

Pengembangan instrumen tes hasil belajar matematika peserta didik sekolah dasar berorientasi pada berpikir tingkat tinggi

Sabina Ndiung ✉, Universitas katolik Indonesia Santu Paulus Ruteng

Mariana Jediut, Universitas katolik Indonesia Santu Paulus Ruteng

✉ sabinandiung@unikastpaulus.ac.id

Abstract: This study aims to develop test items on mathematics learning outcomes by looking level of difficulty items, validity and reliability, and discriminatory power item test index. This study involved 30 students of grade VI SDK Wae Mata, Lembor sub-district. The method used was development. The Instrument developed is oriented on HOTS with an essay test form. Determining level of difficulty test items used formula developed by Candiasa, validity test items used correlation product moment technique, reliability coefficient used Alpha Cronbach, and the discriminatory power item test index used Ferguson formula. Based on results, obtained: (1) difficulty level of items in moderate category; (2) 5 valid items and 2 items are revised; (3) reliability coefficient is 0,79, including in high reliability; and (4) all different power of item test index is good category. Thus, instrument developed is oriented on HOTS, which is reliable and useful as an instrument for collecting data.

Keywords: Development, HOTS, Mathematics learning outcome, Test instrument

Abstrak: Penelitian ini bertujuan mengembangkan instrumen tes hasil belajar matematika dengan melihat tingkat kesukaran butir soal, validitas dan reliabilitas, serta indeks daya beda butir. Penelitian ini melibatkan 30 peserta didik kelas VI SDK Wae Mata Kecamatan Lembor Kabupaten Manggarai Barat. Metode yang digunakan dalam penelitian ini adalah pengembangan. Instrumen yang dikembangkan berorientasi pada HOTS dengan bentuk tes uraian. Untuk menentukan tingkat kesukaran butir tes digunakan formula yang dikembangkan oleh Candiasa, validitas butir tes menggunakan formula Carl Pearson, koefisien reliabilitas menggunakan formula Alpha Cronbach, dan indeks daya beda butir menggunakan formula Ferguson. Berdasarkan hasil uji coba diperoleh (1) tingkat kesukaran butir pada kategori sedang; (2) terdapat 5 butir yang valid sedangkan 2 butir lainnya direvisi; (3) koefisien reliabilitasnya 0,79 pada kategori tinggi; dan (4) indeks daya beda butir semuanya berkategori baik. Dengan demikian instrumen yang dikembangkan yang berorientasi pada HOTS ini dapat diandalkan dan dapat digunakan sebagai instrumen pengumpulan data penelitian.

Kata kunci: Berpikir tingkat tinggi, Instrumen tes, Hasil belajar matematika, Pengembangan

Received 09 April 2020; **Accepted** 29 April 2020; **Published** 01 June 2020

Citation: Ndiung, S. & Jediut, M. (2020). Pengembangan instrumen tes hasil belajar matematika peserta didik sekolah dasar berorientasi pada berpikir tingkat tinggi. *Premiere Educandum : Jurnal Pendidikan Dasar dan Pembelajaran*, 10(1), 94 – 111. Doi.org/10.25273/pe.v10i1.6274



Copyright ©2020 Premiere Educandum : Jurnal Pendidikan Dasar dan Pembelajaran

Published by Universitas PGRI Madiun. This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

PENDAHULUAN

Salah satu komponen penting dalam penyelenggaraan pendidikan adalah penilaian. Peningkatan kualitas pendidikan dapat ditempuh melalui peningkatan kualitas pembelajaran dan kualitas sistem penilaiannya. Konsep dasar penilaian yang dikembangkan mengacu pada revisi terkini dalam Kurikulum 2013 yang diberlakukan difokuskan dengan penyempurnaan pada bagian standar isi dan standar penilaian. Pada standar isi dirancang agar peserta didik mampu berpikir kritis dan analitis sesuai dengan standar internasional yang dilakukan dengan mengurangi materi yang tidak relevan dan pendalaman serta perluasan materi yang relevan bagi peserta didik, sedangkan pada standar penilaian dilakukan dengan mengadaptasi model-model penilaian standar internasional secara bertahap. Penilaian hasil belajar lebih menitikberatkan pada kemampuan berpikir tingkat tinggi (*Higher Order Thinking Skills/HOTS*) (Kemendikbud, 2017).

Kemampuan berpikir tingkat tinggi peserta didik dilakukan berdasarkan hasil studi *Programme for International Student Assessment (PISA)* yang menunjukkan bahwa prestasi literasi membaca (*reading literacy*), literasi matematika (*mathematical literacy*), dan literasi sains (*scientific literacy*) peserta didik Indonesia sangat rendah. Kemampuan berpikir tingkat tinggi dapat dilatih dalam proses pembelajaran matematika di kelas. Namun penerapan pembelajaran yang berorientasi pada pembentukan keterampilan berpikir tingkat tinggi bukan hal yang mudah dilaksanakan oleh guru. Guru harus benar-benar menguasai materi dan strategi pembelajaran dan guru juga dihadapkan pada tantangan dengan lingkungan peserta didik. Pembelajaran akan bermakna jika peserta didik diajak berpikir tingkat tinggi. Keberhasilan penguasaan suatu konsep akan didapatkan ketika peserta didik sudah mampu berpikir tingkat tinggi. Agar HOTS siswa dapat berkembang dengan baik, maka siswa perlu dibiasakan dengan aktivitas-aktivitas yang melatih HOTS itu sendiri dimana peserta didik tidak hanya dapat mengingat dan memahami suatu konsep, namun peserta didik dapat menganalisis serta mensintesis, mengevaluasi, dan mengkreasikan suatu konsep dengan baik, konsep yang telah dipahami tersebut dapat melekat dalam ingatan mereka dalam waktu yang lama, sehingga penting sekali bagi peserta didik untuk memiliki keterampilan berpikir tingkat tinggi. (Arifin & Retnawati, 2017). Kemampuan berpikir tingkat tinggi melibatkan analisis dan sintesis (C4), mengevaluasi (C5), dan mencipta atau kreativitas (C6) (Anderson & Krathwohl, 2001). HOTS sangat penting untuk diterapkan dan dikembangkan dalam pembelajaran di kelas. Dalam hal ini jika peserta didik memiliki kemampuan berpikir tingkat tinggi, maka mereka mampu menyelesaikan masalah matematika dengan baik.

Keberhasilan penguasaan suatu konsep hanya akan diperoleh jika peserta didik memiliki kemampuan berpikir tingkat tinggi, karena konsep yang telah dipahami akan melekat dalam ingatan peserta didik dalam waktu yang lama, sehingga penting sekali bagi peserta didik untuk memiliki keterampilan berpikir tingkat tinggi (Hanifah, 2019). Hasil belajar optimal yang ditunjukkan oleh peserta didik tidak terlepas dari kreativitas mereka dalam menyelesaikan masalah. Kreativitas merupakan komponen penting yang dimiliki oleh peserta didik agar mampu menyelesaikan masalah matematis (Darminto, 2013). Kemampuan ini sangat penting karena dalam kehidupan sehari-hari setiap manusia selalu berhadapan dengan berbagai masalah yang harus diselesaikan termasuk menyelesaikan masalah matematis yang berkaitan dengan soal-soal yang bersifat non rutin. Dengan peserta didik mampu menyelesaikan masalah yang baik maka dapat dikatakan bahwa peserta didik tersebut hasil belajarnya meningkat.

Standar baru yang dibutuhkan peserta didik untuk memiliki kompetensi yang dibutuhkan saat ini yang paling utama adalah mendorong pendidik untuk menyiapkan para peserta didik dengan keterampilan yang dibutuhkan pada kehidupan abad 21 yang dikenal dengan "four Cs" yaitu *critical thinking and problem solving, communication and collaboration, and creativity and innovation* (Partnership for 21st Century learning, 2011).

Kompetensi-kompetensi tersebut penting diajarkan pada peserta didik dalam konteks bidang studi inti dan tema abad ke-21 (Setianingsih, Sa'dijah, & Rahman, 2017). Salah satu kompetensi yang ditekankan pada pendidikan abad 21 adalah *learning and innovation skills*, mencakup berpikir kritis dan mengatasi masalah, yaitu peserta didik mampu menggunakan berbagai alasan seperti berpikir induktif atau deduktif untuk berbagai situasi; menggunakan cara berpikir sistematis; membuat keputusan dan mengatasi masalah; komunikasi dan kolaborasi, yakni peserta didik mampu berkomunikasi dengan jelas dan melakukan kolaborasi dengan anggota kelompok lainnya; dan kreativitas dan inovasi, yaitu peserta didik mampu berpikir kreatif, bekerja secara kreatif dan menciptakan inovasi baru (Ndiung, Dantes, Ardana, & Marhaeni, 2019; Partnership for 21st Century learning, 2011).

