

PENGEMBANGAN INSTRUMEN *HIGHER ORDER THINKING SKILLS* (HOTS) PADA TEMA 7 INDAHNYA KERAGAMAN DI NEGERIKU SUBTEMA 1 SD KELAS IV

D.W.A.T. Udayani¹, I.M. Candiasa², I.M.C. Wibawa²

¹²³Program Studi Pendidikan Dasar
Universitas Pendidikan Ganesha
Singaraja, Indonesia

e-mail: ayu.tini@undiksha.ac.id¹, candiasa@undiksha.ac.id²,
imadecitra.wibawa@undiksha.ac.id³

Abstrak

Penelitian ini bertujuan menghasilkan instrumen tes HOTS Tema 7 Indahny Keragaman di Negeriku Subtema 1 SD kelas IV dan instrumen tes HOTS yang memenuhi kelayakan validitas, validasi isi dapat dilihat dari kisi-kisi atau spesifikasi instrumen yaitu dengan cara mengkaji butir-butir instrumen yang digunakan, apakah sudah mewakili materi secara keseluruhan atau belum representative (Koyan, 2011), reliabilitas, uji reliabilitas digunakan untuk menguji atau mengetahui keajegan alat ukur tersebut apabila alat ukur tersebut diberikan secara berulang kali pada objek atau subjek yang sama hasilnya akan relatif sama (Yusuf, 2015), daya beda, uji daya beda adalah kemampuan pada butir tes dalam membedakan antara peserta didik yang berkemampuan tinggi dan siswa yang memiliki kemampuan rendah (Arikunto, 2015). tingkat kesukaran dan *Partial Credit Model* (PCM), Sumintono (2015) model PCM mengakomodasi analisis butir yang antara satu dengan lainnya memiliki bobot yang berbeda dapat digunakan pada data politomi yang dihasilkan dengan total skor majemuk atau berbeda-beda dikembangkan pada skala politomi yang memiliki pola skor x sebesar $0,1,2,3,\dots, mi$. Pendekatan penelitian adalah *Research and Development* (R&D). *Tryout* dilakukan di SD Gugus VI Sukawati menggunakan 30 responden dalam uji *Small Group* dan 60 responden *field test*. Kisi-kisi instrumen berisi muatan pelajaran Bahasa Indonesia, Ilmu Pengetahuan Alam (IPA), Ilmu Pengetahuan Sosial (IPS), dan Seni Budaya dan Prakarya yang terdiri dari satu Kompetensi Dasar (KD). Kompetensi dasar diturunkan menjadi 10 indikator 2 indikator C4/K1, 1 indikator C4/K2, 3 indikator C4/K4, 1 indikator C5/K2 dan 3 indikator C5/K4. Instrumen dikembangkan dalam bentuk tes soal uraian menentukan kualitas instrumen tersebut harus melalui beberapa uji coba. Uji validitas pakar formula Gregory nilai 1 kriteria baik.

Kata kunci: *Higher Order Thinking Skills (HOTS)*; Instrumen Tes; Sekolah Dasar

Abstract

This study aimed to produce a HOTS test instrument on Theme 7, The Beauty of Diversity in My Nation, Sub-theme 1 Elementary School grade IV, and a HOTS test instrument that meets eligibility validity. Content validation can be seen from the grid or instrument specifications by examining the instrument items used, whether they represent the material as a whole or not representative (Koyan, 2011). Furthermore, reliability. The reliability test is used to test or determine the constancy of the measuring instrument. If the measuring instrument is given repeatedly to the same object or subject, the results will be relatively the same (Yusuf, 2015). Different power. Different power test is the ability of the test items to distinguish between students who have high abilities and students who have low abilities (Arikunto, 2015). Difficulty level and Partial Credit Model (PCM). Sumintono (2015) said that the PCM model accommodates the items analysis which has different weights with one another. It can be used on polytomy data produced with a multiple or different total score developed on a polytomy scale with a pattern of x scores of $0, 1, 2, 3,\dots, mi$. The research approach was Research and Development (R&D). The tryout was conducted at Gugus VI Sukawati Elementary School using 30 respondents in the Small Group test and 60 respondents in the field test. The instrument grid contained lessons in the Indonesian Language, Natural Sciences (IPA), Social Sciences (IPS), and Cultural Arts and Crafts, which consist of one Basic Competency (BC). Basic competence is decreased to 10 indicators 2 indicators C4/K1, 1 indicator C4/K2, 3 indicators C4/K4, 1 indicator C5/K2 and 3 indicators C5/K4. The instrument was developed in the form of a test item

description to determine the instrument quality, which must go through several trials. Test the validity of the expert formula Gregory, value 1 is a good criterion.

Keywords : *Higher Order Thinking Skills (HOTS); Test Instrument; Elementary School*

PENDAHULUAN

Sesuai dengan tuntutan abad 21 yang penuh dengan tantangan serta persaingan sehingga kemampuan HOTS kemampuan berpikir tingkat tinggi peserta didik diasah. Berdasarkan hasil *Studi Programme for Internasional 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 (Ndiung & Jediut, 2020). Hasil penelitian PISA juga menunjukkan rendahnya kemampuan berfikir kritis peserta didik di Indonesia. Soal-soal serta permasalahan yang diujikan oleh PISA adalah permasalahan yang bersifat kontekstual yang dialami pada kehidupan keseharian peserta didik, yang bermuatan kognitif C4 (menganalisis), C5 (mengevaluasi) dan C6 (mencipta). Pada pelaksanaannya soal-soal yang diujikan oleh PISA mampu diselesaikan oleh peserta didik dengan kemampuan berpikir kritis yang baik.

Hal ini sejalan dengan kemampuan berpikir tingkat tinggi dapat dilatih dalam proses pembelajaran di kelas. Pembelajaran akan bermakna jika peserta didik diajak berpikir tingkat tinggi, keberhasilan penguasaan suatu konsep akan didapatkan ketika peserta didik sudah mampu berpikir tingkat tinggi (Ndiung & Jediut, 2020). HOTS merupakan cara berpikir siswa pada tingkat kognitif tinggi dibentuk dari beragam konsep dan metode kognitif serta taksonomi pembelajaran seperti metode *problem solving* (Muthoharoh, 2020). Agar kemampuan HOTS peserta didik dapat berkembang dengan baik, maka peserta didik 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 mengkreasi 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.

