

Pengembangan Instrumen Penilaian Kemampuan Berpikir Tingkat Tinggi dan Literasi Sains Pada Pembelajaran IPA Kelas IV SD

Sarifa Faridatil Imi Al Idrus

STIT Sibawaihi Mutawalli, Indonesia sarifa.faridatil@gmail.com

Mohammad Junaidin Muhdar

STIT Sibawaihi Mutawalli, Indonesia junaidinmuhdar@gmail.com

Abstract: *This study aims to develop an instrument for higher-order thinking skills and science literacy by examining its validity, reliability, discrimination index, difficulty level, and effectiveness. The study involved fourth-grade students at SD IT Abu Bakar Ash-Shiddiq Dane Rase. The method used in this study was development. The developed instrument was in the form of multiple-choice questions to measure higher-order thinking skills and essays to measure science literacy. Data were analyzed by determining validity, reliability, discrimination index, and difficulty level. Based on the results of the judges' test and empirical trials, the content validity was high, the item validity of the higher-order thinking skills instrument had 20 valid items and 5 invalid items. Meanwhile, all items in the science literacy instrument were valid, with reliability coefficients of 0.937 and 0.940, categorized as very high, item discrimination ranged from fairly good to very good, and the difficulty level had only 1 item in easy category, other items are medium and hard. High-level thinking skill instruments are effective while science literacy is less effective.*

Keywords: *Higher Order Thinking Skills, Scientific Literacy, Instrument*

Abstrak: *Penelitian ini bertujuan mengembangkan instrumen kemampuan berpikir tingkat tinggi dan literasi sains dengan melihat validitas, reliabilitas, daya beda dan tingkat kesukaran butir serta keefektifannya. Penelitian ini melibatkan peserta didik kelas IV SD IT Abu Bakar Ash-Shiddiq Dane Rase. Metode yang digunakan dalam penelitian ini adalah pengembangan. Instrumen yang dikembangkan berbentuk pilihan ganda untuk mengukur kemampuan berpikir tingkat tinggi dan uraian untuk mengukur literasi sains. Data dianalisis dengan menentukan validitas, reliabilitas, daya beda dan tingkat kesukaran. Berdasarkan hasil uji *judgest* dan uji coba empiris diperoleh validitas isi bekategori tinggi, validitas butir instrumen kemampuan berpikir tingkat tinggi 20 butir valid dan 5 tidak valid. Sedangkan instrumen literasi sains semua valid, koefisien reliabilitas 0,937 dan 0,940 dengan kategori sangat tinggi, daya beda butir berkategori cukup baik sampai sangat baik, tingkat kesukaran hanya 1 butir pada kategori mudah, butir lain kategorinya sedang dan sukar. Instrumen kemampuan berpikir tingkat tinggi efektif dan literasi sains kurang efektif.*

Kata kunci: *Kemampuan Berpikir Tinggi, Literasi Sains.*

PENDAHULUAN

Salah satu komponen esensial dalam penyelenggaraan pendidikan adalah asesmen atau penilaian. Peningkatan kualitas pendidikan tidak dapat dilepaskan dari peningkatan kualitas pembelajaran dan sistem asesmen yang digunakan. Dalam implementasi Kurikulum Merdeka, asesmen tidak hanya berfungsi untuk mengukur hasil belajar peserta didik, tetapi juga menjadi bagian integral dari proses pembelajaran yang bertujuan memberikan umpan balik bagi peserta didik dan guru untuk meningkatkan kualitas pembelajaran secara berkelanjutan. Pengembangan asesmen pada Kurikulum Merdeka mengacu pada Standar Nasional Pendidikan, khususnya Standar Isi, Standar Proses, dan Standar Penilaian Pendidikan. Kurikulum Merdeka dirancang untuk mengembangkan kompetensi peserta didik secara utuh melalui pembelajaran yang berpusat pada peserta didik, berdiferensiasi, serta berorientasi pada pencapaian kompetensi dan

karakter sesuai Profil Pelajar Pancasila. Materi pembelajaran disusun secara lebih esensial dan fleksibel agar peserta didik memiliki kesempatan yang lebih luas untuk mengembangkan kemampuan berpikir kritis, kreatif, kolaboratif, dan komunikatif sesuai tuntutan abad ke-21.

Sejalan dengan hal tersebut, asesmen dalam Kurikulum Merdeka menekankan penggunaan asesmen formatif dan sumatif yang berkeadilan, objektif, dan edukatif. Asesmen tidak lagi berfokus pada penguasaan materi semata, tetapi juga pada kemampuan peserta didik dalam menganalisis, mengevaluasi, memecahkan masalah, serta menerapkan pengetahuan dalam berbagai konteks kehidupan nyata. Oleh karena itu, pengembangan instrumen asesmen perlu diarahkan pada pengukuran kemampuan berpikir tingkat tinggi (Higher Order Thinking Skills/HOTS) dan literasi sebagai kompetensi penting yang mendukung kesiapan peserta didik menghadapi tantangan global.