Sampai saat ini banyak instrumen hasil belajar baik yang digunakan oleh guru untuk ulangan harian maupun yang digunakan oleh sekolah untuk ulangan umum belum mengarah kepada penyelesaian masalah yang menantang penalaran peserta didik terkait penyelesaian soal-soal yang tidakbiasanya yang menuntut penyelesaian dengan prosedur non rutin. Hal tersebut terbukti dari laporan hasil studi *Programme for International Student Assessment (PISA)* tahun 2018 yang diikuti oleh 79 negara, peserta didik Indonesia berada pada peringkat yang kurang memuaskan yang cenderung stagnan dalam 10-15 tahun terakhir. Hasil studi menunjukkan bahwa kemampuan Literasi berada pada peringkat 72 dari 77 negara, Matematika berada di peringkat 72 dari 78 negara, dan Sains berada pada peringkat 70 dari 78 negara hal ini membuat Indonesia berada pada peringkat 74 (OECD, 2019). Berdasarkan laporan terbaru tersebut, perfoma Indonesia cenderung menurun jika dibandingkan dengan laporan PISA 2015. Hal ini bisa dilihat dari tiga aspek yang dinilai. Perbandingan perolehan skornya dapat dilihat pada **Tabel 1** berikut.

TABEL 1. Perbandingan Skor PISA tahun 2015 dan tahun 2018

No	Aspek	Skor PISA 2015	Skor PISA 2018
1	Membaca	397	371
2	Matematika	386	379
3	Kinerja Sains	403	396

Hasil studi ini mengindikasikan bahwa kualitas pembelajaran matematika masih perlu ditingkatkan agar kemampuan pemecahan masalah matematika yang ditunjukkan oleh peningkatan hasil belajar peserta didik semakin baik untuk menghadapi tantangan abad 21. Pencapaian yang baik tentu didukung oleh kemampuan peserta didik yang sejauh ini belum menunjukkan perubahan yang belum maksimal. Pernyataan ini sejalan dengan temuan penelitian (Tryanasari, Aprilia, & Cahya, 2017) yang mengatakan bahwa kendala utama pembelajaran literasi terletak pada faktor pemahaman guru itu sendiri serta faktor kebijakan sekolah yang belum membentuk tim Gerakan Literasi Sekolah (GLS).

Selanjutnya PISA membagi capaian kemampuan literasi siswa dalam enam tingkatan kecakapan, mulai level 1 (terendah) sampai level 6 (tertinggi) untuk matematika dan sains. Level-level tersebut menggambarkan tingkat penalaran dalam menyelesaikan masalah. Mayoritas siswa Indonesia belum mencapai level 2 untuk matematika (75,7%) dan sains (66,6%), yang memprihatinkan 42,3% siswa bahkan belum mencapai level kecakapan terendah (level 1) untuk matematika dan 24,7% untuk sains (OECD, 2013).

Kemampuan menyelesaikan soal PISA dibagi menjadi 6 level yaitu kemampuan mengingat pada level 1, kemampuan memahami level 2, kemampuan menerapkan level 3, kemampuan menganalisis level 4, kemampuan mengevaluasi level 5, dan kemampuan mencipta pada level 6. Dari hasil penelitian (Simalango, Darmawijoyo, & Aisyah, 2018) menyatakan bahwa penyelesaian soal PISA tahun 2012 level 4,5 dan 6 siswa mengalami kesulitan dalam memahami soal, mengkonversikan permasalahan nyata dalam bentuk matematika serta pemecahannya. Dalam menyelesaikan soal level 4 siswa merasa

kesulitan dalam menginterpretasikan dalam situasi nyata. Dengan pembiasaan untuk menyelesaikan soal berpikir tinggi maka dapat membentuk siswa yang aktif, positif dalam menyelesaikan masalah. Oleh karena itu, siswa di Indonesia diharapkan memiliki setiap kemampuan berpikir tersebut agar mampu menyelesaikan soal-soal yang menuntut siswa menggunakan setiap kemampuan tersebut. Salah satunya adalah soal-soal PISA. Hasil yang diperoleh oleh siswa Indonesia pada PISA 2009 untuk soal di bawah level 2 persentasi siswa Indonesia yang memberikan jawaban benar lebih tinggi dari rata-rata persentasi siswa dari negara-negara OECD, tapi untuk soal level 5 dan 6 persentasi siswa yang memberikan jawaban benar mendekati nol, jauh dari rata-rata persentasi siswa dari negara-negara OECD (Johar, 2012).

Guna mencapai tujuan pendidikan saat abad 21, Kemendikbud akan melakukan inovasi dan terobosan yang diperlukan untuk mempercepat proses dan melakukan pelompatan bidang pendidikan. Salah satu yang sedang dikaji Kemendikbud adalah pembenahan sistem penilaian yang lebih bermanfaat bagi perbaikan praktik pengajaran di kelas maupun perumusan kebijakan pendidikan. Berdasarkan siaran pers Nomor : 078/Sispres/A6/IV/2020 (edukasi.kompas.com) salah satu strategi Mendikbud pembelajaran holistik adalah standar penilaian global. Asesmen Kompetensi Minimum (AKM) akan digunakan untuk mengukur kinerja sekolah berdasarkan literasi dan numerasi yang merupakan kompetensi inti yang menjadi fokus tes PISA, *Trends in International Mathematics and Science Study* (TIMSS), dan *Progress international in Reading Literacy Study* (PIRLS). Oleh karena itu, dalam mengembangkan instrumen tes guru perlu memperhatikan tingkatan proses kognitif peserta didik yang memacu kreativitas dalam berpikir khususnya dalam menyelesaikan masalah matematika dalam kehidupan sehari-hari.

Dalam meningkatkan kualitas pendidikan membutuhkan perbaikan dalam proses pembelajaran di sekolah dengan menggunakan metode kerja yang sistematis salah satunya adalah inovasi dari sistem penilaian. Artinya bahwa dalam meningkatkan sistem pembelajaran di sekolah, diperlukan beberapa informasi untuk menggambarkan hasil yang sistematis dari guru, sekolah serta lembaga pendidikan yang seharusnya profesional dalam melaksanakan penilaian (Pandora & Mardapi, 2017). Evaluasi dapat mendorong siswa untuk lebih giat belajar secara terus menerus dan juga mendorong guru untuk lebih meningkatkan kualitas proses pembelajaran serta mendorong sekolah untuk lebih meningkatkan fasilitas dan kualitas belajar siswa. Sehubungan dengan hal tersebut, optimalisasi sistem evaluasi memiliki dua makna, pertama adalah sistem evaluasi yang memberikan informasi yang optimal. Kedua adalah manfaat yang dicapai dari evaluasi. Manfaat yang utama dari evaluasi adalah meningkatkan kualitas pembelajaran dan selanjutnya akan terjadi peningkatan kualitas pendidikan (Mahirah, 2017). Oleh karena itu, penting dilakukan pengembangan suatu instrumen yang mampu meningkatkan kemampuan pemecahan masalah melalui soal-soal tes yang mengukur berpikir tingkat tinggi yang bermuara pada pencapaian hasil belajar yang maksimal. Instrumen yang dikembangkan dalam penelitian ini bertujuan untuk mengetahui tingkat kesukaran, validitas, dan reliabilitas butir tes hasil belajar matematika pada peserta didik kelas V sekolah dasar.

Tes Hasil Belajar Matematika Berorientasi HOTS

Asesmen merupakan suatu proses pengumpulan data peserta didik baik yang dilakukan selama proses pembelajaran, maupun terhadap hasil belajar. Data-data yang dikumpulkan tersebut selanjutnya dianalisis dan hasil analisis tersebut berfungsi sebagai umpan balik terhadap proses pembelajaran, maupun sebagai bahan pengambilan keputusan terhadap status peserta didik. Dalam mengajar guru selalu menuntut peserta didik belajar dan jarang memberikan pelajaran tentang cara peserta didik untuk belajar, guru juga menuntut peserta didik menyelesaikan masalah, tapi jarang mengajarkan peserta didik tentang cara menyelesaikan masalah (Arends, 2013). Lebih lanjut Arends

mengemukakan bahwa penilaian (*assessment*) hasil belajar merupakan proses mengumpulkan informasi tentang peserta didik dan kelas yang bertujuan untuk membuat keputusan-keputusan instruksional. Penilaian hasil belajar mengacu pada ranah afektif, kognitif, dan psikomotorik (Anderson & Krathwohl, 2001) yang mengkategorikan bahwa ranah kognitif dibedakan menjadi dua yaitu dimensi proses kognitif dan dimensi pengetahuan. Dimensi proses kognitif berisikan enam kategori Taksonomi Bloom yakni: mengingat (C1), memahami/mengerti (C2), menerapkan (C3), menganalisis (C4), mengevaluasi (C5), dan menciptakan (C6). Sedangkan dimensi pengetahuan berisikan empat kategori yaitu pengetahuan faktual (K1), pengetahuan konseptual (K2), pengetahuan prosedural (K3), dan pengetahuan metakognitif (K4).