Pada pembelajaran guru melatih siswa berupa kemampuan serta keterampilan berpikir tingkat tinggi atau *Higher Order Thinking Skill (HOTS)*. Sehingga penerapan HOTS pada latihan soal peserta didik memiliki kelebihan dalam kemampuan berpikir nalar untuk menjawab pertanyaan-pertanyaan yang lebih rumit dan atau memecahkan suatu kasus masalah yang lebih rumit (Dasar, 2016). Upaya dalam penerapan HOTS kepada siswa dengan adanya pembiasaan dengan kegiatan yang melatih HOTS itu sendiri, guru harus memberikan soal-soal atau latihan yang mampu meningkatkan HOTS peserta didik (Suhady, 2020). Pada penerapan HOTS siswa di arahkan untuk belajar lebih aktif serta mampu meningkatkan berkemampuan berpikir tingkat tinggi (Riza, 2021). Aktivitas HOTS melatih kemampuan siswa dalam menganalisa gagasan atau ide dalam merancang sebuah hipotesis, mengkonstruksi informasi yang diperoleh dengan memberikan penjelasan-penjelasan terhadap informasi yang kompleks. Proses berfikir secara kritis dan kreatif dengan HOTS dapat mendorong siswa dalam melakukan penalaran tingkat tinggi sehingga siswa tidak terpaku pada satu pola jawaban yang dihasilkan dari proses menghafal dan adanya konsep ilmunya (Chuseri et al., 2021).

Kemampuan berpikir tingkat tinggi (HOTS) dipandang sebagai aspek penting dalam proses pembelajaran, kemampuan berpikir tingkat tinggi (HOTS) meningkatkan kemampuan siswa dalam bersaing di jenjang pendidikan lebih tinggi serta menyiapkan siswa untuk bersaing di masa depan sehingga dapat meningkatkan kemampuan berpikir tingkat tinggi (HOTS) siswa, pendidik memiliki

peranan penting untuk menjadikan peserta didik berada pada level tersebut (Sani, 2019). HOTS salah satu bentuk instrumen tes yang dapat dimanfaatkan dalam mengetahui kemampuan berpikir tingkat tinggi, yaitu kemampuan berpikir yang tidak hanya mengingat, menyatakan kembali, atau merujuk. Soal-soal HOTS pada konteks asesmen mengukur kemampuan: 1) transfer antar konsep, 2) memproses dan menerapkan informasi, 3) mengkaitkan berbagai informasi, 4) menggunakan informasi untuk menyelesaikan masalah, dan, (5) menelaah ide dan informasi secara kritis. Sejalan dengan hal ini mulai tahun 2021 pemerintah telah membuat suatu program yaitu asesmen nasional yang disitilahkan dengan AKM.

AKM (Asesmen Kompetensi Minimum) menjadi salah satu diantara 4 kebijakan program merdeka belajar yang bertujuan untuk mengukur kompetensi berpikir atau bernalar siswa ketika membaca teks (literasi) dan menghadapi persoalan yang membutuhkan pengetahuan matematika (numerasi). AKM pada jenjang sekolah dasar diikuti oleh siswa yang berada pada kelas V. oleh karena itu wali kelas mengupayakan peserta didik agar mampu terbiasa sedini mungkin terlatih dalam pemahaman berfikir tingkat tinggi misalnya adalah dengan memberikan soal berjenis HOTS. Adapun rumusan masalah dalam penelitian ini adalah: bagaimana instrument tes HOTS tema 7 Indahya keragaman di negeriku subtema 1 SD kelas IV. Adapun tujuan penelitian ini adalah: mengetahui instrument tes HOTS tema 7 Indahya keragaman di negeriku subtema 1 SD kelas IV.

METODE

Metode penelitian yang digunakan dalam penelitian ini adalah *Research and Development* (R&D). Metode penelitian dan pengembangan merupakan metode penelitian dengan tujuan menghasilkan produk serta menguji keefektifan produk (Sugiyono, 2007). Rancangan penelitian ini mengikuti metode penelitian pengembangan yang terdiri atas 2 komponen utama yaitu, (1) model

pengembangan dan (2) prosedur pengembangan.

Metode pengumpulan data terdiri dari pedoman wawancara, lembar validasi ahli, serta instrumen tes HOTS tema 7 Indahya keragaman di negeriku subtema 1 SD kelas IV. Metode analisis data terdiri dari uji validitas, reliabilitas, daya beda, indeks kesukaran dan teori respon butir moderen *Partial Credit Model* (PCM).

Pada analisis validitas dibantu dengan aplikasi Ministep dengan model Rasch (model 1P), reliabilitas, daya beda dan indeks kesukaran instrumen dibantu program komputer *Microsoft Office Excel 2016*. Sedangkan dalam teori respon butir *Partial Credit Model* (PCM) berbantuan aplikasi winstep.

Validasi isi dapat dilihat dari kisi-kisi atau spesifikasi instrumen yaitu dengan cara mengkaji butir-butir instrumen yang digunakan, apakah sudah mewakili materi secara keseluruhan atau belum representative (Koyan, 2011). Instrumen instrumen tes HOTS tema 7 Indahya keragaman di negeriku subtema 1 SD kelas IV divalidasi oleh 2 orang dosen ahli dan 3 orang guru SD kelas IV ahli. Rumus validitas isi menggunakan *formula Gregory* sebagai berikut.

$$\text{Validitas Isi} = \frac{D}{(A+B+D)} = \frac{10}{(0+0+10)} = 1 \quad (1)$$

Koefisien validitas isi 10 butir soal uraian Tema 7 Indahya Keragaman di Negeriku Subtema 1 adalah 1. Hal ini berarti validitas isi tes sangat tinggi.

Pada uji *One to One* dilaksanakan bertujuan untuk mengetahui tingkat kepraktisan soal dengan mengujikan instrumen kepada tiga orang pesertadidik kemudian peserta didik diarahkan mengisi kuesioner. Selanjutnya dianalisis dengan menggunakan rumus kepraktisan. Berikut adalah hasil analisis dari uji *One to One* pada tiga panelis dengan taraf kemampuan peserta didik yang heterogen berdasarkan hasil wawancara guru wali kelas, dengan rata-rata = 92%. Berdasarkan hasil tersebut butir instrumen dapat dikatakan sangat praktis.

Pengujian validitas butir instrumen dilakukan dengan uji *Small Group* kepada 30 orang peserta didik SD kelas IV dengan aplikasi Ministep dengan model

Rasch (model 1P), menghasilkan *output* tabel Nilai *Outfit MNSQ ZSTD* serta *Pt Measure Corr*. Menganalisis validitas instrument dikembangkan melihat nilai *Outfit MNSQ, ZSTD* serta *Pt Measure Corr* (Yulianto & Widodo, 2020).

Analisis pengujian validitas instrumen pada penelitian ini menggunakan pemodel Rasch, hal yang dilihat yaitu tingkat validitas respon pada item berdasarkan: (1) Nilai *Outfit Mean Square* (MNSQ) yang diterima $0,5 < \text{MNSQ} < 1,5$, (2) Nilai *Outfit Z-Standard* (ZSTD) kesesuaian nilai uji z yang diterima $-2,0 < \text{ZSTD} < + 2,0$, (3) *Point Measure Correlation* (Pt Mean Corr) $0,4 < \text{Pt Mean Corr} < 0,85$. Seluruh butir soal dapat dinyatakan valid atau diterima, sesuai dengan 3 kriteria, maka tidak dilaksanakan perbaikan.