Kemampuan berpikir tingkat tinggi (High Order Thinking Skills/HOTS) merupakan suatu keterampilan yang amat dibutuhkan peserta didik demi mempertahankan eksistensi di era globalisasi dan teknologi informasi yang kian pesat perkembangannya (Jamaluddin, et al 2020: 13). Kemampuan berpikir tingkat tinggi dapat dilatih melalui pembelajaran IPA. Pendidikan IPA berperan sentral dalam mengembangkan pengetahuan, sikap, dan keterampilan berpikir peserta didik. Melalui pendidikan IPA peserta didik dapat mengenal, menyikapi dan mengapresiasi ilmu pengetahuan dan teknologi, serta menanamkan kebiasaan berpikir dan berperilaku ilmiah yang kritis, kreatif dan mandiri (BNSP, 2006). Namun, dalam dalam penerapannya pembelajaran yang berorientasi membentuk keterampilan berpikir tingkat tinggi bukan hal yang mudah bagi guru. Materi dan strategi pembelajaran harus benar-benar dikuasai oleh guru. Ndiung & Jediut (2020:94) menyatakan bahwa pembelajaran akan bermakna jika peserta didik diarahkan untuk berpikir tingkat tinggi.

Hal ini berarti peserta didik dianggap berhasil menguasai suatu konsep apabila telah mampu berpikir tingkat tinggi. Agar kemampuan berpikir tingkat tinggi peserta didik dapat berkembang, khususnya dalam pembelajaran IPA, peserta didik hendaknya dibiasakan untuk melakukan aktivitas-aktivitas yang dapat mengasah kemampuan berpikir tingkat tinggi itu sendiri. Dalam hal ini kemampuan peserta didik tidak hanya terbatas pada mengingat dan memahami suatu konsep, namun peserta didik mampu untuk menganalisis, mensintesis, mengevaluasi dan mengkreasikan suatu konsep dengan baik, konsep yang telah dipahami dapat diingat dalam waktu yang lama, sehingga sangat penting bagi peserta didik untuk memiliki keterampilan berpikir tingkat tinggi (Arifin & Retnawati, 2017).

Hal ini berarti peserta didik dianggap berhasil menguasai suatu konsep apabila telah mampu berpikir tingkat tinggi. Agar kemampuan berpikir tingkat tinggi peserta didik dapat berkembang, khususnya dalam pembelajaran IPA, peserta didik hendaknya dibiasakan untuk melakukan aktivitas-aktivitas yang dapat mengasah kemampuan berpikir tingkat tinggi itu sendiri. Dalam hal ini kemampuan peserta didik tidak hanya terbatas pada mengingat dan memahami suatu konsep, namun peserta didik mampu untuk menganalisis, mensintesis, mengevaluasi dan mengkreasikan suatu konsep dengan baik, konsep yang telah

dipahami dapat diingat dalam waktu yang lama, sehingga sangat penting bagi peserta didik untuk memiliki keterampilan berpikir tingkat tinggi (Arifin & Retnawati, 2017).

Proses pembelajaran yang selama ini dilakukan belum dapat mengembangkan kemampuan berpikir tingkat tinggi peserta didik. Proses pembelajaran yang menuntut peserta didik untuk mampu mengingat dan memperbanyak informasi tanpa pemahaman yang baik masih marak terjadi. Antara, et al (2021:249) menyatakan bahwa proses pembelajaran yang dilakukan saat ini memaksa peserta didik untuk mengingat dan menimbun informasi tanpa memahami informasi tersebut. Akibatnya setelah peserta didik lulus, mereka hanya pandai dalam teori tapi kurang dalam hal pengaplikasian. Karena dalam proses pembelajaran peserta didik kurang diarahkan untuk menggunakan kemampuan berpikirnya dalam memecahkan masalah. Hal ini sejalan dengan pendapat Hidayati, et al (2017) yang menyatakan bahwa guru masih kesulitan dalam melatih peserta didik berpikir tingkat tinggi.

Sampai saat ini kemampuan berpikir tingkat tinggi peserta didik di Indonesia khususnya dalam pembelajaran IPA masih rendah. Hal ini dibuktikan dengan 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 dan cenderung tidak mengalami peningkatan 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). Hasil studi menunjukkan bahwa kualitas pembelajaran IPA masih perlu ditingkatkan agar kemampuan berpikir tinggi peserta didik dalam pemecahan masalah IPA semakin baik.

Selain kemampuan berpikir tingkat tinggi, literasi sains merupakan kemampuan yang harus dimiliki peserta didik dalam pembelajaran IPA. Pemahaman tentang IPA yang meliputi pemahaman terhadap alam melalui penguasaan ilmu dasar IPA seperti kimia, biologi, fisika serta pemahaman tentang hakikat IPA sebagai suatu penyelidikan ilmiah menjadi fokus utama dalam kajian literasi sains (Mardhiyyah, et al 2016:148). Literasi sains berkaitan dengan kemampuan peserta didik dalam memahami berbagai informasi, ilmu pengetahuan serta fakta yang ada dalam kehidupan sehari-hari. (Suastra, 2017) menyatakan bahwa pada dasarnya literasi sains bukan hanya sekedar mampu membaca, menulis, dan mengkomunikasikan melainkan kesadaran serta kemampuan peserta didik dalam menyelesaikan masalah, menggunakan konsep-konsep sains.

Literasi sains tidak hanya menitikberatkan pada pengetahuan sains saja, tetapi juga keterampilan sains. Berikut aspek keterampilan literasi sains meliputi keterampilan proses sains, pengambilan keputusan dalam isu-isu sosial ditinjau dari segi sains (socio scientific issue), serta pemecahan masalahnya (Holbrook & Rannikmae, 2007; Sadler & Zeidler, 2009). Aspek keterampilan sains tersebut hendaknya dilatih kepada peserta didik dalam pembelajaran IPA. Berdasarkan data PISA 2012, Indonesia berada pada peringkat ke-64 dari 65 negara peserta

dengan perolehan rata-rata nilai komponen literasi sains peserta didik Indonesia sebesar 382. Angka tersebut masih di bawah skor rata-rata PISA, yaitu sebesar 501 (OECD, 2014). Hal ini menunjukkan bahwa kompetensi literasi sains peserta didik masih perlu ditingkatkan.