TABEL 2. *Enam Level Kemampuan Matematika dalam PISA*

Level	Kompetensi Matematika
6	Para siswa dapat melakukan konseptualisasi dan generalisasi dengan menggunakan informasi berdasarkan modelling dan penelaahan dalam suatu situasi yang kompleks. Mereka dapat menghubungkan sumber informasi berbeda dengan fleksibel dan menerjemahkannya. Para siswa pada tingkatan ini telah mampu berpikir dan bernalar secara matematika. Mereka dapat menerapkan pemahamannya secara mendalam disertai dengan penguasaan teknis operasi matematika, mengembangkan strategi dan pendekatan baru untuk menghadapi situasi baru. Mereka dapat merumuskan dan mengkomunikasikan apa yang mereka temukan. Mereka melakukan penafsiran dan berargumentasi secara dewasa.
5	Para siswa dapat bekerja dengan model untuk situasi yang kompleks, mengetahui kendala yang dihadapi, dan melakukan dugaan-dugaan. Mereka dapat memilih, membandingkan, dan mengevaluasi strategi untuk memecahkan masalah yang rumit yang berhubungan dengan model ini. Para siswa pada tingkatan ini dapat bekerja dengan menggunakan pemikiran dan penalaran yang luas, serta secara tepat menghubungkan pengetahuan dan keterampilan matematikanya dengan situasi yang dihadapi. Mereka dapat melakukan refleksi dari apa yang mereka kerjakan dan mengkomunikasikannya.
4	Para siswa dapat bekerja secara efektif dengan model dalam situasi yang konkret tetapi kompleks. Mereka dapat memilih dan mengintegrasikan representasi yang berbeda, dan menghubungkannya dengan situasi nyata. Para siswa pada tingkatan ini dapat menggunakan keterampilannya dengan baik dan mengemukakan alasan dan pandangan yang fleksibel sesuai dengan konteks. Mereka dapat memberikan penjelasan dan mengkomunikasikannya disertai argumentasi berdasar pada interpretasi dan tindakan mereka.
3	Para siswa dapat melaksanakan prosedur dengan baik, termasuk prosedur yang memerlukan keputusan secara berurutan. Mereka dapat memilih dan menerapkan strategi memecahkan masalah yang sederhana. Para siswa pada tingkatan ini dapat menginterpretasikan dan menggunakan representasi berdasarkan sumber informasi yang berbeda dan mengemukakan alasannya. Mereka dapat mengkomunikasikan hasil interpretasi dan alasan mereka.
2	Para siswa dapat menginterpretasikan dan mengenali situasi dalam konteks yang memerlukan inferensi langsung. Mereka dapat memilih informasi yang relevan dari sumber tunggal dan menggunakan cara representasi tunggal. Para siswa pada tingkatan ini dapat mengerjakan algoritma dasar, menggunakan rumus, melaksanakan prosedur atau konvensi sederhana. Mereka mampu memberikan alasan secara langsung dan melakukan penafsiran harafiah.
1	Para siswa dapat menjawab pertanyaan yang konteksnya umum dan dikenal serta semua informasi yang relevan tersedia dengan pertanyaan yang jelas. Mereka bisa mengidentifikasi informasi dan menyelesaikan prosedur rutin menurut instruksi eksplisit. Mereka dapat melakukan tindakan sesuai dengan stimuli yang diberikan

Matematika merupakan alat penting bagi peserta didik menghadapi masalah dan tantangan dalam aspek pribadi, pekerjaan, sosial, dan ilmiah kehidupan mereka. Dengan demikian, penting untuk memiliki pemahaman tentang sejauh mana peserta didik cukup

siap untuk menerapkan matematika dalam memecahkan masalah (OECD, 2013). Demikian halnya (Brookhart, 2010) mengemukakan bahwa kemampuan berpikir tingkat tinggi (HOTS) meliputi kemampuan logika dan penalaran (*logic and reasoning*), analisis (*analysis*), evaluasi (*evaluation*), dan kreasi (*creation*), pemecahan masalah (*problem solving*), dan pengambilan keputusan (*judgement*).

Kemampuan matematika siswa dalam PISA dibagi menjadi enam level (tingkatan), level 6 sebagai tingkat pencapaian yang paling tinggi dan level 1 yang paling rendah (OECD, 2013). Setiap level tersebut menunjukkan tingkat kompetensi matematika yang dicapai siswa. Secara lebih rinci level-level yang dimaksud tergambar pada **Tabel 2**.

Pengembangan instrumen hasil belajar matematika dalam penelitian ini terkait dengan kemampuan peserta didik dalam menyelesaikan masalah pada soal-soal yang bersifat non rutin yang mengacu pada pendapat Polya (Van De Walle, 2013) yang mengemukakan bahwa terdapat empat langkah penting pemecahan masalah matematika, yaitu: 1) memahami masalah. Pada tahap ini, kegiatan pemecahan masalah diarahkan untuk membantu peserta didik menetapkan apa yang diketahui pada permasalahan dan apa yang ditanyakan; 2) membuat rencana untuk menyelesaikan masalah. Dalam memilih strategi yang sesuai dengan permasalahan yang akan dipecahkan; 3) melaksanakan rencana yang telah dibuat. Melaksanakan strategi atau pendekatan sesuai dengan yang telah direncanakan; dan 4) memeriksa ulang jawaban yang diperoleh. Langkah ini penting dilakukan untuk mengecek apakah hasil yang diperoleh sudah sesuai dengan ketentuan dan tidak terjadi kontradiksi dengan yang ditanya. Pemecahan masalah adalah usaha mencari jalan keluar dari suatu kesulitan (Yuwono, 2016). Masalah dalam pembelajaran matematika biasanya ditemukan pada saat diberikan soal/pertanyaan dikarenakan ketidakmampuan menemukan aturan untuk menyelesaikannya.

Aspek kemampuan yang diukur dalam tes hasil belajar matematika pada level pengukuran kemampuan tingkat tinggi mengacu pada taksonomi Bloom revisi (Gunawan & Palupi, 2012), mencakup proses kognitif pada aspek analisis, evaluasi, dan mencipta sedangkan pada aspek pengetahuannya mencakup pengetahuan konseptual, prosedural, dan metakognitif. Instrumen tes hasil belajar matematika yang dikembangkan dalam penelitian ini mengacu pada soal-soal pemecahan masalah non rutin tentang menyelesaikan masalah pecahan yang dikembangkan mengukur keterampilan berpikir tingkat tinggi peserta didik kelas VI sekolah dasar pada proses kognitif analisis (C4) dengan dimensi pengetahuan konseptual (K2) dan pengetahuan prosedural (K3).

Konstruksi Tes

Tes merupakan merupakan sejumlah pertanyaan yang memiliki jawaban yang benar atau salah. Tes diartikan juga sebagai sejumlah pertanyaan yang membutuhkan jawaban atau tanggapan dengan tumengukur tingkat kemampuan seseorang. Ada sembilan langkah yang perlu ditempuh dalam mengembangkan tes hasil belajar, yaitu: (1) menyusun spesifikasi tes, (2) menulis soal tes, (3) menelaah soal tes, (4) melakukan uji coba tes, (5) menganalisis butir soal, (6) memperbaiki tes, (7) merakit tes, dan (8) melaksanakan tes, dan (9) menafsirkan hasil tes (Mardapi, Kunaidi, & Kartowagiran, 2011). Dalam menyusun spesifikasi tes kegiatan yang dilakukan adalah menentukan tujuan tes, menyusun kisi-kisi tes, memilih bentuk tes, dan menentukan panjang tes. Selanjutnya setelah soal tes disusun maka dilakukan uji coba tes yang bertujuan memperbaiki kualitas tes. Melalui uji coba diperoleh data tentang validitas, reliabilitas, tingkat kesukaran, daya beda butir, dan lain sebagainya. Kegiatan perancangan tes tercakup di dalamnya yakni: (1) penetapan tujuan, (2) penyiapan tabel spesifikasi, (3) menyeleksi format item yang sesuai, (4) menulis item, dan (5) mengedit item. Kegiatan uji coba tes meliputi kegiatan: (1) analisis item pengujian uji coba pertama, (2) analisis item pengujian uji coba kedua, dan (3) penyiapan format tes. Prosedur ini harus diikuti untuk menghasilkan instrumen tes yang baik (Pandra & Mardapi, 2017).

Salah satu komponen yang penting dalam mengembangkan tes hasil belajar adalah langkah desain. Oleh karena itu, seyogyanya dalam pengembangan tes mengikuti langkah-langkah di atas dimana pengembangan tes hasil belajar matematika di kelas lima dalam penelitian ini menggunakan langkah-langkah berikut ini, yaitu: (1) menentukan bentuk tes, yaitu tes esai, (2) menulis tes, (3) mereview dan merevisi pertanyaan, (4) uji coba lapangan, (5) mengumpulkan hasil tes, (6) penyekoran, dan (7) melaporkan hasil tes.

Tingkat Kesukaran Butir Tes

Tingkat kesukaran item atau indeks kesulitan adalah angka yang menunjukkan proporsi peserta didik yang menjawab betul dalam satu soal. Menurut Daryanto, 2007 (Anderson & Krathwohl, 2001), soal yang baik adalah soal yang tidak terlalu mudah dan tidak terlalu sukar. Soal yang terlalu mudah tidak merangsang peserta didik untuk berpikir lebih dalam memecahkannya. Sebaliknya soal yang terlalu suka akan menyebabkan peserta didik menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi karena di luar jangkauannya. Sehingga dapat dikatakan bahwa tingkat kesukaran soal adalah angka yang menunjukkan bahwa apakah soal yang diujikan termasuk mudah, mudah, dan sukar. Perhitungan tingkat kesukaran butir soal dihitung dengan seberapa besar derajat suatu kesukaran suatu soal oleh peserta didik dan soal yang dikatakan sukar apabila hanya sedikit peserta didik yang bisa menjawabnya. Untuk menentukan tingkat kesukaran butir soal tes politomi (Candiasa, 2010) dihitung dengan formula:

$$P = \frac{\sum U + \sum L - (2N \times S_{\min})}{2N(s_{\max} - S_{\min})} \quad (1)$$

Keterangan: $\sum U$ = total skor kelompok atas; $\sum L$ = total skor kelompok bawah; s_{\max} = Skor maksimal butir; S_{\min} = Skor miniman butir; N = Banyak peserta Tes.