Uji reliabilitas merupakan keterandalan alat ukur atau kejegan alat ukur, artinya kapanpun alat ukur tersebut digunakan akan menghasilkan hasil ukur yang relative tetap. Reliabilitas digunakan untuk menguji atau mengetahui keajegan alat ukur tersebut apabila alat ukur tersebut diberikan secara berulang kali pada objek atau subjek yang sama hasilnya akan relatif sama (Yusuf, 2015). Uji reliabilitas dilakukan hanya pada butir soal yang telah diuji validitas butir tes dan telah dinyatakan *valid*, dengan demikian uji reliabilitas bisa dilakukan setelah dilakukannya uji validitas. Uji reliabilitas pada instrument yang bersifat politomi pengujian reliabilitas menggunakan *formula Alpha Cronbach* Adapun rumusnya adalah sebagai berikut.

$$r_{11} = \left(\frac{k}{k-1} \right) \left(1 - \frac{\sum Si^2}{St^2} \right) \quad (2)$$

(Sudijono, 2011)

Uji daya beda daya beda adalah kemampuan pada butir tes dalam membedakan antara peserta didik yang berkemampuan tinggi dan siswa yang memiliki kemampuan rendah (Arikunto, 2015). Analisis daya beda merupakan pengkajian butir tes yang dimaksudkan bertujuan mengetahui kesanggupan butir tes membedakan siswa yang tergolong mampu dengan peserta didik yang tergolong tidak mampu (Candiasa, 2015).

Indeks diskriminasi butir soal uraian dapat dihitung dengan rumus perbandingan antara selisih mean kelompok atas dan mean kelompok bawah dengan skor maksimal tiap butir soal (Bajpai & Bajpai, 2014) yaitu, *Diskrimination Indkes* :

$$\frac{\text{Average upper group} - \text{Average lower group}}{\text{Full item score}} \quad (3)$$

Pada umumnya uji taraf kesukaran butir tes dilakukan dengan memperhitungkan banyak responden yang menjawab butir tersebut dengan benar. Rumus yang digunakan adalah:

$$I = \frac{\sum U + \sum L - (2N \times S_{min})}{2N (S_{max} - S_{min})} \quad (4)$$

(Candiasa, 2011)

Untuk menentukan kelompok atas dan kelompok bawah maka terlebih dahulu seluruh skor peserta tes diranking dari yang terbesar hingga yang terkecil, selanjutnya ditentukan 27% skor tertinggi sebagai kelompok atas dan 27% skor terendah sebagai kelompok bawah. Dalam penelitian ini untuk uji taraf kesukaran butir tes menggunakan bantuan *Microsoft Office Excel*.

Pada penelitian ini menggunakan analisis dapat yang meliputi validitas isi, reliabilitas tes, analisis daya beda, taraf kesukaran butir tes, dan probabilitas respons siswa dengan tingkat kemampuan dalam memperoleh skor tertentu pada tiap butir soal, untuk analisis yaitu *Partial Credit Model* (PCM) menggunakan aplikasi *Ministep*. Menurut Sumintono (2015) model PCM mengakomodasi analisis butir yang antara satu dengan lainnya memiliki bobot yang berbeda dapat digunakan pada data politomi yang dihasilkan dengan total skor majemuk atau berbeda-beda dikembangkan pada skala politomi yang memiliki pola skor x sebesar $0, 1, 2, 3, \dots, mi$.

Sehingga peluang seorang peserta pada tingkat kemampuan θ meraih skor x diatas x-1 dapat dihitung dengan persamaan sebagai berikut (Hambleton., 1985):

$$\frac{P_{ix}(\theta)}{P_{ix}(\theta)+P_{ix}(\theta)} = \frac{\exp(D(\theta-b_i))}{1+\exp(D(\theta-b_i))} \text{ untuk } x = 0, 1, 2, 3, \dots, m_i \quad (5)$$

$P_{ix}(\theta)$ dan $P_{ix}(\theta)$ mengacu pada peluang peserta siswa dengan kemampuan θ meraih skor x dan $x-1$. Sehingga peluang siswa dengan kemampuan θ untuk memperoleh skor x pada item i dengan tingkat kesukaran item sebesar δ dapat dituliskan sebagai berikut:

$$P_{nix} = \frac{1}{1+\exp(\theta_n - \delta_{ij})} \text{ untuk } x = 0$$

$$P_{nix} = \frac{\exp(\theta_n - \delta_{ij})}{1+\exp(\theta_n - \delta_{ij})} \text{ untuk } x = 0 \quad (6)$$

PCM pada skala politomi yang merupakan pengembangan dari model Rasch pada skala dikotomi. Asumsi pada PCM yakni setiap butir mempunyai daya beda yang sama, namun indeks kesukaran dalam setiap langkah tidak perlu terurut, suatu langkah dapat lebih sulit dibandingkan langkah berikutnya. Skor kategori pada PCM menunjukkan banyaknya langkah untuk menyelesaikan dengan benar butir tersebut. Skor yang lebih tinggi menunjukkan kemampuan yang lebih besar dari pada skor kategori yang lebih rendah. Jika i adalah butir politomi dengan kategori skor x sebesar $0, 1, 2, \dots, m_i$.

Probabilitas respons individu n dengan tingkat kemampuan θ memperoleh skor kategori j pada butir soal i , dinyatakan oleh (Masters, 2016). dengan persamaan, sebagai berikut:

$$P_{nix}(\theta) = \frac{\exp \sum_{j=0}^x (\theta - \delta_{ij})}{\sum_h^{m_i} \exp \sum_{j=0}^h (\theta - \delta_{ij})} \text{ untuk } x = 0, 1, 2, \dots, m_i \quad (7)$$

HASIL DAN PEMBAHASAN

Prosedur Pengembangan Instrumen tes HOTS tema 7 Indahya Keragaman di Negeriku Subtema 1 SD kelas IV

