



PENGEMBANGAN INSTRUMEN ASESMEN BERBASIS *OPEN-ENDED QUESTION* PADA KURIKULUM MERDEKA UNTUK MENGUKUR KEMAMPUAN BERPIKIR KREATIF PESERTA DIDIK

Oleh

Rini Rifani¹, Teguh Prasetyo², Syukri Indra³, Wiworo Retnadi Rias Hayu⁴
Universitas Djuanda Bogor
E-mail: teguh@unida.ac.id

Diterima 14 September 2024, direvisi 2 Pebruari 2025, diterbitkan 3 Pebruari 2025

Abstrak

Penelitian ini bertujuan untuk mendeskripsikan pengembangan dan menganalisis kriteria instrumen penilaian berbasis pertanyaan terbuka dalam Kurikulum Merdeka, yang dirancang untuk mengukur keterampilan berpikir kreatif siswa sekolah dasar kelas 4. Penelitian ini mengadopsi model pengembangan instrumen sembilan langkah: mendefinisikan tujuan instrumen, menentukan cakupan materi, mengembangkan indikator item, menyusun item instrumen, melakukan validitas konten, merevisi instrumen, melakukan uji coba, menganalisis hasil uji coba, dan menyusun instrumen akhir. Analisis validitas konten menggunakan indeks Aiken V, yang dilakukan oleh empat validator, menunjukkan bahwa tujuh item penilaian dinyatakan valid dan lima lainnya tidak valid. Uji coba produk skala besar menghasilkan hasil sebagai berikut: (1) uji validitas 10 butir soal menunjukkan semua butir soal valid, (2) butir soal reliabel dengan Cronbach's Alpha sebesar 0,822 yang menunjukkan reliabilitas sangat tinggi, (3) analisis indeks kesukaran menunjukkan delapan butir soal memiliki tingkat kesulitan sedang dan dua butir soal memiliki tingkat kesulitan mudah, dan (4) analisis indeks daya pembeda menunjukkan dua butir soal memiliki tingkat kesulitan cukup baik, enam butir soal memiliki tingkat kesulitan baik, dan dua butir soal memiliki tingkat kesulitan sempurna.

Kata Kunci: Pengembangan, Instrumen Asesmen, Soal Terbuka, Berpikir Kreatif

Abstract

This study aims to describe the development and analyze the criteria for an Open-ended-question-based assessment instrument within the Merdeka Curriculum, designed to measure the creative thinking skills of 4th-grade elementary school students. The research adopts a nine-step instrument development model: defining the instrument's purpose, determining the scope of material, developing item indicators, constructing the instrument items, conducting content validity, revising the instrument, pilot testing,

analyzing the trial results, and assembling the final instrument. The content validity analysis using Aiken's V index, conducted by four validators, indicated that assessment seven items were declared valid and five others were invalid. Large-scale product testing produced the following results: (1) validity tests of 10 questions showed all questions were valid, (2) the instrument items were reliable with a Cronbach's Alpha of 0.822, indicating very high reliability, (3) difficulty index analysis revealed eight items were of medium difficulty and two items were easy, and (4) discrimination index analysis showed two items were moderately good, six items were good, and two items were perfect.

Keywords: Development, Assessment Instruments, Open-ended Questions, Creative Thinking

I. PENDAHULUAN

Kurikulum merupakan suatu rancangan dalam kegiatan pembelajaran yang terdapat banyak komponen yang saling berkaitan dan saling mendukung. Secara umum, kurikulum efektif harus terorganisir dengan baik, relevan dan terkini, serta mampu membekali peserta didik dengan pemahaman dan keahlian yang diperlukan agar berhasil dalam bidang pilihannya. Hingga saat ini, kurikulum Indonesia telah mengalami sebelas kali perubahan, tujuan perubahan kurikulum ini adalah untuk melakukan penyempurnaan dari kurikulum sebelumnya dan untuk saat ini Indonesia secara bertahap sedang menerapkan kurikulum merdeka. Kurikulum merdeka dikembangkan untuk mengatasi tantangan persaingan ketat sumber daya manusia yang ada secara mendunia di abad ke-21 (Amalia, 2022).