Jika nilai P mendekati 0, maka item tersebut sukar, dan jika nilainya mendekati 1, maka item tersebut mudah. Suatu item dikatakan terlalu mudah atau terlalu sukar tidak menggambarkan informasi tentang tes karena item tersebut tidak mampu membedakan kemampuan peserta tes sehingga item tersebut gugur. Level tingkat kesukaran butir tes (P) yang baik adalah berada pada rentangan 0.3 sampai dengan 0.7 sebagai gambaran kemampuan maksimal peserta tes (Anderson & Krathwohl, 2001). Tetapi rentangan dari 0.3 sampai dengan 0.7 dapat ditolerir untuk tujuan pengembangan item. Setiap item yang dikembangkan memiliki nilai P yang berbeda tergantung tujuan pengembangannya.

Validitas dan Realibilitas Instrumen

Validitas sebagai berikut *validity of a est has been defined as the extent to which the test measures what it was designed to measures*. Suatu tes dikatakan memiliki validitas yang tinggi apabila alat tersebut menjalankan fungsi ukur secara tepat atau memberikan hasil ukur yang sesuai dengan maksud dikakukannya pengukuran tersebut (Sudaryono, 2011). Artinya hasil ukur dari pengukuran tersebut merupakan besaran yang mencemrinkan secara tepat fakta atau keadaan sesungguhnya dari apa yang diukur. Tes yang akan digunakan dalam penelitian ini adalah tes uraian atau tes esai. Validitas instrumen politomi, seperti tes uraian atau disebut juga tes esai ditentukan dengan korelasi *product moment* dari Carl Pearson dengan rumus seperti berikut.

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}} \quad (2)$$

yang mana X : menunjukkan skor yang diperoleh peserta didik pada butir yang akan diuji validitasnya; Y : menunjukkan skor total yang diperoleh responden untuk semua butir tes; N : banyak responden.

Validitas butir soal tes oleh koefisien korelasi product moment (r_{xy}). Kriteria penerimaan butir soal adalah sebagai berikut: (1) jika $r_{xy(\text{hitung})} > r_{xy(\text{tabel})}$ maka butir soal dinyatakan valid, sedangkan (2) jika $r_{xy(\text{hitung})} < r_{xy(\text{tabel})}$ maka butir soal dinyatakan tidak valid atau direvisi.

Sementara, reliabilitas adalah konsistensi hasil pengukuran yang ditunjukkan pada waktu yang berbeda pada subjek yang sama. Tes dikatakan reliabel jika skor perolehan memiliki korelasi yang tinggi dengan skor total. Nilai validitas dan reliabilitas suatu instrumen dipengaruhi oleh subjek yang diukur, pengguna instrument, dan instrument itu sendiri. sehingga validitas dan reliabilitas harus selalu diuji sebelum instrument digunakan (Yusup, 2018). Reliabilitas instrument dapat diuji dengan beberapa uji reliabilitas. Pada penelitian ini, menggunakan teknik uji reliabilitas *internal consistency* yaitu *Alpha Cronbach* karena dilakukan untuk menguji tes yang memiliki benar lebih dari 1 dengan bentuk instrumennya adalah esai (Yusup, 2018). Reliabilitas mengacu pada konsistensi hasil pengukuran yang ditunjukkan oleh instrumen tersebut. Formula *Alpha Cronbach* digunakan dalam menghitung reliabilitas tes uraian, dengan rumus sebagai berikut:

$$r_{11} = \left(\frac{n}{n-1} \right) \left(1 - \frac{\sum S_i^2}{S_t^2} \right) \quad (3)$$

Keterangan: r_{11} : koefisien reliabilitas Alpha Cronbach; n : banyaknya butir instrumen yang valid; $\sum s_i^2$: jumlah varian skor tiap item; s_t^2 : varian total.

Pengujian Daya Beda Butir Tes Hasil Belajar Matematika

Analisis daya beda adalah pengkajian butir tes yang dimaksudkan untuk mengetahui kesanggupan butir tes untuk membedakan peserta didik yang tergolong mampu dengan siswa yang tergolong tidak mampu. Indeks daya beda butir dinyatakan dengan d . Dalam penelitian ini digunakan indeks daya beda butir tes politomi. Indeks daya beda butir untuk tes non dikotomi (politomi) dihitung dengan formula Ferguson sebagai berikut.

$$d = \frac{N^2 - \sum f_i^2}{N^2 - \frac{N^2}{n+1}} = \frac{(n+1)(N^2 - \sum f_i^2)}{nN^2} \quad (4)$$

yang mana: d = Indeks daya beda; N = banyak peserta tes; f_i = frekuensi pada tiap-tiap skor; n = banyak butir. Kriteria besarnya koefisien daya beda klasifikasikan menjadi empat kategori, adalah sebagai berikut.

TABEL 3. *Klasifikasi Daya Beda Butir*

Kategori Daya Beda	Koefisien
Baik	0,40 – 1,00
Sedang (tidak perlu revisi)	0,30 – 0,39
Perlu Revisi	0,20 – 0,29
Tidak baik	-1,00 – 0,19

Apabila indeks daya beda Ferguson satu butir bernilai positif cukup besar, maka butir tersebut banyak dipilih dengan benar oleh peserta didik yang memperoleh skor tinggi. Artinya daya beda butir tersebut tinggi karena mampu membedakan siswa yang mampu dan siswa yang tidak mampu. Sebaliknya, apabila indeks daya beda butir atau bahkan negatif, maka butir tersebut banyak dijawab salah oleh kebanyakan peserta didik

yang mendapatkan skor tinggi. Artinya, daya beda butir tersebut rendah, karena tidak mampu membedakan peserta didik yang mampu dan peserta didik yang kurang mampu. Daya beda butir dengan formula Ferguson kriteria penerimaannya jika $d > 0,40$.

METODE

Penelitian ini merupakan penelitian pengembangan, yaitu pengembangan instrument hasil belajar matematika peserta didik sekolah dasar. Produk yang akan dihasilkan dari penelitian ini adalah instrumen tes hasil belajar matematika bagi peserta didik kelas V sekolah dasar yang berorientasi pada proses berpikir tingkat tinggi. Artinya bahwa konstruksi butir tes yang dikembangkan berdasarkan tingkatan pengetahuan dari taxonomy Bloom Revisi yaitu pada aspek proses kognitif menganalisis (C4), mengevaluasi (C5) dan mencipta (C6) dan pada aspek pengetahuan konseptual (K2), procedural (K3), dan metakognisi (K4) (Gunawan & Palupi, 2012). Pengembangan instrumen hasil belajar matematika dalam penelitian ini terkait dengan kemampuan peserta didik dalam menyelesaikan masalah pada soal-soal yang bersifat non rutin yang memuat aspek proses kognitif analisis (C4) dan aspek pengetahuan konseptual (K2) dan prosedural (K3) yang secara rinci dapat dilihat pada kisi-kisi tes pada Tabel 5.

Item yang dikembangkan sebanyak 7 butir soal-soal pemecahan masalah yang berorientasi pada pengukuran berpikir tingkat tinggi. Subjek penelitian ini adalah 30 peserta didik kelas VI SDK Wae Mata Kecamatan Lembor Kabupaten Manggarai Barat. Dalam penelitian ini pengembangan tes menggunakan langkah-langkah berikut ini, yaitu: (1) menentukan bentuk tes, yaitu tes esai, (2) menulis tes, (3) mereview dan merevisi pertanyaan, (4) uji coba lapangan, (5) mengumpulkan hasil tes, (6) penyekoran, dan (7) melaporkan hasil tes.

Tes yang dikembangkan dalam penelitian ini adalah tes tertulis dengan bentuk uraian yang berorientasi pada pengukuran keterampilan berpikir tingkat tinggi. Bentuk tes uraian dikembangkan untuk mengetahui proses penyelesaian masalah dilakukan peserta didik. Validasi instrumen penelitian dilakukan dengan menentukan validitas tes, reliabilitas instrumen tes, indeks kesukaran butir tes dan indeks daya beda butir. Untuk mengetahui validitas empirik butir tes hasil belajar matematika ditempuh dengan dua prosedur, yaitu melalui penilaian ahli (*expert judge*) dan analisis validitas konten butir. Validitas isi butir tes hasil belajar matematika melibatkan ahli sesuai dengan keahlian masing-masing berkaitan dengan isi, konstruksi, dan bahasa. Adapun lima orang ahli tersebut berlatar belakang pendidikan matematika, Bahasa, dan Psikologi yang berkualifikasi S3 pada Universitas Pendidikan Ganesha Singaraja Bali. Hasil telaah pakar dianalisis dengan menggunakan pendekatan *content validity Ratio* (CVR) yang dikembangkan oleh Lawshe (Newman, Lim, & Pineda, 2013). Penilaian ahli atau penilaian pakar dilakukan dengan tujuan untuk menelaah ketepatan isi butir tes ditinjau dari aspek relevansi isi dengan tujuan, konstruksi, dan kebakasaannya. Proses penilaian ahli atau pakar dijabarkan ke dalam lima indikator, yaitu: (1) kesesuaian antara butir dengan indikator, (2) kesesuaian antara indikator dengan materi, (3) pernyataan soal tidak ambigu, dan (4) kesesuaian penggunaan bahasa dengan responden, dan (5) butir soal mengukur keterampilan berpikir tingkat tinggi. Pembobotan dilakukan dengan mengacu pada lima indikator di atas dengan kriteria pembobotan dapat dilihat pada **Tabel 4** berikut.