(1) Tahap *Preliminary* pada tahap *preliminary* dilaksanakan pengkajian terhadap sumber referensi yang sesuai dengan penelitian ini. Mengumpulkan beberapa informasi serta teori, akan dilangsungkan kegiatan penentuan tempat

penelitian serta subjek yang akan di uji coba dengan menghubungi kepala gugus serta menuju kepada kepala sekolah dari masing-masing sekolah dan guru kelas IV di sekolah yang akan dijadikan lokasi uji coba serta melaksanakan observasi dan wawancara terhadap guru kelas IV mengenai kegiatan pembelajaran tema 7 Indahya Keberagaman di Negeriku Subtema 1 yang telah terlaksana pada muatan pelajaran Pendidikan Pancasila dan Kewarganegaraan (PKN), Bahasa Indonesia, Ilmu Pengetahuan Alam (IPA), Ilmu Pengetahuan Sosial (IPS) dan Dasar Seni Budaya dan Prakarya (SDP) mengenai permasalahan yang dihadapi serta khususnya mengenai tersedianya latihan soal-soal HOTS yang digunakan pada pembelajaran tematik. Dalam penelitian ini pada tahap melihat potensi masalah melaksanakan wawancara langsung dengan guru kelas IV bertujuan agar memperoleh informasi mengenai Instrumen tes yang umumnya digunakan. Jenis data yang digunakan yaitu data kuantitatif. Data kuantitatif merupakan jenis data yang dapat diukur secara langsung, yang berupa informasi atau penjelasan yang dinyatakan dengan bilangan atau berbentuk angka (Sugiyono, 2009). Pada penelitian ini data kuantitatif yang digunakan adalah hasil skor perolehan siswa setelah uji coba produk yang nantinya digunakan dalam perhitungan mencari validitas butir, reliabilitas tes, taraf kesukaran butir dan daya beda butir, serta probabilitas siswa menjawab benar.

(2) Tahap *Self Evaluation* pada tahap *Self Evaluation* dilaksanakan penilaian oleh diri sendiri terhadap desain instrument tes HOTS Tema 7 Indahya Keberagaman di Negeriku Subtema 1 kelas IV yang akan dibuat oleh peneliti. Pada tahapan ini meliputi: Analisis, pada langkah ini penelitian melaksanakan analisis siswa, analisis kurikulum dan analisis perangkat atau bahan yang akan dikembangkan. Desain, pada tahapan ini kegiatan yang dilaksanakan yaitu peneliti mendesain kisi-kisi soal instrument HOTS Tema 7 Indahya Keberagaman di Negeriku Subtema 1, soal-soal instrument HOTS muatan pelajaran Pendidikan Pancasila dan Kewarganegaraan (PKN), Bahasa Indonesia, Ilmu Pengetahuan

Alam (IPA), Ilmu Pengetahuan Sosial (IPS) dan Dasar Seni Budaya dan Prakarya (SBDP) dan kunci jawaban Intrumen.

Berikut penjelasan lebih lanjut mengenai tahapan desain yaitu: Menyusun kisi-kisi soal pada tahapan menyusun kisi-kisi memiliki tujuan agar perangkat tes yang akan disusun tidak menyimpang dari bahan atau materi serta aspek tes yang akan diukur. Dalam hal ini, kisi-kisi penyusunan soal HOTS Tema 7 Indahya Keberagaman di Negeriku bertujuan membantu dalam menulis butir soal yang tepat untuk mengukur kemampuan HOTS siswa. Adapun langkah-langkah dalam menyusun kisi-kisi adalah seperti berikut. (a) Menentukan Kompetensi Dasar (KD), (b) Memiliki materi pokok yang berkaitan dengan KD yang akan diuji, (c) Merumuskan Indikator soal, (d) Menentukan level kognitif dalam penyusunan kisi-kisi soal, (d) Menulis butir pertanyaan sesuai dengan kisi-kisi soal.

Butir-butir pertanyaan ditulis sesuai dengan kaidah penulisan butir soal HOTS pada muatan pelajaran Pendidikan Pancasila dan Kewarganegaraan (PKN), Bahasa Indonesia, Ilmu Pengetahuan Alam (IPA), Ilmu Pengetahuan Sosial (IPS) dan Dasar Seni Budaya dan Prakarya (SBDP). Kaidah penulisan butir soal HOTS Tema 7 Indahya Keberagaman di Negeriku Subtema 1 berbeda dari kaidah penulisan lainnya. Perbedaannya terlihat pada level kognitif yang digunakan, sedangkan pada aspek konstruksi dan bahasa relatif sama.

Membuat pedoman penskoran (rubrik) atau kunci jawaban. Setiap butir soal yang ditulis hendaknya dilengkapi dengan pedoman penskoran atau kunci jawaban. Desain produk ini sebagai prototype. Masing-masing prototype fokus pada tiga karakteristik yaitu: konten, konstruksi dan bahasa.

(3) Tahap *Prototyping* (Validasi, Evaluasi dan Revisi). Pada tahap *prototyping* produk yang telah dibuat atau didesain akan dievaluasi. Tahap evaluasi ini produk akan diuji cobakan dalam 3 kelompok, yaitu *Expert Review*, *One-to one*, dan *small group*. Hasil desain pada prototype pertama yang akan dikembangkan atas dasar *self evaluation*

diberikan pada pakar (*Expert review*), peserta didik *one-to-one* dan *small group* secara paralel. Dari hasil ketiganya dijadikan bahan untuk melaksanakan revisi.

Pada validitas isi yang akan dilaksanakan, butir tes yang telah disusun menyesuaikan dengan kisi-kisi tes hasil belajar Tema 7 Indahya Keberagaman di Negeriku Subtema 1 selanjutnya ditelaah oleh pakar/ahli yang dapat menghasilkan pendapat berupa relevan atau tidak relevannya suatu item. Penelaahan instrumen oleh pakar atau ahli dilaksanakan dengan tujuan melihat keselarasan indikator dengan tujuan pengembangan instrument, kesesuaian indikator dengan pokok materi, melihat kesesuaian instrument dengan indikator butir, melihat ketepatan konsep butir soal, melihat kebenaran isi, ketepatan kunci, bahasa dan budaya (Retnawati, 2016).

(4) Tahap Uji Coba Lapangan (*Field Test*). *One to one* pada tahapan *one to one* peneliti memanfaatkan 3 orang siswa sebaya non subjek uji coba sebagai tester. Ketiga siswa yang menjadi tester terbagi menjadi 3 yakni siswa dengan kemampuan rendah, siswa dengan kemampuan sedang dan siswa dengan kemampuan tinggi. Berdasarkan komentar ketiga siswa tersebut desain produk yang telah dibuat direvisi dan diperbaiki.

Kelompok Kecil (*Small Group*), hasil revisi dari *expert* dan saran di *one-to-one* pada prototype pertama dijadikan dasar untuk merevisi desain prototype kedua. Pada tahap ini dilakukan uji coba pada kelompok kecil terdiri dari 3 orang, karakteristik siswa terdiri dari siswa dengan kemampuan tinggi, siswa dengan kemampuan sedang dan siswa dengan kemampuan rendah. Berdasarkan hasil tes dan tanggapan siswa inilah soal direvisi dan diperbaiki lagi. Hasil dari tahap ini diharapkan menghasilkan instrument tes yang mampu mengukur HOTS Tema 7 Indahya Keberagaman di Negeriku Subtema 1 siswa kelas IV sekolah dasar.