Penerapan Kurikulum Merdeka, menuntut adanya penyesuaian metode pembelajaran dan penilaian yang digunakan. Salah satu tantangan utama dalam penerapan kurikulum ini adalah mengukur kemampuan berpikir kreatif siswa secara efektif. Meskipun Kurikulum Merdeka dirancang untuk mendukung pengembangan keterampilan abad 21, termasuk berpikir

kreatif, instrumen penilaian yang relevan dan terstandar masih sangat terbatas. Hal ini terlihat dari temuan di lapangan yang menunjukkan bahwa guru pada jenjang sekolah dasar masih kesulitan mengukur kemampuan berpikir kreatif siswa secara akurat. Guru cenderung mengandalkan metode observasi dan rubrik yang kurang spesifik sehingga hasil penilaian kurang mencerminkan kemampuan siswa secara keseluruhan. Oleh karena itu, pengembangan instrumen penilaian berbasis *Open-Ended Question* pada Kurikulum Merdeka menjadi penting dalam memastikan kemampuan berpikir kreatif siswa dapat diukur secara objektif dan komprehensif. Penelitian ini diharapkan dapat mengisi kesenjangan tersebut dengan menyediakan instrumen asesmen pembelajaran yang valid, reliabel, dan praktis untuk mendukung penerapan Kurikulum Merdeka pada jenjang sekolah dasar.

Keterampilan berpikir kreatif (*Creative Thinking Skill*) merupakan kemampuan yang dimiliki seseorang dalam menjabarkan informasi baru dan memadukan gagasan atau pikiran yang unik untuk memecahkan suatu masalah (Qomariyah & Subekti, 2021). Panjaitan & Surya (2017), berpendapat bahwa

berpikir kreatif adalah berpikir secara intuitif, tajam, merangsang imajinasi, mengungkap banyak hal baru, menemukan gagasan-gagasan yang mengejutkan, serta menemukan inspirasi yang belum diketahui. Sedangkan Uloli (2021) berargumentasi bahwa berpikir kreatif bukan hanya menghasilkan ide-ide baru, namun juga cara-cara beragam dalam membangun dan mengkomunikasikan ide-ide tersebut

Namun, pada kenyataannya berpikir kreatif dalam pembelajaran masih sulit dilakukan oleh peserta didik. Hal ini diakibatkan oleh penggunaan metode pengajaran yang digunakan kurang tepat dan ketidakmampuan guru dalam membuat berbagai macam soal sehingga untuk menilai kemampuan berpikir kreatif, guru cenderung menilai dengan cara mengamati dan tidak menggunakan instrumen asesmen yang tepat. Terdapat berbagai macam cara untuk menilai berpikir kreatif peserta didik, satu diantaranya adalah dapat diadakan dengan memberikan pertanyaan berbentuk *Open-ended Question*. Menurut Hidayat & Widjajanti, (2018) kemampuan berpikir kreatif dapat dinilai dengan memberikan pertanyaan *Open-ended* kepada peserta didik

Asesmen pembelajaran memiliki peranan penting dalam menumbuhkan berpikir kreatif peserta didik. Jadi, untuk mengukur kemampuan tingkat berpikir kreatif peserta didik dibutuhkan adanya instrumen asesmen dengan soal yang dapat mendukung kemampuan berpikir kreatif peserta didik. Soal yang dapat mengukur kemampuan tingkat berpikir kreatif peserta didik adalah soal pertanyaan terbuka (*Open-ended Question*). Hanya saja penggunaan pertanyaan terbuka di SD belum banyak

dimanfaatkan bakal mengukur kemampuan tingkat berpikir kreatif peserta didik.

Berdasarkan temuan wawancara bersama guru SD Negeri Harjasari 2 Kota Bogor pada Tanggal 20-21 Desember 2023 didapatkan informasi, sesungguhnya dalam mengukur kemampuan tingkat berpikir kreatif peserta didik guru belum menggunakan instrumen asesmen *Open-ended Question*, guru hanya menilai tingkat berpikir kreatif peserta didik dengan ketersediaan rubrik yang terdapat dalam buku dan memberikan pertanyaan permantik di awal dan di akhir pembelajaran saja. Hal ini menunjukkan bahwa guru belum menggunakan instrumen asesmen yang relevan untuk menilai kemampuan tingkat berpikir kreatif peserta didik, karena instrumen yang digunakan guru belum sesuai dengan tujuan atau kemampuan berpikir kreatif yang akan diukur, serta ujian yang dipakai dalam instrumen belum dapat mengukur kemampuan tingkat berpikir kreatif, sehingga guru lebih cenderung menilai berpikir kreatif peserta didik dengan cara mengamati.

Adapun berdasarkan penelitian Yohana & Rachmadiyah (2023) dan Samaratunga & Kusuma (2023), memberikan informasi bahwa *Open-ended Question* di SD hanya dapat mengukur kemampuan tingkat berpikir kritis peserta didik saja. Selanjutnya penelitian yang berbeda dengan Wahyuni & Palupi (2022), menyebutkan bahwa *Open-ended Question* dapat menganalisis kemampuan tingkat berpikir kreatif peserta didik pada tingkat SD. Berdasarkan hal ini maka dapat ditarik kesimpulan bahwa masih sedikit penelitian tentang asesmen dengan jenis soal *Open-ended* yang digunakan pada

berpikir kritis khusus berpikir kreatif peserta didik. Selain itu, penelitian-penelitian terdahulu masih berfokus pada kurikulum 2013 belum berfokus pada kurikulum baru yaitu kurikulum merdeka di SD. Penelitian ini bertujuan untuk Pengembangan Instrumen Asesmen Berbasis *Open-ended Question* pada Kurikulum Merdeka untuk Mengukur Kemampuan Berpikir Kreatif Peserta Didik Kelas IV SD.