Butir instrument yang dikembangkan disini adalah butir instrument hasil belajar matematika yang mengukur berpikir tingkat tinggi peserta didik sekolah dasar. Kriteria yang digunakan untuk menentukan valid tidaknya butir tes adalah nilai minimum *Content Validity Ratio* (CVR) dalam tabel kritis berdasarkan jumlah penilai atau panelis. Untuk jumlah pakar sebanyak 5 orang, butir tes kemampuan numerik dinyatakan valid apabila memiliki $CVR \geq 0,99$. Adapun kisi-kisi dan soal-soal tes yang dikembangkan dapat dilihat pada **Tabel 5** dan **Tabel 6**.

TABEL 4. *Bobot penilaian pakar*

No	Kriteria	Deskriptor	Skor
1	Tidak Relevan	Jika dalam butir pernyataan, tidak ada indikator atau hanya satu indikator terpenuhi	1
2	Kurang Relevan	Jika dalam butir pernyataan, dua atau tiga indikator terpenuhi	2
3	Relevan	Jika dalam butir pernyataan empat atau lima indikator terpenuhi	3

TABEL 5. *Kisi-kisi tes hasil belajar matematika*

	Kompetensi Dasar	Indikator Soal	Ranah Kognitif	Nomor Butir
1.1	Menjelaskan dan melakukan penjumlahan dan pengurangan dua pecahan dengan penyebut berbeda	Memecahkan persoalan yang berkaitan dengan penjumlahan pecahan dengan penyebut beda	C ₄ K ₂	1
4.1		Menyelesaikan masalah yang berkaitan dengan pengurangan pecahan dengan penyebut beda	C ₄ K ₂	2
4.1	Menyelesaikan masalah yang berkaitan dengan penjumlahan dan pengurangan dua pecahan dengan penyebut berbeda	Menguraikan masalah yang berkaitan dengan pengerjaan hitung campuran berbagai bentuk pecahan	C ₄ K ₃	3
3.2		Menjelaskan dan melakukan perkalian dan pembagian pecahan dan desimal	Memecahkan masalah yang berkaitan dengan perkalian pecahan	C ₄ K ₂
4.2	Menyelesaikan masalah yang berkaitan dengan perkalian dan pembagian pecahan dan desimal	Memecahkan masalah yang berkaitan dengan pembagian pecahan	C ₄ K ₃	6,7
		Memecahkan masalah yang berkaitan dengan pecahan sebagai perbandingan	C ₄ K ₂	4

Keterangan: *C* mengukur proses kognitif dan *K* mengukur dimensi pengetahuan.

Selanjutnya, pengujian validitas butir secara empiris dilakukan setelah uji coba instrumen yaitu validitas internal konsistensi butir tes. Pengujian validitas internal konsistensi butir tes hasil belajar matematika menggunakan metode *Part Whole* yakni dengan mengkorelasikan antara skor butir tes dengan skor total yang berkorelasi tinggi. Tes yang akan digunakan dalam penelitian ini adalah tes uraian atau tes esai atau disebut juga tes politomi. Validitas instrumen politomi, seperti tes uraian atau disebut juga tes esai ditentukan dengan korelasi *product moment* dari Carl Pearson.

Validitas konten butir tes hasil belajar matematika menggunakan metode *Part Whole* yakni dengan mengkorelasikan antara skor butir tes dengan skor total yang berkorelasi tinggi. Tes yang akan digunakan dalam penelitian ini adalah tes uraian atau tes esai. Validitas instrumen politomi, seperti tes uraian atau disebut juga tes esai ditentukan dengan korelasi *product moment* dari Carl Pearson. Validitas butir ditentukan koefisien korelasi *product moment* (r_{xy}). Kriteria penerimaan butir soal adalah sebagai berikut: (1) jika $r_{xy}(\text{hitung}) > r_{xy}(\text{tabel})$ maka butir soal dinyatakan valid, sedangkan (2) jika $r_{xy}(\text{hitung}) < r_{xy}(\text{tabel})$ maka butir soal dinyatakan tidak valid yang selanjutnya butir yang tidak valid tersebut direvisi.

Selanjutnya, instrumen yang baik wajib memenuhi ketentuan reliabilitas instrumen. Reliabilitas menunjukkan suatu pengertian bahwa instrumen cukup dipercaya untuk dipergunakan sebagai alat pengumpulan data, karena instrumen tersebut sudah baik. Reliabilitas mengacu pada konsistensi hasil pengukuran yang ditunjukkan oleh instrumen

tersebut. Instrumen yang memiliki reliabilitas tinggi akan memberikan hasil yang relatif sama, sekalipun instrumen tersebut digunakan dalam kurun waktu yang berbeda. Untuk melihat keajegan suatu perangkat tes dilakukan pengujian reliabilitas terhadap hasil tes tersebut. Reliabilitas perangkat tes dapat dilihat dari besaran koefisien reliabilitas perangkat tes tersebut. Untuk menghitung reliabilitas instrumen tes hasil belajar matematika dengan bentuk tes uraian, digunakan rumus *Alpha Cronbach*. Klasifikasi derajat reliabilitas menurut Guilford, adalah sebagai berikut.

- $r_{11} \leq 0,20$: derajat reliabilitas sangat rendah
- $0,20 < r_{11} \leq 0,40$: derajat reliabilitas rendah
- $0,40 < r_{11} \leq 0,60$: derajat reliabilitas sedang
- $0,60 < r_{11} \leq 0,80$: derajat reliabilitas tinggi
- $0,80 < r_{11} \leq 1,00$: derajat reliabilitas sangat tinggi

TABEL 6. Soal-soal Tes hasil Belajar Matematika

No	Permasalahan (Soal Tes)
1	Klara akan membuat kue yang memerlukan 5 kilogram tepung terigu. Ia hanya memiliki takaran yang ukurannya $\frac{1}{2}$ kg dan $\frac{3}{4}$ kg untuk menakar tepung tersebut. Paling sedikit berapa takaran tepung yang diperlukan untuk membuat kue tersebut?
2	Pak Andreas dapat menyelesaikan pembuatan sebuah lemari dalam waktu 6 hari. Jika pekerjaan itu dikerjakan bersama-sama dengan pak Yohanes, maka dapat diselesaikan dalam dua hari. Seandainya pekerjaan itu diselesaikan oleh pak Yohanes sendiri, berapa hari akan selesai?
3	Sebuah semangka yang beratnya 1 kg terdiri dari 93% air. Setelah dijemur di bawah matahari untuk beberapa saat, kandungan airnya menjadi 90%. Berapa berat semangka sekarang?
4	Perbandingan panjang dan lebar suatu persegi panjang adalah 5 : 3. Jika luas persegi panjang itu adalah 240 cm ² , maka tentukan panjang dan lebar dari persegi panjang itu.
5	Luas kebun Krisetia adalah 750m ² dan $\frac{2}{5}$ bagiannya akan ditanami sayur-sayuran. Berapa m ² luas kebun yang ditanami sayur-sayuran?
6	Jika jus jeruk yang berada dalam galon yang mempunyai volume $8\frac{1}{2}$ liter akan dituangkan ke dalam gelas-gelas, yang mana setiap kapasitas sebuah gelas adalah 250 ml. Tentukan berapa paling sedikit gelas yang digunakan untuk menampung jus jeruk tersebut?
7	Sepuluh buah ember dapat mengisi $\frac{5}{8}$ bagian dari drum air. Jika 4 buah ember dan 5 buah baskom dapat mengisi sisa drum, berapa banyak baskom dapat mengisi drum air itu?



(Adaptasi: Aisyah & Zulkardi, 2010)

Untuk menentukan tingkat kesukaran butir soal menggunakan formula tingkat kesukaran butir soal. Kriteria untuk indeks kesukaran menurut Sudjana adalah sebagai berikut. Butir dengan *P* 0,00 sampai 0,30 tergolong sukar; Butir dengan *P* 0,31 sampai 0,70 tergolong sedang; dan Butir dengan *P* 0,71 sampai 1,00 tergolong mudah. Level tingkat kesukaran butir tes (*P*) yang baik adalah berada pada rentangan 0.3 sampai dengan 0.7 sebagai gambaran kemampuan maksimal peserta tes.

Selanjutnya, ditentukan indeks daya beda butir. Dalam penelitian ini, indeks daya beda butir untuk tes non dikotomi (politomi) dihitung dengan formula Ferguson. Apabila indeks daya beda bernilai positif cukup besar, maka butir tersebut banyak dipilih dengan

benar oleh peserta didik yang memperoleh skor tinggi. Sebaliknya, apabila indeks daya beda butir atau bahkan negatif, maka butir tersebut banyak dijawab salah oleh kebanyakan peserta didik yang mendapatkan skor tinggi. Daya beda butir dengan formula Ferguson kriteria penerimaannya jika $IDB > 0,40$. Teknik Analisis data untuk mengetahui tingkat kesukaran, menentukan validitas, reliabilitas instrumen dan daya beda butir menggunakan bantuan Program *MS Excel 2016*.