Validitas, Reliabilitas, Daya Beda, Tingkat Kesukaran dan *Partial Credit Model* (PCM)

Kisi-Kisi instrumen tes HOTS tema 7 Indahnya Keragaman di Negeriku Subtema 1 SD kelas IV

Bertujuan untuk menyusun kisi-kisi instrumen tes HOTS Tema 7 Indahnya Keragaman di Negeriku Subtema 1 SD kelas IV dan menyusun instrumen tes HOTS Tema 7 Indahnya Keragaman di Negeriku Subtema 1 SD kelas IV yang memenuhi standar kualitas tes. Pada penyusunan kisi-kisi instrumen tes HOTS yaitu berupa soal uraian.

Berdasarkan hasil prasurvei berupa wawancara serta observasi yang telah dilaksanakan kepada guru peserta didik kelas IV di SD Negeri Gugus VI Sukawati di kabupaten Gianyar, masih banyak guru yang menggunakan soal dengan jenis soal LOTS. Soal hanya mampu mengukur kemampuan dasar peserta didik dengan soal berkategori LOTS. Guru mengemukakan bahwa belum memahami secara utuh mengenai soal berjenis HOTS, karena belum banyaknya contoh instrumen tes HOTS juga menjadi bentuk kendala utama yang dihadapi guru, mengakibatkan peserta didik dalam pelaksanaan simulasi AKM masih dalam bimbingan guru terkait dengan penyelesaian soal kemampuan bernalar dan berpikir tingkat tinggi atau soal HOTS.

Pengembangan item berbasis HOTS berdampak baik pada peserta didik, pada pelaksanaan pembelajaran tentunya peran guru yang memiliki pemahaman baik tentang proses kognitif dan ketrampilan berfikir tingkat tinggi (Pratiwi et al., 2019). Agar HOTS peserta didik dapat berkembang dengan baik, peserta didik dapat dibiasakan dalam menyelesaikan instrumen HOTS, hal ini menyebabkan potensi HOTS dalam diri peserta didik berkembang (Retnawati et al., 2018). Pada proses pembelajaran serta evaluasi HOTS dapat dimaknai sebagai kemampuan proses berfikir kompleks yang mencakup menguraikan materi, mengkritisi dan menciptakan solusi pemecahan masalah (Budiarta et al., 2018).

Pada penyusunan instrument HOTS yang diperlukan adalah menyusun kisi-kisi. Kisi-kisi merupakan suatu format atau matriks memuat mengenai syarat-syarat tentang soal-soal yang hendak disusun (Syofyan, 2016). Kisi-kisi dapat diartikan test *blue-print* atau *table of specification* mengenai deskripsi kompetensi dan materi yang akan diujikan. Wujudnya kisi-kisi adalah sebuah tabel yang memuat tentang perperincian materi dan tingkah laku beserta imbalan serta proporsi yang dikehendaki oleh penilai, tiap kotak diisi dengan bilangan yang menunjukkan jumlah soal (Suharsimi, 2018). Kisi-kisi instrumen tes HOTS merupakan format yang mencantumkan kriteria yang diperlukan dalam menyusun instrumen tes HOTS. Tujuan penyusunan kisi-kisi adalah untuk menentukan ruang lingkup materi dan sebagai petunjuk dalam menulis instrument soal (Syofyan, 2016). Kisi-kisi yang baik idealnya mencakup 3 hal yakni memuat isi silabus, komponen dijabarkan dengan jelas serta materi yang dipilih bisa dibuatkan soalnya (Tanujaya et al., 2017).

Kisi-kisi instrumen tes HOTS pada penelitian ini dibuat berdasarkan Kompetensi Inti (KI) dan Kompetensi dasar (KD) pada kurikulum 2013. Kompetensi Inti (KI) dan Kompetensi dasar (KD) mewakili dari masing-masing muatan materi pelajaran pada tema Tema 7 Indahnya Keberagaman di Negeriku Subtema 1 kelas IV semester 2 yaitu Bahasa Indonesia, Ilmu Pengetahuan Alam (IPA), Ilmu Pengetahuan Sosial (IPS) dan SBDP. Penyusunan kisi-kisi pada penelitian ini diawali dengan menganalisis kompetensi Dasar (KD). Analisis berdasarkan program Semester dan program tahunan yang telah ditetapkan oleh guru wali kelas IV. Kompetensi Dasar (KD) yang digunakan dalam kisi-kisi HOTS ini adalah seluruh Kompetensi Dasar yang terdapat dalam Tema 7 Indahnya Keberagaman di Negeriku Subtema 1 kelas IV semester 2. Kompetensi Dasar dirumuskan menjadi indikator yang dijadikan instrumen, memenuhi tingkat kognitif C4 (Menganalisis) dan C5 (mengevaluasi)

Berdasarkan hasil analisis Kompetensi Dasar (KD), yang telah dirumuskan menjadi indikator pada setiap

muatan materi seluruhnya memenuhi syarat dalam penyusunan soal HOTS. Selanjutnya disusunlah kisi-kisi tersebut yang memuat beberapa komponen yakni sebagai berikut Indikator Pencapaian Kompetensi (IPK), Dimensi Proses Berpikir, Dimensi Pengetahuan, Cakupan Materi serta nomor butir soal. Kisi-kisi yang telah disusun kemudian diserahkan kepada 5 ahli pakar untuk diuji pakar. Berdasarkan analisis seluruh judges kisi-kisi yang telah disusun sudah memenuhi dan sesuai kriteria karena para judges tidak memberikan masukan ataupun komentar terkait kisi-kisi yang telah disusun.

Kualitas instrumen tes HOTS Tema 7 Indahnya Keragaman di Negeriku Subtema 1 SD kelas IV

Kualitas instrumen tes uraian dapat ditinjau berdasarkan hasil validitas, reliabilitas, daya beda, indeks kesukaran serta *Partial Credit Model* (PCM). Pada instrumen tes HOTS Uraian Tema 7 Indahnya Keragaman di Negeriku Subtema 1 Kelas IV seluruh butir tes yang telah disusun oleh peneliti kemudian diserahkan pada para pakar judges. Hal tersebut bertujuan dalam mengetahui validitas isi dari instrumen. Uji pakar ini peneliti menggunakan 5 orang ahli yang terdiri dari 2 orang dosen ahli serta 3 orang guru kelas IV dari asal sekolah yang berbeda. Berdasarkan analisis uji pakar seluruh butir soal terbilang valid tanpa adanya komentar dari para pakar sehingga seluruh instrumen tidak memerlukan perbaikan. Hasil uji pakar dianalisis dengan menggunakan formula Gregory. Hasil dari validitas isi ini memperoleh nilai $CVR = 1$ yang berarti butir tes memiliki relevansi yang tinggi baik dari segi materi dan kriteria HOTS. Setelah dianalisis validitas isi, butir instrumen kemudian diuji *One To One* kepada 3 orang siswa untuk diperoleh nilai kepraktisan instrumen. Kepada 3 orang siswa yang dipilih berdasarkan tingkat kemampuan siswa, ada yang berkemampuan tinggi, sedang dan rendah. Ketiga siswa ini diberikan instrumen untuk menjawab instrumen tersebut setelah itu ketiga siswa diberikan sebuah kuesioner yang berkaitan dengan

kepraktisan instrumen. Hasil dari kuesioner kemudian dianalisis dengan rumus kepraktisan yang diperoleh hasil yakni 92% masuk dalam kategori sangat praktis. Setelah melaksanakan uji pakar *One To One* maka dilanjutkan dengan tahapan *Small Group*. Pada *small group* ada 5 analisis yang dilakukan yakni uji validitas, uji reliabilitas, uji daya beda dan uji tingkat kesukaran.

