0,30$ (diterima dengan revisi) | - | 0 | 0 |
| $DP < 0,20$ (diganti/benar-benar revisi) | - | 0 | 0 |

Berdasarkan Tabel 1.19, dapat diketahui bahwa daya pembeda butir soal berkisar pada kategori diterima tanpa revisi sebanyak 5 butir soal (100%).

Pembahasan

1. Pengembangan Instrumen Penilaian *HOT* yang Valid dan Reliabel

Penelitian ini mengembangkan instrumen penilaian berbasis *HOT* pelajaran Matematika kelas IV sekolah dasar yang valid dan reliabel, adapun instrumen yang dikembangkan adalah soal pilihan ganda dan soal uraian berdasarkan revisi taksonomi Bloom pada level kognitif menganalisis (C4), mengevaluasi (C5), dan mengkreasi (C6) pada materi bangun datar.

Berdasarkan teori dari Kementerian Pendidikan dan Kebudayaan (2016: 17), untuk memperoleh soal *HOT* diperlukan langkah-langkah dalam menyusun soal *HOT*, yaitu: (a) Menganalisis Kompetensi Dasar (KD) yang dapat dibuat soal-soal *HOT*; (b) Menyusun kisi-kisi soal *HOT*; (c) Memilih stimulus yang *HOT*, menarik dan kontekstual; (d) Menulis butir pertanyaan sesuai dengan kisi-kisi soal; dan (e) Membuat pedoman penskoran atau kunci jawaban.

Penelitian ini telah mengikuti semua langkah-langkah penyusunan instrumen penilaian *HOT* pada setiap tahap pengembangan, yaitu kegiatan menganalisis kompetensi dasar yang dapat dibuat soal *HOT*, membuat kisi-kisi soal *HOT*, mencari dan memilih stimulus yang menarik dan kontekstual, merancang instrumen soal,

dan membuat rubrik penskoran dan penilaian. Pengembangan instrumen penilaian berbasis *HOT* pelajaran Matematika kelas IV sekolah dasar.

Pada soal pilihan ganda untuk level menganalisis (C4) terdapat 8 butir soal, indikator yang dikembangkan dari 8 butir soal tersebut adalah menuntut kemampuan peserta didik untuk dapat merinci, menafsirkan, dan menentukan, sehingga mengembangkan kemampuan peserta didik untuk dapat memahami konsep dengan baik dan mampu memecahkan masalah.

Pada level mengevaluasi (C5) terdapat 4 butir soal, indikator yang dikembangkan dari 4 butir soal tersebut adalah menuntut kemampuan peserta didik untuk dapat menafsirkan dan menyusun, sehingga mengembangkan kemampuan peserta didik untuk mampu memecahkan masalah dan mampu berhipotesis.

Pada level mengkreasi (C6) terdapat 3 butir soal, indikator yang dikembangkan dari 3 butir soal tersebut adalah menuntut kemampuan peserta didik dapat menentukan, merancang dan membuat, sehingga mengembangkan kemampuan peserta didik untuk mampu memecahkan masalah.

Sedangkan untuk soal uraian pada level menganalisis (C4) terdapat 2 butir soal, indikator yang dikembangkan dari 2 butir soal tersebut adalah menuntut kemampuan peserta didik untuk dapat merinci, sehingga mengembangkan kemampuan peserta didik untuk mampu memecahkan masalah.

Pada level mengevaluasi (C5) terdapat 2 butir soal, indikator yang dikembangkan dari 2 butir soal tersebut adalah menuntut kemampuan peserta didik untuk dapat memprediksi dan membuat kesimpulan, sehingga mengembangkan kemampuan peserta didik untuk dapat memahami konsep dengan baik dan mampu memecahkan masalah. Pada level mengevaluasi (C6) terdapat 1 butir soal, indikator yang dikembangkan dari 1 butir soal tersebut adalah menuntut kemampuan peserta didik untuk dapat merancang, sehingga mengembangkan kemampuan peserta didik untuk mampu memecahkan masalah.