HASIL PENELITIAN

Berpedoman pada kisi-kisi tes, disusun butir tes hasil belajar matematika dalam bentuk uraian. Setelah disusun, kemudian diujicobakan kepada siswa kelas VI sekolah dasar di luar subyek penelitian, serta validasi isi oleh lima orang ahli yang berlatar belakang pendidikan matematika, Bahasa, dan Psikologi yang berkualifikasi S3. Uji coba empiris instrumen tes yang dikembangkan dilaksanakan pada peserta didik kelas VI SDK Wae Mata Kecamatan Lembor kabupaten Manggarai Barat yang melibatkan 30 peserta didik. Ujicoba bertujuan untuk memperoleh informasi tentang tingkat kesukaran butir, validitas, koefisien reliabilitas, dan indeks daya beda butir dari instrument tes hasil belajar yang dikembangkan. Setelah uji coba, dilakukan revisi sesuai dengan saran para pakar. Setelah dilakukan uji coba instrumen selanjutnya menentukan tingkat kesukaran, validitas dan reliabilitas instrumen untuk melihat kelayakan instrumen tersebut digunakan sebagai instrumen penelitian.

Tingkat Kesukaran Butir Soal

Tingkat kesukaran butir yang disimbolkan dengan P untuk butir soal tes politomi atau butir soal uraian dihitung dengan formula.

$$P = \frac{\sum U + \sum L - (2N \times S_{\min})}{2N(S_{\max} - S_{\min})} \quad (5)$$

Dari data 30 peserta tes ditentukan kelompok atas dan kelompok bawah masing-masing 27%, maka kelompok atas dan kelompok bawah masing-masing terdiri dari $27/100 \times 30$ orang = 8,1 orang dibulatkan menjadi 8 orang. Setelah skor masing-masing butir diranking, maka diperoleh indeks kesukaran butir seperti terangkum dalam **Tabel 7** berikut.

TABEL 7. Indeks kesukaran butir soal tes hasil belajar matematika

Nomor Butir	Indeks Kesukaran	Kategori
1	0.60	Sedang
2	0.66	Sedang
3	0.53	Sedang
4	0.50	Sedang
5	0.53	Sedang
6	0.63	Sedang
7	0.58	Sedang

Berdasarkan data pada **Tabel 7** di atas diperoleh bahwa indeks tingkat kesukaran masing-masing butir soal berada pada kategori sedang. Dengan demikian dapat dikatakan bahwa kategori butir soal yang telah disusun *level of difficulty item (P)* berada pada rentangan 0.3 to 0.7 yaitu pada kategori sedang. Butir-butir item yang berdasarkan hasil analisis dalam kategori yang baik dalam arti tingkat kesukarannya itemnya sedang, seyogyanya butir item tersebut segera dicatat di bank soal, selanjutnya butir-butir soal tersebut dapat dikeluarkan lagi dalam tes-tes hasil belajar pada waktu-waktu yang akan

datang. Sebaliknya, untuk butir dalam kategori mudah hendaknya guru menelaah kembali dengan cara melacak dan menelusuri sehingga dapat diketahui faktor yang menyebabkan hampir semua peserta tes dapat menjawab butir tes. Jika butir soal termasuk kategori mudah maka pengecoh butir soal itu tidak berfungsi karena semua peserta tes memahami materi yang ditanyakan. Selanjutnya untuk butir-butir item yang termasuk dalam kategori sukar hendaknya dikaji ulang, melacak dan menelusuri sehingga dapat diketahui faktor penyebab butir tersebut sulit dijawab peserta tes (Yani, Asri, & Burhan, 2013). Hasil penelitian mereka juga menyimpulkan bahwa ditinjau dari tingkat kesukaran, instrument yang dikembangkan secara keseluruhan dikategorikan sebagai soal yang sedang sebab persentasi soal terbanyak dalam kategori soal yang baik, dan menyarankan untuk soal-soal yang dikategorikan sebagai soal yang baik hendaknya disimpan dan digunakan pada masa yang akan datang sedangkan untuk soal-soal dikategorikan sebagai soal yang kurang baik hendaknya dilakukan perbaikan-perbaikan.

Validitas Instrumen Tes Hasil Belajar Matematika

Untuk mengetahui validitas instrumen hasil belajar matematika secara empirik dilakukan dengan dua prosedur yaitu analisis validitas isi butir tes hasil belajar matematika dan analisis validitas konstruk butir. Analisis validitas isi butir dengan data empirik hasil telaah ahli menggunakan pendekatan rasio validitas isi (CVR) yang dikembangkan oleh *Lawshe*.

Hasil validasi Isi instrumen hasil belajar matematika oleh lima orang pakar dapat dijelaskan dalam Tabel 5. Kriteria yang digunakan untuk menentukan valid tidaknya butir tes adalah nilai minimum *content validity ratio* (CVR) dalam tabel kritis berdasarkan jumlah penilai atau panelis. Untuk jumlah pakar 5 orang, butir tes hasil belajar matematika dinyatakan valid apabila memiliki $CVR \geq 0,99$.

Berdasarkan hasil analisis validitas instrumen tes hasil belajar matematika diperoleh dari 7 butir tes uraian yang dikembangkan semuanya dinyatakan valid. Secara rinci hasil analisis validitas isi butir tes hasil belajar matematika yang dikembangkan dapat dilihat pada Tabel 8 berikut. Untuk mengetahui validitas butir hasil belajar matematika secara empirik dilakukan dengan dua prosedur yaitu analisis validitas isi butir tes hasil belajar matematika dan analisis validitas konstruk butir. Analisis validitas isi butir dengan data empirik hasil telaah ahli menggunakan pendekatan rasio validitas isi (CVR) yang dikembangkan oleh *Lawshe* dengan rumus:

$$CVR = \frac{n_e - \frac{N}{2}}{\frac{N}{2}} \quad (6)$$

Keterangan: CVR = Rasio validitas isi; n_e = Jumlah panelis yang memberikan penilaian 3 (relevan); N = Banyaknya panelis.

Data penilaian ahli terhadap tes hasil belajar matematika yang dikembangkan masing-masing dapat dilihat pada **Tabel 8** berikut.

TABEL 8. Penilaian ahli terhadap butir tes hasil belajar matematika

No Butir	Penilaian Ahli					Jumlah Nilai 3
	A	B	C	D	E	
1	3	3	3	3	3	5
2	3	3	3	3	3	5
3	3	3	3	3	3	5
4	3	3	3	3	3	5
5	3	3	3	3	3	5
6	3	3	3	3	3	5
7	3	3	3	3	3	5

Mengacu data pada **Tabel 8** di atas, untuk butir 1: $N = 5$, $n_e = 5$, sehingga diperoleh $CVR = 1.00$ (dalam kategori valid). CVR untuk setiap butir adalah dijelaskan pada **Tabel 9** berikut.

TABEL 9. CVR tes hasil belajar matematika

Nomor Butir	N	n_e	CVR	CVR-tabel	Keterangan
1	5	5	1,00	0,99	valid
2	5	5	1,00	0,99	valid
3	5	5	1,00	0,99	valid
4	5	5	1,00	0,99	valid
5	5	5	1,00	0,99	valid
6	5	5	1,00	0,99	Valid
7	5	5	1,00	0,99	Valid

Selanjutnya, pengujian validitas butir secara empiris dilakukan setelah ujicoba instrumen yaitu validitas internal konsistensi butir tes. Ujicoba tes hasil belajar matematika dilakukan pada peserta didik kelas VI SDK Wae Mata Kecamatan Lembor Kabupaten Manggarai Barat dengan jumlah responden 30 peserta didik. Pengujian validitas internal konsistensi butir tes hasil belajar matematika menggunakan metode *Part Whole* yakni dengan mengkorelasikan antara skor butir tes dengan skor total yang berkorelasi tinggi. Tes yang akan digunakan dalam penelitian ini adalah tes uraian atau tes esai atau disebut juga tes politomi. Validitas instrumen politomi, seperti tes uraian atau disebut juga tes esai ditentukan dengan korelasi *product moment* dari Carl Pearson (Bayuni, Candiasa, & Koyan, 2013).

Validitas konten butir tes hasil belajar matematika menggunakan metode *Part Whole* yakni dengan mengkorelasikan antara skor butir tes dengan skor total yang berkorelasi tinggi. Tes yang akan digunakan dalam penelitian ini adalah tes uraian atau tes esai. Validitas instrumen politomi, seperti tes uraian atau disebut juga tes esai ditentukan dengan korelasi *product moment* dari Carl Pearson dengan rumus seperti berikut.

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}} \quad (7)$$

Validitas butir ditentukan koefisien korelasi product moment (r_{xy}). Kriteria penerimaan butir soal adalah sebagai berikut: (1) jika $r_{xy}(\text{hitung}) > r_{xy}(\text{tabel})$ maka butir soal dinyatakan valid, sedangkan (2) jika $r_{xy}(\text{hitung}) < r_{xy}(\text{tabel})$ maka butir soal dinyatakan tidak valid atau drop. Analisis validitas butir terhadap 7 butir tes hasil belajar matematika yang ada diperoleh bahwa ada 2 butir tes yang tidak valid yaitu butir tes nomor 3 dan nomor 7, sedangkan 5 butir tes lainnya dinyatakan valid. Dengan demikian banyaknya butir tes yang dipergunakan untuk mengumpulkan data penelitian tentang hasil belajar matematika peserta didik adalah 5 butir tes uraian pada materi menyelesaikan masalah operasi hitung pecahan. Koefisien korelasi butir soal nomor 1 sampai dengan 7 dapat dilihat dalam **Tabel 10** di bawah ini.