Instrumen yang selama ini digunakan guru belum dapat mengembangkan kemampuan berpikir tingkat tinggi dan literasi sains peserta didik. Ndiung & Jediut (2020:95) menyatakan bahwa sampai saat ini sebagian besar instrumen hasil belajar yang digunakan oleh sekolah untuk ulangan umum belum mengarah kepada penyelesaian masalah yang menantang penalaran peserta didik terkait penyelesaian soal-soal yang tidak biasanya yang menuntut penyelesaian dengan prosedur non rutin. Handayani., et al (2021: 16) melaporkan hasil analisis terhadap soal yang dibuat oleh guru di sekolah dasar yang menunjukkan bahwa pemenuhan soal yang menuntun pada kemampuan berpikir tingkat tinggi di setiap sekolah kurang dari 8%. Idealnya tes formatif yang dilaksanakan oleh guru 25% mencakup keterampilan berpikir tingkat tinggi (Standar Penilaian BAN, 2012). Sebagian besar soal yang ditemukan hanya mencakup pada kemampuan mengingat, memahami, dan penerapan (C1-C3). Guru kurang memperhatikan tuntutan dari C1-C6 dalam pembuatan soal. Hal ini mengakibatkan kemampuan berpikir tingkat tinggi peserta didik menjadi rendah sehingga ketika dihadapkan dengan soal-soal yang kompleks mereka kesulitan dalam menyelesaikannya. Selanjutnya kompetensi literasi sains peserta didik menjadi rendah karena mereka tidak dilatih untuk mengemukakan pendapat atau gagasan-gagasan yang ada dalam pikiran mereka, sehingga ketika dihadapkan dengan soal-soal yang menuntut pengungkapan makna dan keterkaitan materi dengan lingkungan sekitar peserta tidak dapat menyelesaikannya (Mardhiyyah, 2015:149).

Berbagai persoalan yang telah diuraikan diatas perlu penyelesaian untuk meningkatkan kemampuan berpikir tingkat tinggi dan literasi sains peserta didik. Salah satu upaya yang dapat dilakukan guru untuk memastikan capaian kemampuan berpikir tingkat tinggi dan literasi sains peserta didik pada pembelajaran IPA adalah menggunakan instrumen yang relevan. Ndiung & Jediut (2020:96) menyatakan bahwa melalui pembiasaan untuk menyelesaikan soal-soal berpikir tingkat tinggi maka dapat membentuk peserta didik yang aktif, positif dalam menyelesaikan masalah. Tes merupakan suatu media program penilaian pendidikan yang berbentuk butir-butir soal untuk mengukur kemampuan peserta didik yang jawabannya dapat dinyatakan benar atau salah (Mudjijo, 1995). Suatu instrumen tes dapat dikatakan baik jika memiliki kriteria sebagai berikut: (1) memiliki validitas yang cukup tinggi; (2) memiliki reliabilitas yang baik; dan (3) memiliki nilai kepraktisan (Thoha, 1996). Oleh karena itu, guru seyogyanya dapat menyusun instrumen yang relevan. Dengan demikian, kemampuan berpikir tingkat tinggi dan literasi sains peserta didik dapat diukur secara valid.

Dalam meningkatkan kualitas pendidikan dibutuhkan perbaikan dalam proses pembelajaran di sekolah melalui metode kerja yang sistematis salah satunya adalah inovasi dari sistem penilaian. Hal ini berarti 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 (Pandra &

Mardapi, 2017). Evaluasi dapat menjadi motivasi bagi peserta didik 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 peserta didik.

Berkaitan dengan hal tersebut, optimalisasi sistem eIValuasi memiliki dua makna, pertama adalah sistem evaluasi yang memberikan informasi yang optimal. Kedua adalah manfaat yang dicapai dari eIValuasi. Manfaat yang utama dari eIValuasi adalah meningkatkan kualitas pembelajaran yang berimplikasi terhadap 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 kemampuan berpikir tingkat tinggi dan yang mampu meningkatkan kompetensi literasi sains peserta didik. Berbagai pertimbangan tersebut yang mendorong peneliti untuk memilih topik penelitian “Pengembangan Instrumen Kemampuan Berpikir Tingkat Tinggi dan Literasi Sains Pada Pembelajaran IPA kelas IV SD”.

Tujuan penelitian ini yaitu: (1) untuk mengembangkan instrumen penilaian kemampuan berpikir tingkat tinggi dan literasi sains pada pembelajaran IPA kelas IV SD, (2) untuk menganalisis validitas instrumen penilaian kemampuan berpikir tingkat tinggi dan literasi sains pada pembelajaran IPA kelas IV SD, (3) untuk menganalisis reliabilitas instrumen penilaian kemampuan berpikir tingkat tinggi dan literasi sains pada pembelajaran IPA kelas IV SD, (4) untuk menganalisis daya beda instrumen penilaian kemampuan berpikir tingkat tinggi dan literasi sains pada pembelajaran IPA kelas IV SD, (5) untuk menganalisis tingkat kesukaran instrumen penilaian kemampuan berpikir tingkat tinggi dan literasi sains pada pembelajaran IPA kelas IV SD, (6) untuk menganalisis keefektifan instrumen penilaian kemampuan berpikir tingkat tinggi dan literasi sains pada pembelajaran IPA kelas IV SD.