Uji validitas merupakan uji yang berfungsi untuk melihat apakah suatu alat ukur tersebut valid (sahih) atau tidak valid (Miftahul, 2021). Alat ukur yang dimaksud disini merupakan instrumen tes uraian HOTS Tema 7 Indahnya Keragaman di Negeriku Subtema 1 Kelas IV. Pada uji validitas semua soal berkategori valid berdasarkan nilai *Outfit MNSQ*, *ZSTD* serta *Pt Measure Corr* (Yulianto & Widodo, 2020). Analisis pengujian validitas instrumen pada penelitian ini menggunakan pemodel Rasch, hal yang dilihat yaitu tingkat validitas respon pada item yaitu, nilai *Outfit Mean Square* (MNSQ) yang diterima $0,5 < MNSQ < 1,5$, nilai *Outfit Z-Standard* (ZSTD) kesesuaian nilai uji z yang diterima $-2,0 < ZSTD < + 2,0$ dan *point Measure Correlation* (Pt Mean Corr) $0,4 < Pt Mean Corr < 0,85$. Seluruh butir soal dapat dinyatakan valid atau diterima, sesuai dengan 3 kriteria, maka tidak dilaksanakan perbaikan sehingga bisa dilanjutkan pada uji reliabilitas.

Uji reliabilitas adalah indeks yang menunjukkan sejauh mana suatu alat pengukur dapat dipercaya atau diandalkan. Sehingga uji reliabilitas dapat digunakan untuk mengetahui konsistensi alat ukur, apakah alat ukur tetap konsisten jika pengukuran tersebut diulang (Janna & Herianto, 2021). Alat ukur dikatakan reliabel jika menghasilkan hasil yang sama meskipun dilakukan pengukuran berkali-kali. Pada uji reliabilitas nilai reliabilitas instrumen yakni 0,832 yang berarti instrumen reliabel. Pada reliabilitas instrumen yang bersifat politomi ditentukan dengan *Alpha Cronbach*.

Setelah instrumen terbilang reliabel maka dilanjutkan dengan uji daya beda. Daya beda merupakan kemampuan pada butir tes dalam membedakan antara peserta didik yang berkemampuan tinggi

dan siswa yang memiliki kemampuan rendah (Arikunto, 2015). Analisis daya beda merupakan pengkajian butir tes yang dimaksudkan bertujuan mengetahui kesanggupan butir tes membedakan siswa yang tergolong mampu dengan siswa yang tergolong tidak mampu (Candiasa, 2011). Pada pengujian daya beda seluruh soal yang berkategori baik sekali ada 10 butir yakni butir soal nomor 1, 2, 3, 4, 5, 6, 7, 8, 9 dan 10.

Setelah uji daya beda kemudian dilakukan uji tingkat kesukaran. Tingkat kesukaran butir soal merupakan kesanggupan atau kemampuan peserta didik dalam menjawab, bukan asumsi dari guru dalam menyusun soal, hasil ini berdasarkan butir soal yang sulit, sedang atau mudah bagi guru tidak berarti sulit atau mudah bagi peserta didik (Loka, 2019). Pada uji tingkat kesukaran yang diperoleh hasil bahwa ada 1 soal yang berkategori mudah yakni butir soal nomor 10 sementara itu soal yang berkategori sedang ada 8 butir soal yakni butir soal nomor 1, 2, 3, 4, 5, 6, 7 dan 8. Soal yang berkategori sukar ada 1 soal yakni butir soal nomor 9. Berdasarkan hasil uji coba sehingga instrument uraian HOTS Tema 7 Indahya Keragaman di Negeriku Subtema 1 Kelas IV berjumlah 10 butir siap digunakan pada *Field Test*.

Field test adalah pengujian instrumen yang dilaksanakan pada kelompok besar. Pengujian ini ada 5 analisis yang dilakukan yakni analisis validitas, reliabilitas, daya beda, tingkat kesukaran serta *Partial Credit Model* (PCM). Pada uji validitas semua instrumen berkategori valid, jadi kesepuluh butir soal tersebut bisa dilanjutkan pada uji reliabilitas, pada uji reliabilitas nilai reliabilitas yang diperoleh yakni 0,974 yang menandakan bahwa instrumen tersebut tergolong reliabel. Setelah uji reliabilitas, instrumen tersebut kemudian diuji daya pembedanya yang mendapatkan hasil yakni ada 8 soal berkategori baik sekali yakni butir soal nomor 1, 2, 5, 6, 7, 8, 9, 10 dan ada 2 soal yang berkategori baik yakni butir soal nomor 3 dan 4. Pengujian dilanjutkan pada uji tingkat kesukaran instrumen yang mendapatkan hasil yakni soal yang berkategori mudah ada 2 soal yakni butir

soal nomor 3 dan 7 sementara itu soal yang berkategori sedang ada 6 soal yakni butir soal nomor 1, 2, 5, 6, 8, 9, 10 dan soal yang berkategori susah ada 1 yakni butir soal nomor 4.

Pada analisis *Partial Credit Model* (PCM) bebantuan aplikasi Winstep, didapat tes memiliki separation 3,02 dan item reliability 0,90. Hal ini menandakan bahwa tes telah berfungsi cukup baik karena, memiliki jangkauan tingkat kesulitan yang beragam. Sedangkan untuk subjek hanya memiliki separation 1,77 dan *person* reliability 0,76 hal tersebut, menandakan subjek kurang variatif karena hanya memiliki *range ability* yang sempit. Item butir soal yang paling sulit adalah item nomer 4 (P4) yang berada di posisi paling atas. Secara teoritis berdasarkan peta persebaran diatas soal P4 hanya subjek no 54, 43, 46, 55, 52, 60 dan 51 yang memiliki probabilitas menjawab benar karena posisinya ada diatas dari soal (P4). Sementara itu soal yang paling mudah adalah soal (P7) karena sebagian besar subjek memiliki probabilitas menjawab dengan benar kecuali bagi subjek no 06, 28, 09, 16, 49 dan 59 karena posisinya berada di bawah dari soal (P7).