TABEL 10. Rekapitulasi hasil uji validitas butir tes hasil belajar matematika

Nomor Butir	r_{xixi}	r_{tabel}	Kriteria	Keputusan
1	0,713	0,361	Valid	Digunakan
2	0,766	0,361	Valid	Digunakan
3	0,325	0,361	Tidak Valid	Direvisi

4	0,718	0,361	Valid	Digunakan
5	0,637	0,361	Valid	Digunakan
6	0,675	0,361	Valid	Digunakan
7	0,300	0,361	Tidak Valid	Direvisi

Reliabilitas Tes Hasil Belajar Matematika

Instrumen yang baik wajib memenuhi ketentuan reliabilitas instrumen. Reliabilitas menunjukkan suatu pengertian bahwa instrumen cukup dipercaya untuk dipergunakan sebagai alat pengumpulan data, karena instrumen tersebut sudah baik. Reliabilitas instrument dapat diuji dengan beberapa uji reliabilitas. Pada penelitian ini, menggunakan teknik uji reliabilitas *internal consistency* yaitu *Alpha Cronbach* karena dilakukan untuk menguji tes yang memiliki benar lebih dari 1 dengan bentuk instrumennya adalah esai (Yusup, 2018). Reliabilitas mengacu pada konsistensi hasil pengukuran yang ditunjukkan oleh instrumen tersebut. Instrumen yang memiliki reliabilitas tinggi akan memberikan hasil yang relatif sama, sekalipun instrumen tersebut digunakan dalam kurun waktu yang berbeda. Untuk melihat keajegan suatu perangkat tes dilakukan pengujian reliabilitas terhadap hasil tes tersebut. Reliabilitas perangkat tes dapat dilihat dari besaran koefisien reliabilitas perangkat tes tersebut. Untuk menghitung reliabilitas instrumen tes hasil belajar matematika dengan bentuk tes uraian, dengan formula *Alpha Cronbach* (Bayuni et al., 2013)

$$r_{11} = \left(\frac{n}{n-1} \right) \left(1 - \frac{\sum S_i^2}{S_t^2} \right) \quad (8)$$

Hasil analisis reliabilitas butir tes hasil belajar matematika diperoleh koefisien reliabilitasnya adalah 0,79. Selanjutnya varians tiap butirnya dapat dihitung dengan menggunakan rumus di atas. Hasil perhitungan varians butir masing-masing nomor dapat dilihat dalam **Tabel 11** berikut.

TABEL 11. Varians butir soal yang valid

Nomor Butir	S_i^2
1	2,947
2	5,586
4	5,178
5	3,168
6	2,582
$\sum S_i^2$	19,731
$\sum S_t^2$	53,241

$$r_{11} = \left(\frac{n}{n-1} \right) \left(1 - \frac{\sum S_i^2}{S_t^2} \right) = \left(\frac{5}{5-1} \right) \left(1 - \frac{19,731}{53,241} \right) = \left(\frac{5}{4} \right) (1 - 0,3638) = 1,25 \times 0,6362 = 0,79$$

Menurut Guilford reliabilitas instrumen 0,79 termasuk reliabilitas tinggi. Hasil analisis menunjukkan bahwa instrumen tes hasil belajar matematika telah memenuhi karakteristik yang memadai untuk digunakan pada penelitian.

Daya Bada Butir Tes Hasil Belajar Matematika

Indeks daya beda butir untuk tes non dikotomi (politomi) atau soal uraian dihitung dengan formula Ferguson sebagai berikut.

$$d = \frac{N^2 - \sum f_i^2}{N^2 - \frac{N^2}{n+1}} = \frac{(n+1)(N^2 - \sum f_i^2)}{nN^2} \quad (8)$$

Berdasarkan hasil analisis daya beda butir diperoleh bahwa instrumen yang disusun berkategori baik. Selengkapnya, dapat dilihat pada **Tabel 12** berikut.

TABEL 12. Hasil analisis daya beda butir

Butir	Indeks Daya Beda	Keterangan
1	0,866	Baik
2	0,843	Baik
4	0,813	Baik
5	0,776	Baik
6	0,798	Baik

Berdasarkan Tabel hasil analisis daya beda butir di atas, semua butir yang valid memiliki daya pembeda yang baik. Hal ini terjadi karena validitas dan daya pembeda memiliki ukuran yang hampir sama.

PEMBAHASAN

Dikembangkannya instrumen hasil belajar matematika pada siswa kelas V sekolah dasar bertujuan untuk mengetahui tingkat kesukaran butir soal. Validitas, dan reliabilitas dari instrument. Untuk memperoleh instrument yang baik, maka prosedur pengembangan instrument hasil belajar matematika dalam penelitian ini melewati prosedur sesuai dengan kaidah ilmiah. Adapun tahapan yang dilalui dalam penelitian ini sebelum analisis dhasil akhir data penelitian mengacu pada pendapat Mardapi yaitu (1) menentukan bentuk tes, yaitu tes esai, (2) menulis tes, (3) mereview dan merevisi pertanyaan, (4) uji coba lapangan, (5) mengumpulkan hasil tes, (6) penyekoran, dan (7) melaporkan hasil tes.

Berdasarkan hasil analisis data tingkat kesukaran butir dengan menggunakan program *MS Excel* diperoleh bahwa dari 7 butir soal yang ada masuk dalam kategori sedang yang ditunjukkan dalam Tabel 2. Dari 7 butir item yang dikembangkan tingkat kesukarannya berada pada rentangan 0,53 sampai dengan 0,66 pada kategori sedang. Hal ini menunjukkan bahwa butir item dengan kategori ini kebanyakan digunakan guru untuk peserta tes. Demikian halnya penelitian (Bayuni et al., 2013) taraf kesukaran butir tes diperoleh hasil yaitu kategori sedang 62,5%, kategori sukar 7,5%, dan kategori mudah 30% yang menunjukkan bahwa sebanyak 62,5 % butir soal dapat digunakan untuk kepentingan penelitian. Dengan demikian, soal yang mengukur berpikir tingkat tinggi adalah butir soal yang baik yang kebanyakan digunakan dalam mengukur hasil belajar matematika peserta didik.

Selanjutnya, untuk dapat menentukan reliabilitas butir soal maka terlebih dahulu dicarikan validitas butirnya. Berdasarkan hasil analisis terhadap hasil pekerjaan peserta didik diperoleh bahwa dari 7 butir soal yang ada diperoleh yang valid 5 butir soal yang ditunjukkan pada Tabel 7. Adapun 2 butir soal lainnya tidak valid. Dua butir soal yang tidak valid tersebut direvisi terkait kebahasaannya. Hal ini dikarenakan peserta didik kurang memahami petunjuk soal sehingga mereka melakukan kesalahan yang seharusnya tidak terjadi sehingga redaksi kalimat disederhanakan lagi sesuai karakteristik peserta didik kelas V sekolah dasar. Kurang lebih 70% dari 30 peserta tes melakukan kesalahan khususnya dalam menentukan hasil pengukuran volume. Sehingga dari 7 butir soal yang diujicobakan hanya ada 5 butir yang dinyatakan valid yaitu butir soal nomor 1, 2, 4, 5, dan 6 sedangkan butir soal no 3 dan 7 tidak valid sehingga harus direvisi.

Setelah validitas butir di ketahui, maka dilanjutkan dengan menentukan koefisien reliabilitasnya. Berdasarkan hasil analisis menggunakan formula *Alpha Cronbach*,

diperoleh bahwa koefisien reliabilitas butir item hasil belajar matematika yang dikembangkan adalah 0,79 berada pada kategori sangat tinggi. Hal ini menunjukkan bahwa kelima butir item tersebut dapat digunakan dalam penelitian. Seperti halnya pendapat Gregory, 2007 (Retnawati, 2017) mengatakan bahwa semakin besar reliabilitas suatu instrument, akan semakin kecil kesalahan pengukuran, demikian pula sebaliknya, semakin kecil reliabilitas skor, akan semakin besar kesalahan pengukurannya. Reliabilitas butir item dengan reliabilitas tinggi jika menunjukkan konsistensi hasil pengukuran yang dicapai oleh orang yang sama ketika mereka diuji ulang dengan tes yang sama pada kesempatan yang berbeda. Seperti halnya hasil penelitian (Hanifah, 2019) mengemukakan bahwa salah satu cara untuk mengetahui apakah peserta didik sudah memiliki keterampilan berpikir tingkat tinggi yaitu dengan cara melakukan penilaian. Penilaian yang berupa tes dapat digunakan untuk mengasah kemampuan berpikir peserta didik, dan berpengaruh dalam menentukan keterampilan berpikir peserta didik. Banyak penelitian memperlihatkan bahwa peserta didik di Indonesia memiliki kemampuan HOTS yang rendah dibandingkan dengan negara lainnya, salah satu faktor yang penyebabnya adalah peserta didik kurang terdapat dilatih untuk memiliki kemampuan berpikir tingkat tinggi, agar peserta didik dalam memahami materi yang dipelajari dengan baik. Oleh karena itu perlu sekiranya guru lebih mengotimalkan teknik penilaian yang berupa tes yang dapat digunakan untuk mengasah kemampuan berpikir peserta didik, dalam hal ini HOTS. Demikian halnya (Sari, 2015) dalam judul penelitian pengembangan soal matematika model PISA untuk mengetahui argumentasi siswa SMP, menyimpulkan bahwa (1) Telah dihasilkan prototipe perangkat soal matematika model PISA untuk mengetahui argumentasi siswa sebanyak 12 butir soal yang berbentuk uraian non objektif (open construct response). (2) Prototipe perangkat soal matematika model PISA untuk mengetahui argumentasi siswa yang dihasilkan memiliki efek potensial terhadap argumentasi siswa. Hal ini menunjukkan bahwa dalam melakukan pengukuran hasil belajar peserta didik khususnya dalam pembelajaran matematika di sekolah dasar dibiasakan untuk menggunakan soal-soal tes yang mengukur berpikir tingkat tinggi untuk mengasah kemampuan pemecahan masalah matematika.