METODE

Penelitian ini menggunakan model penelitian dan pengembangan pendidikan (educational research and development) yang bertujuan untuk mengembangkan instrumen kemampuan berpikir tingkat tinggi dan literasi sains pada pembelajaran IPA kelas IV SD. Sebagaimana dijelaskan oleh Borg and Gall (1989) kegiatan research and development adalah suatu proses penelitian yang tujuannya untuk mengembangkan dan memvalidasi produk-produk pendidikan. Prosedur pengembangan yang digunakan dalam penelitian ini, diadaptasi dari Thiagarajan dengan 4 four D Model (model 4D) yang terdiri atas empat tahap meliputi (1) tahap define; (2) tahap design; (3) tahap develop; dan (4) tahap disseminate.

Pada penelitian ini, data dikumpulkan menggunakan lembar validasi pakar, tes pilihan ganda kemampuan berpikir tingkat tinggi, dan tes uraian literasi sains. Berikut teknik pengumpulan data dapat dilihat pada tabel 1 dibawah ini.

Tabel 1 Teknik dan Instrumen Pengumpulan Data

No	Data	Teknik	Jenis instrumen	Validasi instrument
1	Data kemampuan berpikir tingkat tinggi	Tes	Pilihan ganda	Validitas isi Reliabilitas tes Keefektifan tes
2	Data literasi sains	Tes	Uraian	Validitas isi Reliabilitas tes Keefektifan tes

Metode analisis data pada penelitian ini dilakukan dengan beberapa cara yaitu: Pertama analisis validitas isi instrumen dilakukan dengan menguji validitas isi instrumen kemampuan berpikir tingkat tinggi dan literasi sains yang telah dikembangkan pada 5 pakar yaitu 2 dosen ahli dan 3 praktisi lapangan (guru) kelas V sekolah dasar menggunakan lembar validasi pakar. Hasil validasi pakar dianalisis menggunakan analisis Content Validity Ratio (CVR) dengan rumus berikut:

$$CVR = \frac{2ne}{N} - 1$$

(Azwar, 2013:114)

Keterangan :

CVR : *Content Validity Ratio* (Rasio Validitas Isi / RVI).

ne : Jumlah panelis/validator yang menyetujui kevalidan instrument

N : Jumlah panelis / validator seluruhnya.

Setelah mengidentifikasi setiap sub pertanyaan instrumen dengan menggunakan CVR, maka selanjutnya yaitu menghitung CVI (Content Validity Index). CVI digunakan untuk menghitung rata-rata dari keseluruhan nilai CVR untuk komponen instrumen kemampuan berpikir tingkat tinggi dan literasi sains. Adapun rumus CVI, yaitu sebagai berikut:

$$CVI = \frac{\sum CVR}{k}$$

(Azwar, 2013:115)

Keterangan:

CVI : *Content Validity Index*

CVR : *Content Validity Ratio* (Rasio Validitas Isi/RVI)

k : banyaknya item

Setelah mendapatkan hasil perhitungan validitas isi, untuk mengetahui di kategori mana koefisien validitas itu berada, maka diketahui berdasarkan tabel 2 berikut ini.

Tabel 2 Kriteria Koefisien Validitas Isi

Koefisien	Validitas
0,80 – 1,00	Validitas isi sangat tinggi
0,60 – 0,79	Validitas isi tinggi
0,40 – 0,59	Validitas isi sedang
0,20 – 0,39	Validitas isi rendah
0,00 – 0,19	Validitas isi sangat rendah

Instrumen yang dinyatakan valid dan memiliki koefisien validitas isi dari sedang sampai tinggi oleh pakar, kemudian siap untuk diuji coba kepada peserta didik secara terbatas sebelum di sebarluaskan. Selanjutnya dilakukan uji validitas empiris menggunakan analisis butir, yaitu dengan mengkorelasikan skor tiap butir dengan skor total yang merupakan jumlah skor tiap butir soal. Validitas butir soal ditentukan dengan rumus korelasi product moment sebagai berikut.

$$r_{xy} = \frac{N \sum XY - \sum X \sum Y}{\sqrt{(N \sum X^2 - (\sum X)^2)(N \sum Y^2 - (\sum Y)^2)}} \\ \text{(Koyan, 2012:126)}$$

Keterangan:

r_{xy} = korelasi antara skor butir dengan skor total
 X = skor butir
 Y = skor total
 N = banyaknya responden

Kriteria yang digunakan adalah dengan membandingkan harga r_{xy} dengan harga tabel kritik r product moment. Jika r_{xy} lebih besar dari pada r_{xy} tabel (pada taraf signifikansi 5 %), maka butir tes dinyatakan valid begitu pula sebaliknya. Dalam penelitian ini validitas butir dihitung menggunakan program microsoft excel 2010.

Kedua analisis reliabilitas, nilai-nilai untuk uji reliabilitas diambil dari skor-skor item yang dinyatakan valid. Uji reliabilitas instrumen diuji dengan Analisis reliabilitas menggunakan rumus KR-20 untuk tes kemampuan berpikir tingkat tinggi dan alpha cronbach untuk tes literasi sains.