Pada data politomi ini bisa digolongkan berfungsi dengan baik karena berdasarkan data observed count setiap category label terdapat beberapa subjek yang mendapatkan nilai itu (0,1,2,3) selain itu rating scale juga dikatakan baik jika dilihat dari category measure yakni terjadinya peningkatan yang bertahap dari -2,17 menjadi -0,52 lalu 0,63 hingga 2,00. Pada nilai Andrich Threshold diperoleh hasil sebagai hasil tersebut didapat data sebagai berikut tahap 1 memiliki tingkat kesulitan tahap sebesar - 0,89 (tinggi), tahap 2 memiliki tingkat kesulitan tahap sebesar 0,41 (tinggi), tahap 3 memiliki tingkat kesulitan tahap sebesar 0,48 (sedang). Rata-rata ability peserta didik dengan contoh penjabaran pada butir soal nomor 4 yakni peserta didik yang memiliki rata-rata ability -0,24 memiliki probabilitas mendapatkan skor 0, kemudian peserta didik yang memiliki rata-rata ability 0,21 memiliki probabilitas mendapatkan skor 1, lalu peserta didik dengan rata-rata ability 0,56 memiliki probabilitas mendapatkan

skor 2 sementara itu peserta didik yang memiliki rata-rata ability 0,58 memiliki probabilitas mendapatkan skor 3 dan begitu seterusnya untuk butir soal yang lain tentunya disesuaikan dengan hasil pada *output*.

Setelah melaksanakan analisis statistik yang menunjukkan bahwa kualitas tes telah valid, reliabel, memiliki daya beda dan tingkat kesukaran yang heterogen serta analisis statistika modern *Partial Credit Model* (PCM), secara substansi instrumen tes HOTS Uraian Tema 7 Indahya Keragaman di Negeriku Subtema 1 Kelas IV sudah mencerminkan tes HOTS dan kompetensi dasar pada masing-masing muatan materi yaitu, Pendidikan Pancasila dan Kewarganegaraan (PKN), Bahasa Indonesia, Ilmu Pengetahuan Alam (IPA), Ilmu Pengetahuan Sosial (IPS) dan Dasar Seni Budaya dan Prakarya (SDP), yang termuat dalam Tema 7 Indahya Keragaman di Negeriku Subtema 1 Kelas IV. Ditinjau berdasarkan jenis soal uraian HOTS bahwa tidak seluruh soal HOTS memiliki tingkat kesukaran yang sulit, soal HOTS memiliki keuntungan dalam meningkatkan motivasi belajar peserta didik karena penilaian HOTS menghubungkan materi pelajaran di kelas dengan konteks dunia nyata agar pembelajaran lebih bermakna (Hidayati, 2017)

Peran penting keterampilan berpikir tingkat tinggi dalam pembelajaran terletak pada proses pembelajaran (Hidayati, 2017). Peserta didik terbiasa berpikir kritis dan kreatif baik dalam pengambilan keputusan dan pemecahan masalah yang berkaitan dengan menganalisis, mengevaluasi dan mencipta (Anderson, L.W., and Krathwohl, 2001). Penelitian dalam proses pembelajaran telah menemukan bahwa pemahaman dan keterampilan paling baik dikembangkan ketika para peserta didik diizinkan untuk bergulat dengan ide-ide baru, membuat dan mempertahankan penyelesaian soal (Van de Walle & Bay Williams, 2009).

PENUTUP

Kisi-kisi instrumen berbasis HOTS Tema 7 Indahya Keberagaman di Negeriku Subtema 1 kelas IV semester 2 di Gugus VI Sukawati memuat muatan

pelajaran Bahasa Indonesia, Ilmu Pengetahuan Alam (IPA), Ilmu Pengetahuan Sosial (IPS), SDP yang masing-masing terdiri oleh 1 Kompetensi Dasar (KD). Kompetensi dasar diturunkan menjadi 10 indikator dengan 2 indikator yang termasuk C4/K1, 1 indikator yang termasuk C4/K2, 3 indikator yang termasuk C4/K4, 1 indikator yang termasuk C5/K2 dan 3 indikator yang termasuk C5/K4.

Kualitas tes Tema 7 Indahya Keberagaman di Negeriku Subtema 1 kelas IV semester 2 di Gugus VI Sukawati berbasis HOTS harus memenuhi beberapa uji. Uji validitas pakar menggunakan formula Gregory diperoleh nilai 1 dengan kriteria baik. Uji Validitas butir analisis model Rasch menggunakan winstep, semua butir telah valid. Uji Reliabilitas diperoleh 0,974 dengan kategori reliabel. Uji daya beda dari jumlah butir tes 10 diperoleh 2 butir soal berkategori baik dan 8 butir soal berkategori sangat baik. uji tingkat kesukaran dari butir tes berjumlah 10 soal diperoleh 2 butir soal mudah, 7 butir soal sedang dan 1 butir soal sukar. Hasil analisis menggunakan *Partial Credit Model* (PCM) terdapat tes memiliki separation 3,02 dan item reliability 0,90. Hal ini menandakan bahwa tes telah berfungsi cukup baik karena, memiliki jangkauan tingkat kesulitan yang beragam. Sedangkan untuk subjek hanya memiliki separation 1,77 dan *person reliability* 0,76 hal tersebut, menandakan subjek kurang variatif karena hanya memiliki *range ability* yang sempit. Item butir soal yang paling sulit adalah item nomer 4 (P4) yang berada di posisi paling atas. Secara teoritis berdasarkan peta persebaran diatas soal P4 hanya subjek no 54, 43, 46, 55, 52, 60 dan 51 yang memiliki probabilitas menjawab benar karena posisinya ada diatas dari soal (P4). Sementara itu soal yang paling mudah adalah soal (P7) karena sebagian besar subjek memiliki probabilitas menjawab dengan benar kecuali bagi subjek no 06, 28, 09, 16, 49 dan 59 karena posisinya berada di bawah dari soal (P7).

Diharapkan kedepannya dalam membuat tes guru lebih menekankan pada tes yang berbasis HOTS. Bagi guru-guru

yang ingin mengembangkan instrument HOTS diharapkan terlebih dahulu melakukan uji validasi butir soal sebelum digunakan dan bila perlu menggunakan tahapan pengembangan butir soal. Bagi peneliti lain apabila ingin mengembangkan instrumen tes berbasis HOTS, disarankan untuk memperbaiki keterbatasan pada penelitian ini dengan mengembangkan butir soal HOTS yang lebih bervariasi, relevan dengan lingkungan siswa serta cakupan materi yang lebih luas.