SIMPULAN

Berdasarkan hasil penelitian dan pembahasan, maka dapat disimpulkan bahwa instrumen tes hasil belajar matematika bagi peserta didik kelas V sekolah dasar yang dikembangkan ini yang berorientasi pada High Order Thinking Skill (HOTS). Instrumen yang dikembangkan berupa soal tes HOTS dengan bentuk uraian. Validiasi instrumen penelitian yang dikembangkan dilakukan dengan menentukan validitas tes, reliabilitas instrumen tes, indeks kesukaran butir dan indeks daya beda butir dengan perolehan sebagai berikut: butir instrumen tes hasil belajar matematika yang dikembangkan ini memiliki tingkat kesukaran pada rentangan $0.30 \leq p \leq 0.70$ termasuk dalam kategori sedang; validitas butir tes hasil belajar matematika yang dikembangkan terdapat 5 butir yang valid pada kategori sedang sedangkan 2 butir lainnya direvisi dari aspek kebahasaannya; koefisien reliabilitas instrumen tes hasil belajar matematika berada pada rentangan $0,60 < r_{11} \leq 0,80$ yaitu 0,79 pada kategori tinggi; indeks daya beda butir soal tes hasil belajar matematika yang dikembangkan ini semuanya berkategori baik; dengan demikian butir soal tes hasil belajar matematika yang dikembangkan yang berorientasi pada High Order Thinking Skill bagi peserta didik kelas V sekolah dasar ini dapat diandalkan dan dapat digunakan sebagai instrumen pengumpulan data penelitian.

DAFTAR PUSTAKA

1. 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: Addison Wesley Longman, Inc.

2. Arends, R. I. (2013). *Learning to teach* (9th ed.). Copyright © by McGraw-Hill Education.
3. Arifin, Z., & Retnawati, H. (2017). Pengembangan instrumen pengukur higher order thinking skills matematika siswa SMA kelas X. *Pythagoras: Jurnal Pendidikan Matematika*, 12(1), 98–108. <https://doi.org/10.21831/pg.v12i1.14058>
4. Bayuni, K. H., Candiasa, I. M., & Koyan, I. W. (2013). Pengembangan Tes Matematika dengan Teknik Part-Whole pada Siswa SD Kelas IV Se-Kecamatan Gianyar. *E-Journal Program Pascasarjana Universitas Pendidikan Ganesha*, 3(2), 1–21.
5. Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. Alexandria: ASCD.
6. Candiasa, I. M. (2010). *Pengujian instrumen penelitian disertai aplikasi Iteman dan Bigsteps*. Singaraja: Undiksha Press.
7. Darminto, B. P. (2013). Improving the Ability of Students' Mathematical Problem Solving. *Jurnal Pendidikan Matematika Dan Sains Tahun*, 1(2), 101–107.
8. Gunawan, I., & Palupi, A. R. (2012). Revisi taksonomi Bloom ranah kognitif: kerangka landasan untuk pembelajaran, pengajaran, dan asesmen. *Premiere Educandum: Jurnal Pendidikan Dasar Dan Pembelajaran*, 2(2), 98–117.
9. Hanifah, N. (2019). Pengembangan instrumen penilaian Higher Order Thinking Skill (HOTS) di sekolah dasar. *Conference Series*, 1(1), 1–8. Retrieved from <http://ejournal.upi.edu/index.php/crecs/article/view/14286>
10. Johar, R. (2012). Domain Soal PISA untuk Literasi Matematika. *Jurnal Peluang*, 1(1), 30–41.
11. Kemendikbud. (2017). *Modul Penyusunan Higher Order Thinking Skill (HOTS)*. Jakarta: Direktorat Jenderal Pendidikan Dasar dan Menengah Departemen Pendidikan dan Kebudayaan.
12. Mahirah, B. (2017). Evaluasi belajar peserta didik (siswa). *Jurnal Idaarah*, 1(2), 257–267.
13. Mardapi, D., Kunaidi, & Kartowagiran, B. (2011). Pengembangan instrumen pengukur hasil belajar nirbias dan terskala baku. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 15(2), 326–341.
14. Ndiung, S., Dantes, N., Ardana, I. M., & Marhaeni, A. (2019). Treffinger Creative Learning Model with RME Principles on Creative Thinking Skill by Considering Numerical Ability. *International Journal of Instruction*, 12(3), 731–744. <https://doi.org/doi:10.29333/iji.2019.12344a>
15. Newman, I., Lim, J., & Pineda, F. (2013). Content Validity using Mixed Methods Approach: Its application and development through the use of a Table of specifications methodology. *Journal of Mixed Methods Research*, 7(3), 243–260. <https://doi.org/10.1177/1558689813476922>
16. OECD. (2013). *PISA 2012 assessment and analytical framework: mathematics, reading, science, problem solving and financial literacy*. German: OECD Publishing.
17. OECD. (2019). *PISA 2018 insights and interpretations*. Retrieved from [https://www.oecd.org/pisa/PISA 2018 Insights and Interpretations final PDF.pdf](https://www.oecd.org/pisa/PISA%2018%20Insights%20and%20Interpretations%20final%20PDF.pdf)
18. Pandra, V., & Mardapi, D. (2017). Development of Mathematics Achievement Test for Third Grade Students at Elementary School in Indonesia. *International Electronic Journal of Mathematics Education*, 12(8), 769–776.
19. Partnership for 21st Century learning. (2011). *P21 Framework Definition*. Retrieved from <http://www.p21.org/our-work/p21-framework>
20. Retnawati, H. (2017). Validitas dan reliabilitas konstruk skor tes kemampuan calon mahasiswa. *Jurnal Ilmu Pendidikan*, 23(2), 126–135.
21. Sari, E. F. P. (2015). *Pengembangan soal matematika model PISA untuk mengetahui argumentasi siswa di sekolah menengah pertama*. 124–147. Retrieved from <http://ejournal.unsri.ac.id>
22. Setianingsih, R., Sa'dijah, C., & Rahman, A. (2017). Investigating Fifth- Grade Students' Construction of Mathematical Knowledge through Classroom Discussion.

- International Electronic Journal of Mathematics Education*, 12(3), 383–396.
23. Simalango, M. M., Darmawijoyo, & Aisyah, N. (2018). Analisis kesulitan siswa dalam menyelesaikan soal-soal PISA taun 2012 level 4, 5, dan 6 di SMPN 1 Indralaya. *Jurnal Pendidikan Matematika*, 12(1), 43–58.
 24. Sudaryono. (2011). Implementasi Teori Responsi Butir (Item Response Theory) pada Penilaian Hasil Belajar Akhir di Sekolah. *Jurnal Pendidikan Dan Kebudayaan*, 17(6), 719–732.
 25. Tryanasari, D., Aprilia, S., & Cahya, W. A. (2017). Pembelajaran literasi di SDN Rejosari 1 kecamatan Kawedan kabupaten Magetan. *Premiere Educandum: Jurnal Pendidikan Dasar Dan Pembelajaran*, 7(2), 173–179. <https://doi.org/10.25273/pe.v7i2.1641>
 26. Van De Walle, J. A. (2013). *Elementary and Middle School Mathematics: Teaching Development* (8th ed.). Pearson Education, Inc.
 27. Yani, A., Asri, A. F., & Burhan, A. (2013). Analisis tingkat kesukaran, daya pembeda, dan fungsi distraktor soal ujian semester ganjil. *E-Journal Universitas Sriwijaya*, 98–115.
 28. Yusup, F. (2018). Uji validitas dan reliabilitas instrumen peelitian kuantitatif. *Jurnal Tarbiyah: Jurnal Ilmiah Kependidikan*, 7(1), 17–23.
 29. Yuwono, A. (2016). Problem solving dalam pembelajaran matematika. *Union: Jurnal Pendidikan Matematika*, 4(1), 6–7.

PROFIL SINGKAT

Sabina Ndiung adalah dosen program studi pendidikan guru sekolah dasar, fakultas keguruan dan ilmu pendidikan, Universitas Katolik Indonesia (Unika) Santu Paulus Ruteng. Ia juga merupakan editor dari jurnal literasi pendidikan dasar (JLPD) Unika Santu Paulus Ruteng.

Mariana Jediut adalah dosen program studi pendidikan guru sekolah dasar, fakultas keguruan dan ilmu pendidikan, Universitas Katolik Indonesia Santu Paulus Ruteng.