Rumus KR-20:

$$r_{1.1} = \left(\frac{k}{k-1} \right) \left(\frac{SD_t^2 - \sum pq}{SD_t^2} \right) \\ \text{(Koyan, 2011:133)}$$

Keterangan:

$r_{1.1}$ = koefisien reliabilitas tes
 p = proporsi responden yang menjawab benar
 q = proporsi responden yang menjawab salah
 SD_t^2 = varian total tes
 k = jumlah butir tes

Rumus *alpha cronbach*:

$$r_{1.1} = \left(\frac{k}{k-1} \right) \left(\frac{SD_t^2 - \sum SD_i^2}{SD_{tot}^2} \right) \\ \text{(Koyan, 2011:135)}$$

Keterangan:

$r_{1.1}$ = koefisien reliabilitas
 k = banyak butir dalam instrumen
 SD_i^2 = varians skor tiap butir
 SD_{tot}^2 = varian skor total

Selanjutnya hasil perhitungan atau pengujian reliabilitas yang telah diperoleh dibandingkan dengan kriteria reliabilitas. Berdasarkan cara perhitungan reliabilitas yang sama. Untuk kriteria derajat reliabilitas tes atau instrumen evaluasi digunakan kriteria yang dikemukakan oleh Guilford (dalam Koyan, 2011: 136) dapat di lihat pada table 3 sebagai berikut.

Tabel 3 Kriteria Koefesien Reliabilitas Instrumen	
Koefesien	Reliabilitas
$0,80 \leq 1,00$	Reliabilitas sangat tinggi
$0,60 \leq 0,79$	Reliabilitas tinggi
$0,40 \leq 0,59$	Reliabilitas sedang
$0,20 \leq 0,39$	Reliabilitas rendah
$0,00 \leq 0,19$	Reliabilitas sangat rendah

Ketiga analisis daya beda dilakukan dengan menentukan kelompok atas dan kelompok bawah dengan yaitu 50% kelompok atas dan 50% kelompok bawah karena jumlah responden kurang dari 30 orang. Untuk mengetahui daya beda suatu tes dalam bentuk pilihan ganda dengan skor 0-1 digunakan indeks Johnson. Bentuk rumus Indeks Johnson yang digunakan adalah sebagai berikut:

$$IJ = \frac{RU - RL}{f}$$

(Dantes, 2012:110)

Keterangan:

- IJ = Indeks Johnson (berdasarkan pada *Upper Lower Index*)
 RU = Jumlah subjek golongan atas yang menjawab benar terhadap butir tes.
 RL = Jumlah subjek golongan bawah yang menjawab benar terhadap butir tes.
 f = Jumlah subjek masing-masing golongan.

Sedangkan menghitung nilai daya beda untuk soal uraian digunakan rumus berikut.

$$DB = \frac{\text{Rerata kelompok atas} - \text{rerata kelompok bawah}}{\text{skor maksimum soal}}$$

(Sudijono, 2006)

Kriteria Daya Beda (D) menurut Koyan (2011:141):

Skor	Kriteria
0,00-0,19	Kurang Baik
0,20-0,39	Cukup baik
0,40-0,70	Baik
0,71-1,00	Sangat baik

Jika “D” negatif, soal tersebut sangat buruk dan harus dibuang. Tes yang baik apabila memiliki D antara 0,15-0,20 atau lebih (Koyan, 2011:141).

Keempat, analisis tingkat kesukaran dilakukan dengan indeks kesukaran rata-rata untuk instrumen kemampuan berpikir tingkat tinggi. Rumus indeks kesukaran rata-rata adalah sebagai berikut:

$$D = \frac{SB}{N}$$

(Dantes, 2012:104).

Keterangan:

P IKR = Proporsi jawaban benar (Indeks kesukaran rata-rata)

SB = Subjek yang menjawab benar pada butir tes yang diuji

N = Jumlah semua subjek yang menjawab butir tes. Sedangkan rumus untuk menghitung tingkat kesukaran tes uraian adalah sebagai berikut.

$$\frac{\text{mean skor maksimum yang ditetapkan}}{(\text{Daryanto \& Mulyo, 2012})}$$

Keterangan:

Mean = rata-rata

Skor maksimum yang ditetapkan = 4

Kriteria tingkat kesukaran (Koyan, 2011:140)

Skor	Kriteria
0,00-0,29	Sukar
0,30-0,70	Sedang
0,71-1,00	Mudah

Kelima analisis keefektifan dilakukan pada peserta didik kelas IV SD gugus 4 Sukasada. keefektifan instrumen ditinjau dari peningkatan kemampuan berpikir tingkat tinggi dan literasi sains peserta didik berdasarkan hasil pretest dan posttest yang dihitung dengan rumus indeks gain berikut ini.

$$N - gain = \frac{\text{Skor Posttest} - \text{Skor Pretest}}{\text{Skor tuntas} - \text{Skor Pretest}}$$

Pembagian kategori perolehan *N- gain* dalam bentuk persen (%) dapat dilihat pada tabel 4 dibawah ini.

Tabel 4 Kategori Tafsiran Keefektifan *N-Gain*

Persentase (%)	Tafsiran
>76	Efektif
56-75	Cukup Efektif
40-55	Kurang Efektif
<40	Tidak Efektif

Modifikasi dari (Hake, R.R, 1999)

PEMBAHASAN

Hasil dari penelitian pengembangan ini adalah telah dikembangkan instrumen penilaian kemampuan berpikir tingkat tinggi dan literasi sains pada pembelajaran IPA kelas IV SD. Prosedur pengembangan terdiri dari empat tahap meliputi (1) define, (2) design, (3) develop, dan (4) disseminate.