DAFTAR RUJUKAN

- Anderson, L.W., and Krathwohl, D. . (2001). *A Taxonomy of Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives*. Longman.
- Arikunto, S. (2015). *Dasar-Dasar Evaluasi Pendidikan*. Bumi Aksara.
- Bajpai, R., & Bajpai, S. (2014). Goodness of Measurement: Reliability and Validity. *International Journal of Medical Science and Public Health*, 3(2), 112. <https://doi.org/10.5455/ijmsph.2013.191120133>
- Budiarta, K., Harahap, M. H., Faisal, & Mailani, E. (2018). Potret implementasi pembelajaran berbasis high order thinking skills (HOTS) di Sekolah Dasar Kota Medan. *Jurnal Pembangunan Perkotaan*, 6(2), 102–111.
- Candiasa, I. M. (2011). *Pengujian Instrumen Penelitian Disertai Aplikasi ITEMAN dan BIGSTEPS*. Unit Penerbitan Universitas Pendidikan Ganesha.
- Candiasa, I. M. (2015). *Pengujian Instrumen Penelitian Disertai Aplikasi ITEMAN dan BIGSTEPS*. Universitas Pendidikan Ganesha.
- Chuseri, A., Anjarini, T., & Purwoko, R. Y. (2021). Pengembangan Modul Matematika Berbasis Realistik Terintegrasi Higher Order Thinking Skills (Hots) Pada Materi Bangun Ruang. *Alifmatika: Jurnal Pendidikan Dan Pembelajaran Matematika*, 3(1), 18–31. <https://doi.org/10.35316/alifmatika.2021.v3i1.18-31>
- Dasar, D. I. S. (2016). Pengembangan Soal Tes Berbasis Hots Pada Model Pembelajaran Latihan Penelitian Di Sekolah Dasar. *PEDADIDAKTIKA: Jurnal Ilmiah Pendidikan Guru Sekolah Dasar*, 3(1), 74–83.
- H., H. R. K. & S. (1985). *Item response theory*. Kluwer Nijhoff Publisher.
- Hidayati, A. U. (2017). TERAMPIL Jurnal Pendidikan dan Pembelajaran Dasar Volume 4 Nomor Melatih Keterampilan Berpikir Tingkat Tinggi Dalam Pembelajaran Matematika Pada Siswa Sekolah Dasar. *Pendidikan Dan Pembelajaran Dasar*, 4(20), 143–156.
- Janna, N. M., & Herianto. (2021). *Artikel Statistik yang Benar*. 18210047.
- K.O. Litna, N.M.S. Mertasari, & G. Sudirtha. (2021). Pengembangan Instrumen Tes Higher Order Thinking Skills (Hots) Matematika Sma Kelas X. *Jurnal Penelitian Dan Evaluasi Pendidikan Indonesia*, 11(1), 10–20. <https://doi.org/10.23887/jpepi.v11i1.278>
- Koyan, I. W. (2011). *Assesmen dalam Pendidikan*. Universitas Pendidikan Ganesha Press.
- Masters, G. N. (2016). *Partial Credit Model*. taylorfrancis.
- Muthoharoh, M. (2020). Inovasi pembelajaran pendidikan agama islam berbasis hots (higher order thinking skill). *Journal of Islamic Education*, 5(2), 131–143.
- 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. <https://doi.org/10.25273/pe.v10i1.6274>
- Pratiwi, N. P. W., Dewi, N. L. P. E. S., & Paramartha, A. A. G. Y. (2019). The

- Reflection of HOTS in EFL Teachers' Summative Assessment. *Journal of Education Research and Evaluation*, 3(3), 127. <https://doi.org/10.23887/jere.v3i3.21853>
- Retnawati, H. (2016). *Validitas, Reliabilitas, dan Karakteristik Butir*. Parama Publishing.
- Retnawati, H., Djidu, H., Kartianom, Apino, E., & Anazifa, R. D. (2018). Teachers' knowledge about higher-order thinking skills and its learning strategy. *Problems of Education in the 21st Century*, 76(2), 215–230. <https://doi.org/10.33225/pec/18.76.215>
- Riza, S. N. (2021). *Analysis of HOTS Indicators on Physics Problems for Senior High Schools in West Pasaman District*. 14(2), 94. <http://repository.unp.ac.id/34167/>
- Sani, A. (2019). *Pembelajaran HOTS*. [https://books.google.co.id/books?hl=en&lr=&id=GfrDwAAQBAJ&oi=fnd&pg=PR5&dq=peserta+didik+berada+pada+level+tersebut+\(Sani,+2019\).+&ots=kHzXJX4zA-&sig=rPRVWbzXkRD-YaXSf1OLIOzF5uY&redir_esc=y#v=onepage&q=peserta+didik+berada+pada+level+tersebut+\(Sani%2C+2019\)](https://books.google.co.id/books?hl=en&lr=&id=GfrDwAAQBAJ&oi=fnd&pg=PR5&dq=peserta+didik+berada+pada+level+tersebut+(Sani,+2019).+&ots=kHzXJX4zA-&sig=rPRVWbzXkRD-YaXSf1OLIOzF5uY&redir_esc=y#v=onepage&q=peserta+didik+berada+pada+level+tersebut+(Sani%2C+2019))
- Sudijono, A. (2011). *Pengantar Statistik Pendidikan*. PT. RajaGrafindo Persada.
- Sugiyono. (2007). *Metode Penelitian Kuantitatif Kualitatif dan R&D*. Alfabeta.
- Sugiyono. (2009). *Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif, dan R & D*. Alfabeta.
- Suhady, W., Roza, Y., & Maimunah, M. (2020). Pengembangan Soal untuk Mengukur Higher Order Thinking Skill (HOTS) Siswa. *Jurnal Gantang*, 5(2), 143–150. <https://doi.org/10.31629/jg.v5i2.2518>
- Suharsimi. (2018). *Dasar-Dasar Evaluasi Pendidikan*. Bumi Aksara.
- Syofyan, H. (2016). *Penyuluhan dan Pelatihan Pendidikan Tentang Pembuatan Kisi-Kisi Soal untuk Guru-Guru d Yayasan Perguruan Birrul Waalidain Semplak Bogor*. *Jurnal Abdimas*, 3(1), 12–17. <https://journal.umtas.ac.id/index.php/ABDIMAS>
- Tanujaya, B., Mumu, J., & Margono, G. (2017). The Relationship between Higher Order Thinking Skills and Academic Performance of Student in Mathematics Instruction. *International Education Studies*, 10(11), 78. <https://doi.org/10.5539/ies.v10n11p78>
- Van de Walle & Bay Williams, J. M. (2009). *Elementary and middle school mathematics: Teaching developmentally (7th ed.)*. Allyn & Bacon/Merill.
- Yulianto, A., & Widodo, A. (2020). Indonesian Journal of Primary Education Disclosure of Difficulty Distribution of HOTS-Based Test Questions through Rasch Modeling. 4(2), 197–203.
- Yusuf, M. (2015). *Asesmen dan Evaluasi Pendidikan*. Prenadamedia Group.