Hasil penelitian tersebut sudah mendukung beberapa teori Newman dan Wehlage dalam Winarni (2018: 216) menyatakan bahwa “dengan *HOT* peserta didik akan memahami konsep dengan baik, mampu memecahkan masalah, dan mampu berhipotesis”.

Secara umum berdasarkan hasil penelitian yang uraikan sebelumnya, instrumen penilaian berbasis *HOT* pada pelajaran Matematika yang dikembangkan berupa 15 soal pilihan ganda dan 5 soal uraian dari aspek materi, konstruksi dan bahasa dinyatakan valid secara teoritis, empiris dan reliabel, maka layak digunakan. Menurut Scarvia dalam Arikunto (2012:101) menyatakan bahwa “persyaratan bagi

tes yaitu, validitas dan reliabilitas penting, karena keduanya saling menyokong sebuah tes yang baik”.

Penelitian ini selaras dengan hasil penelitian Budiman dan Jailani (2014), yaitu berupa soal tes *HOTS* yang terdiri dari 24 butir soal pilihan ganda dan 19 butir soal uraian dari aspek materi, konstruksi, dan bahasa dinyatakan valid dan layak digunakan.

Penelitian ini juga mendukung hasil penelitian Sofiah, Susanto dan Setiawani (2015) di mana pada uji coba yang terakhir (uji coba *large group*) diperoleh reliabilitas dengan interpretasi sangat tinggi. Berdasarkan hasil validasi dan analisis uji coba, secara umum paket tes yang dikembangkan telah sesuai dengan level berpikir tingkat tinggi dan memenuhi kriteria tes yang baik yaitu valid dan reliabel.

Ucapan terima kasih diucapkan kepada seluruh pihak yang telah mendukung dan membantu penelitian ini, yaitu guru-guru kelas IV sekolah dasar, peserta didik kelas IV sekolah dasar dan Kepala Sekolah yang telah memberikan izin sehingga penelitian ini terlaksana dengan baik.

Kesimpulan

Produk akhir dalam penelitian ini adalah instrumen penilaian berbasis *HOT* pelajaran Matematika kelas IV materi bangun datar yang valid dan reliabel, terdiri dari 15 soal pilihan ganda dan 5 soal uraian. Pengembangan Instrumen penilaian berbasis *HOT* dikembangkan melalui enam langkah pengembangan, yaitu: (1) penelitian dan pengumpulan informasi, (2) perencanaan, (3) pengembangan produk awal, (4) uji coba terbatas, (5) revisi produk awal, dan (6) uji coba lapangan.

Kevalidan instrumen dibuktikan dari hasil validasi ahli yang menunjukkan bahwa instrumen layak dikembangkan berdasarkan telaah teoritis pada aspek materi, aspek konstruksi dan aspek bahasa. Secara empiris instrumen penilaian yang dikembangkan juga telah memenuhi kriteria valid dan reliabel, pada soal pilihan ganda daya pembeda soal baik, tingkat kesukaran sedang, distraktor sudah berfungsi cukup baik, dan untuk soal uraian daya pembeda soal baik, dan tingkat kesukaran soal sedang.

Saran

Berdasarkan hasil penelitian dan kesimpulan, soal pilihan ganda sebaiknya mengandung 40% soal *HOT*, 30% soal *MOT*, dan 30% soal *LOT*, sedangkan pada soal uraian sebaiknya mengandung 100% soal *HOT* agar kemampuan berpikir tingkat tinggi peserta didik dapat meningkat.

Untuk memperoleh instrumen yang baik:

- a. Kepada peneliti selanjutnya pada saat merancang instrumen harus berdasarkan kaidah-kaidah penulisan soal *HOT*, dan
- b. Untuk soal pilihan ganda, disarankan kepada peneliti selanjutnya untuk dapat merancang distraktor yang memenuhi cara pembuatan distraktor, agar distraktor berfungsi dengan baik, yaitu dipilih minimal 5% dari pengikut tes.