Pada tahap define ada tiga kegiatan yang dilakukan yaitu yaitu analisis awal dengan melakukan observasi langsung ke SD Gugus 4 Sukasada dan di temukan bahwa belum ada instrumen yang secara baku mengukur kemampuan berpikir tingkat tinggi dan literasi sains peserta didik setelah mencermati instrumen yang ada di sekolah, kedua melakukan analisis peserta didik dengan mencermati data-data terkait dan ditemukan bahwa tujuan pengembangan instrumen kemampuan berpikir tingkat tinggi dan literasi sains sesuai dengan karakteristik peserta didik, ketiga melakukan analisis konsep dengan menentukan grand theory dari masing-masing variabel untuk menghubungkan dengan tema,

kompetensi inti, kompetensi dasar, dan indikator sesuai instrumen yang dikembangkan.

Pada tahap design ada beberapa kegiatan yang dilakukan yaitu menentukan pola instrumen dengan tes pilihan ganda untuk mengukur kemampuan berpikir tingkat tinggi dan tes uraian untuk mengukur literasi sains, selanjutnya menyusun kisi-kisi instrumen yang dihasilkan dari grand theory kemampuan berpikir tingkat tinggi dan literasi sains, dan mengimplementasikan kisi-kisi tersebut menjadi butir-butir instrumen 25 butir untuk kemampuan berpikir tingkat tinggi dan 10 butir untuk literasi sains.

Pada tahap develop ada dua kegiatan yang dilakukan yaitu penilaian ahli (expert appraisal) dengan menguji validitas instrumen yang telah di kembangkan kepada pakar yaitu 2 dosen ahli pendidikan dasar dan 3 praktisi (guru kelas IV SD Gugus 4 Sukasada). Diperoleh hasil perhitungan validitas CVR dari instrumen kemampuan berpikir tingkat tinggi semua butir valid dengan nilai CVI sebesar 0,79 yang termasuk dalam kategori tinggi dan CVR instrumen literasi sains semua butir valid dengan nilai CVI 0,72 yang termasuk dalam kategori tinggi.

Pada tahap disseminate ada beberapa kegiatan yang dilakukan yaitu validation testing, packaging, diffusion and adoption dengan menguji coba instrumen yang telah divalidasi oleh pakar pada peserta didik kelas IV SD 4 Sukasada dari hasil uji coba instrumen kemampuan berpikir tingkat tinggi (1) 20 butir dinyatakan valid dan 5 butir tidak valid, (2) dinyatakan reliabel dengan koefisien reliabilitas sebesar 0,937 yang termasuk dalam kategori sangat tinggi, (3) daya beda dengan kategori sangat baik sebanyak 6 butir, baik 12 butir, cukup baik 2 butir, dan (4) tingkat kesukaran dengan kategori sukar sebanyak 1 butir, sedang sebanyak 18 butir, dan mudah 1 butir. Sedangkan instrumen literasi sains (1) seluruhnya dinyatakan valid, (2) dinyatakan reliabel dengan koefisien reliabilitas sebesar 0,940 yang termasuk dalam kategori sangat tinggi, (3) daya beda dengan kategori baik sebanyak 3 butir dan cukup baik sebanyak 7 butir, dan (4) tingkat kesukaran dengan kategori sedang seluruhnya. Selanjutnya instrumen yang dinyatakan valid dan reliabel berdasarkan hasil uji coba disebar kepada seluruh peserta didik kelas IV SD Gugus 4 Sukasada dan satu sekolah diluar gugus yaitu SDN 1 Banjar Tegal untuk menguji keefektifan instrumen dan diperoleh hasil instrumen penilaian kemampuan berpikir tingkat tinggi cukup efektif dengan skor N-gain sebesar 59,9%. Sedangkan instrumen penilaian literasi sains kurang efektif dengan skor N-gain sebesar 45,3%.

Berdasarkan hasil tersebut instrumen penilaian kemampuan berpikir tingkat tinggi dan literasi sains dinyatakan valid dan reliabel dengan kategori sangat tinggi. Hal ini menunjukkan bahwa instrumen yang dikembangkan telah memenuhi syarat instrumen yang baik. Iskandar dan Rizal (2017:13) menyatakan bahwa instrumen yang baik setidaknya memenuhi syarat valid dan reliabel. Hasil penelitian aspek validitas dan reliabilitas sejalan dengan hasil penelitian Antara., et al (2021) yang menyatakan bahwa instrumen penilaian berbasis higher order thinking skills (HOTS) yang dikembangkan seluruhnya valid berdasarkan validitas isi dan memiliki nilai $r_{hitung} > r_{tabel}$. Koefisien reliabilitas sebesar 0,91 yang berarti seluruh butir instrumen valid, reliabel, dan dapat digunakan. Selain itu, Handayani., et al juga mengemukakan hasil penelitian yang menunjukkan

bahwa instrumen literasi sains yang dikembangkan 30 butir valid dengan nilai CVI sebesar 0,94 dan koefisien reliabilitas sebesar 0,77 yang berarti instrumen valid, reliabel, dan dapat digunakan.

Selanjutnya berdasarkan aspek daya beda dan tingkat kesukaran instrumen penilaian yang dikembangkan telah memenuhi syarat instrumen yang baik. Arifin (2017:32) menyatakan bahwa indeks daya beda berkisar antara -1,00 sampai +1,00 semakin tinggi daya beda suatu instrumen, maka semakin baik kualitas instrumen tersebut. Lebih lanjut Arikunto (2013) menyatakan bahwa instrumen yang baik adalah instrumen yang memiliki indeks tingkat kesukaran dengan kategori sedang tidak terlalu sukar dan tidak terlalu mudah. Hasil penelitian ini sejalan dengan hasil penelitian Ndiung dan Jediut (2020) yang menunjukkan bahwa instrumen tes hasil belajar matematika peserta didik sekolah dasar berorientasi pada berpikir tingkat tinggi yang dikembangkan semua memiliki tingkat kesukaran pada kategori sedang dan indeks daya beda semua item berkategori baik yang berarti instrumen tersebut dapat digunakan dapat diandalkan untuk sebagai instrumen pengumpulan data dalam suatu penelitian.

Selain itu, Novitasari., et al (2018) yang menunjukkan bahwa tingkat kesukaran instrumen tes literasi sains yang dianalisis memiliki tingkat kesukaran dengan kategori sedang sebanyak 100% yang berarti tingkat kesukaran instrumen baik dan dapat digunakan.

Ditinjau dari keefektifan, instrumen kemampuan berpikir tingkat tinggi yang dikembangkan cukup efektif untuk meningkatkan kemampuan berpikir tingkat tinggi. Namun selisih skor pretest dan posttest masih rendah karena kemampuan berpikir tingkat tinggi bukan kemampuan yang mudah di dapat tanpa proses latihan yang rutin. Woolf (2005) menyatakan bahwa dibutuhkan latihan yang tidak singkat untuk mengubah keterampilan berpikir seseorang. Oleh karena itu kemampuan berpikir peserta didik dapat dilatih melalui latihan-latihan dengan instrumen yang relevan.

Sedangkan instrumen literasi sains masih kurang efektif Hal ini karena instrumen literasi sains yang dikembangkan berupa tes uraian yang menuntut peserta didik untuk memberikan jawaban dalam bentuk argumen dan menuntut pemahaman peserta didik tentang hubungan sains dengan kehidupan sehari-hari mereka. Osborne (2007) menyatakan argumen dalam pembelajaran sains, dapat meningkatkan kemampuan peserta didik dalam menyampaikan pendapat.

Argumen-argumen peserta didik dalam menjawab soal mengindikasikan bahwa peserta didik selama ini terbiasa dengan soal-soal yang menuntut jawaban dalam bentuk hafalan dan ingatan. Hal ini karena soal-soal yang digunakan guru masih dalam level kognitif yang rendah. Ridwan (2013) menyatakan bahwa pembelajaran sains yang diberikan guru hanya ditekankan pada hafalan, ingatan maupun aplikasi rumus praktis.

Hasil penelitian ini sejalan dengan hasil penelitian Mardhiyyah., et al (2016) yang menunjukkan bahwa instrumen literasi sains yang dikembangkan valid dan reliabel namun profil literasi sains peserta didik masih rendah terutama dalam aspek sains, teknologi dan masyarakat.

KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan dapat disimpulkan bahwa: Pengembangan instrumen penilaian kemampuan berpikir tingkat tinggi dan literasi sains dilakukan dengan four D Model (model 4D) yang terdiri dari empat tahapan meliputi (1) define (pendefinisian); (2) design (perancangan); (3) develop (pengembangan); dan (4) disseminate (penyebaran). Validitas instrumen kemampuan berpikir tingkat tinggi dan literasi sains yang di kembangkan memenuhi syarat validitas isi dengan nilai CVI sebesar 0,79 untuk kemampuan berpikir tingkat tinggi dan 0,72 untuk literasi sains yang termasuk dalam kategori tinggi. Sedangkan validitas butir instrumen kemampuan berpikir tingkat tinggi dan literasi sains memenuhi syarat dengan 20 instrumen kemampuan berpikir tingkat tinggi dinyatakan valid dan 10 instrumen literasi sains dinyatakan valid. Reliabilitas instrumen kemampuan berpikir tingkat tinggi dan literasi sains yang di kembangkan memenuhi syarat reliabilitas dengan nilai 0,937 untuk instrumen kemampuan berpikir tingkat tinggi dan 0,940 untuk instrumen literasi sains termasuk dalam kategori sangat tinggi. Daya beda instrumen kemampuan berpikir tingkat tinggi dan literasi sains yang di kembangkan memenuhi syarat kelayakan instrumen yang baik. Tingkat kesukaran instrumen kemampuan berpikir tingkat tinggi dan literasi sains yang di kembangkan memenuhi syarat kelayakan instrumen yang baik. Keefektifan Instrumen kemampuan berpikir tingkat tinggi dan literasi sains yang dikembangkan memperoleh skor N-gain sebesar 59,9% termasuk dalam kategori cukup efektif dan untuk instrumen literasi sains memperoleh skor N-gain sebesar 45,3% termasuk dalam kategori kurang efektif.