Referensi

- Arikunto, Suharsimi. 2012. *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Azwar, Saifuddin. 2012. *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Pelajar.
- Brookhart, Susan, M. 2010. *How to Assess Higher-Order Thinking Skills in Your Classroom*. Virginia USA: ASCD Alexandria.
- Budiman, A dan Jaelani. 2014. Pengembangan Instrumen Asesmen Higher Order Thinking Skills (*HOTS*) Pada Mata Pelajaran Matematika SMP Kelas VIII Semester I. *Jurnal Riset Pendidikan Matematika, Volume 1, Nomor 2, November 2014. Hal: 139-151*. Dunduh tanggal 29 Desember 2017.
- Johnson, Elaine, B. 2014. *Contextual Teaching & Learning*. Bandung: Kaifa.
- Kementerian Pendidikan dan Kebudayaan. 2016. *Panduan bagaimana pendidik dapat menulis soal yang berkriteria untuk berpikir tingkat tinggi*. Jakarta : Departemen Pendidikan dan Kebudayaan.
- Khan, W.B., dan Inamullah, H.M. 2011. A study of lower-order and higher-order questions at secondary level. Pakistan: Canadian Center of Science and Education. *Asian Social Science Vol. 7, No. 9; September 2011*
<http://dx.doi.org/10.5539/ass.v7n9p149>. Diunduh tanggal 8 Maret 2018.
- McHugh, Mary, L. 2012. Interrater Reliability: the Kappa Statistk. *Biochemia Medica, Vol.22, No. 3. p: 276-282*.
- McNeill, M., Gosper, M., dan Xu, J. 2012. Assessment Choices To Target Higher Order Learning Outcomes: The Power Of Academic Empowerment. *Research in Learning Technology, Vol.20*. Diunduh pada tanggal 27 Februari 2018.
- Partnership of 21st Century Skill*. 2011. Designed in Cooperation with the Nation's Arts Educators. www.P21.org. diunduh pada tanggal 27 Februari 2018.
- Peraturan Pemerintah Republik Indonesia Nomor 19. 2005. *Standar Nasional Pendidikan*. <https://kemenag.go.id>.
- Puspendik Kemdikbud. 2015. *TIMMS Info Grafik*.
<https://puspendik.kemdikbud.go.id> diunduh tanggal 9 Maret 2018.
- Ramos, J. L. S., Dolipas, B. B., dan Villamor, B.B. 2013. Higher Order Thinking Skillss and Academic Performance in Physics of College Students: A Regression Analysis. *International Journal of Innovative Interdisciplinary Research, Issue 4, p: 48-60*. Diunduh pada tanggal 27 Februari 2018.

**Pengembangan Instrumen Penilaian berbasis Higher Order Thinking (Hot) Pelajaran
Matematika Kelas IV Sekolah Dasar**

- Retnawati. H. 2016. *Validitas Reliabilitas dan Karakteristik Butir*. Yogyakarta: Parama Publishing.
- Sofiah, S. Susanto, dan Setiawani, S. 2015. Pengembangan Paket Tes Kemampuan Berpikir Tingkat Tinggi Matematika Berdasarkan Revisi Taksonomi Bloom Pada Siswa Kelas V SD. *Jurnal Vol. 1, No. 1, (2015), Halaman: 1-7*.
Dalam <http://repository.unej.ac.id/handle/123456789/64108> diunduh tanggal 13 Mei 2017.
- Sugiono. 2016. *Metode Penelitian dan Pengembangan*. Bandung: PT Remaja Rosdakarya.
- Winarni, Endang. W. 2018. *Pendekatan Ilmiah dalam Pembelajaran Kreatif dan Inovatif*. Bengkulu: FKIP UNIB