REFERENSI

- Arikunto, S. (2013). *Dasar-Dasar Evaluasi Pendidikan*, Jakarta: Bumi Aksara.
- Antara, I.G.W., Arnyana, I.B.P. & Margunayasa, I.G. (2021). Pengembangan Instrumen Penilaian Berbasis HOTS (Keterampilan Berpikir Tingkat Tinggi) Pada Tema Ekosistem Siswa Kelas V SD.Singaraja. *Jurnal Pendidikan Dasar Indonesia*. Jilid 5, No. 4.
- Arifin, Z., dan Retnawati, H. (2017). Pengembangan Instrumen Pengukur Higher Order Thinking Skills Matematika Siswa SMA Kelas X. Pythagoras. *Jurnal Pendidikan Matematika*. Jilid 12, No. 1.
- Arifin, Z. (2017). Kriteria Instrumen dalam Suatu Penelitian. *Jurnal THEOREMS (The Original Research of Mathematics)*. Jilid 2, No. 1.
- Azwar, S. (2013). *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Pelajar.
- BSNP. (2006). *Standar Isi untuk Sekolah Menengah dan Dasar*. Jakarta : Badan Standar Nasional Pendidikan.
- Borg ,W.R. & Gall, M.D. (1989). *Educational Research: an Introduction*. New York: Longman.
- Dantes, N. (2012). *Metode Penelitian*. Yogyakarta: Andi Offset.
- Daryanto & Mulyo R. (2012). *Model Pembelajaran Inovatif*. Yogyakarta: Gava Media.
- Hake, R. (1998). Interactive Engagement Versus Traditional Methods: A Six Thousand Student Survey Of Mechanics Test Data For Introductory Physic Course. *American Association of Physics Teachers*. Jilid 66, No. 1.

- Handayani, S.N.P., Suarni, N.K. & Arnyana, I.B.P. (2021). Pengembangan Instrumen Kemampuan Berpikir Kritis dan Literasi Sains Siswa Kelas V SD. *Jurnal Pendidikan Dasar Indonesia*. Jilid 5, No.1.
- Hidayati, A. U. (2017). Melatih Keterampilan Berpikir Tingkat Tinggi dalam Pembelajaran Matematika pada siswa Sekolah Dasar.Yogyakarta.*Jurnal Pendidikan dan Pembelajaran Dasar*, 4 (2).
- Holbrook, J. & Rannikmae, M. (2007). The nature of science Education for teaching scientific literacy.*International Journal of Science Education*. Jilid 29 No. 11.
- Jamaluddin, A., *et al.* (2020). Development of Critical Thinking Skills Instruments in Learning Sains. *Jurnal Pijar MIPA*. Jilid 15 No. 1.
- Kemendikbud. (2017). Modul Penyusunan Higher Order Thinking Skill (HOTS). Jakarta: Direktort Jenderal Pendidikan Dasar dan Menengah Departemen Pendidikan dan Kebudayaan.
- Koyan, I.W. (2011). *Assesmen dalam Pendidikan*. Singaraja: Universitas Pendidikan Ganesha.
- Koyan, I. W. (2012). *Statistik Pendidikan Teknik Analisis Data Kuantitatif*. Singaraja: Universitas Pendidikan Ganesha.
- Mahirah, B.(2017). Evaluasi belajar peserta didik (siswa). *Jurnal Idaarah*. Jilid 1 No.2.
- Mardhiyyah, A.L., Rusilowati, A., Linuwih, S. (2016). Pengembangan Instrumen Asesmen Literasi Sains Tema Energi. *Journal of Primary Education*. Jilid 5, No.1.
- Mudjijo. (1995). *Tes Hasil Belajar*. Jakarta: Bumi Aksara
- Ndiung, S. & Mariana.J. (2020). Pengembangan Instrumen Tes Hasil Belajar Matematika Peserta Didik Sekolah Dasar Berorientasi pada Berpikir Tingkat Tinggi. Jilid 10, No.1.
- Novitasari, L., Handhika, J. & Huriawati, F. (2018). *Analisis Karakteristik Instrumen Tes Literasi Sains Berbasis Etnophisyc pada Materi Pokok Dinamika Rotasi*. Artikel disajikan dalam Prosiding Seminar Nasional Pendidikan Fisika IV. Universitas PGRI Madiun, 25 Juli.
- Osborne, J. (2007). Science Education For Twenty First Century. *Eurasia Journal of Mathematics and Science Education*. Jilid 3, No. 3.
- OECD.2014. PISA (2012) Result in Focus: What 15 – Year – Olds Know and What They can Do with What They Know (Student Performance in Mathematics, Reading and Sciences). Paris: OECD.
- 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_Insights_and_Interpretations_final_PDF.pdf)
- 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.
- Ridwan, S.M., Mardhiyyah, A.L. & Rusilowati, A. *Pengembangan Instrumen Asesmen Dengan Pendekatan Kontekstual Untuk Mengukur Level Literasi Sains Siswa*. Artikel disajikan dalam Seminar Nasional Evaluasi Pendidikan. Universitas Negeri Semarang, 13 Juli.

- Suastra, I. W. (2017). Pembelajaran Sains Terkini Mendekatakan Siswa dengan Lingkungan Alamiah dan Sosial Budayanya. Singaraja: Universitas Pendidikan Ganesha.
- Sudijono, A. (2006). Pengantar Statistik Pendidikan.: PT Raja Grafindo Persada.
- Thoha, C. (1996). *Teknik Evaluasi Pendidikan*. Jakarta: PT Raja Grafindo persada.
- Woolf, B. P. (2005). Distributed Interactive Learning Environments *New Pedagogy: Inquiry Learning*.