II. PEMBAHASAN

Metode Penelitian

Metode penelitian menggunakan metode pengembangan. Penelitian pengembangan ini menggunakan model pengembangan yang diadaptasi dari model pengembangan instrumen penelitian yang dibuat oleh Retnawati (2016) yang terdiri dari 9 tahap, yaitu: (1) Menentukan tujuan penyusunan instrumen, (2) menentukan ruang lingkup materi, (3) mengembangkan indikator item instrumen, (4) mengembangkan item instrumen, (5) validitas isi, (6) revisi instrumen, (7) percobaan instrumen, (8) menganalisis hasil percobaan, dan (9) perakitan Instrumen.

Terdapat dua tahapan desain uji coba yaitu tahapan validasi ahli dan tahap percobaan. Adapun untuk subjek coba dilakukan kepada seluruh peserta didik kelas IV SD Negeri Harjasari 2 Kota Bogor dengan jumlah 10 peserta didik untuk percobaan instrumen skala kecil, dan 30 peserta didik percobaan instrumen skala besar. Sedangkan untuk teknik dan pengumpulan data menggunakan wawancara, angket dan instrumen asesmen *Open-ended Question* berpikir kreatif. Untuk

analisis data menggunakan teknik sebagai berikut:

a. Teknik Validasi Instrumen Asesmen

Validasi instrumen asesmen dilakukan dengan cara membuat dan menyerahkan lembar penilaian yang berisikan persoalan kepada validator. Tujuan lembar penilaian diserahkan kepada validator ini adalah untuk memberikan penilaian tentang layak atau tidaknya instrumen yang sudah dibuat (Apipah et al., 2023). Setelah lembar penilaian diserahkan, selanjutnya validator mengisi lembar tersebut dengan memberikan penilaian yang sesuai dengan ketersediaan skor sebagai berikut:

Tabel. 1 Skor Kriteria Kelayakan Validasi

Kategori	Skor
Sangat Relevan	4
Cukup Relevan	3
Kurang Relevan	2
Tidak Relevan	1

(Sumber: Retnawati, 2016)

Selanjutnya skor yang diperoleh kemudian dihitung menggunakan rumus indeks aiken berikut:

$$V = \frac{\sum S}{n(c - 1)}$$

Keterangan:

- V : Indeks kesepakatan validator
- s : Skor yang ditetapkan validator dikurangi skor terendah
- n : Banyaknya validator
- c : Banyaknya kategori yang dipilih validator

Validasi instrumen asesmen dihitung menggunakan bantuan program Microsoft Excel 2019 dengan acuan yang didasarkan pada Aiken dalam tulisan Restu et al., (2022) validitas suatu item instrumen dapat dikatakan baik dengan 4 orang penilai

dengan 4 kategori jawaban adalah jika diperoleh indeks aiken lebih besar atau sama dengan 0,92 artinya item instrumen tersebut valid, namun sebaliknya, jika indeks aiken kurang dari 0,92 artinya item tersebut tidak valid.

b. Teknik Analisis Validitas

Analisis validitas empiris ditetapkan oleh pengujian item instrumen yang diukur berdasarkan percobaan. analisis validasi dilakukan dengan bantuan program *Microsoft Excel 2019*. Setiap item instrumen dapat dilihat tingkat validitasnya dengan mengamati angka *corrected item-total correlation*, yaitu perbandingan nilai rhitung dengan r_{tabel} . Didasarkan pada kriteria ketentuan validitas jika (Slamet & Wahyuningsih, 2014):

- 1) $r_{hitung} \geq r_{tabel}$ artinya item instrumen valid.
- 2) $r_{hitung} < r_{tabel}$ artinya item instrumen tidak valid.

c. Teknik Analisis Reliabilitas

Analisis derajat konsistensi dilakukan dengan menggunakan uji *internal consistency* dengan teknik *Cronbach Alpha*. Analisis reliabilitas dihitung menggunakan program bantuan SPSS 16.0. Item instrumen dapat dinyatakan reliabel apabila jumlah *Cronbach Alpha* $\geq 0,70$. Adapun menurut Arikunto kriteria interpretasi derajat konsistensi dijelaskan pada Tabel 3 di bawah.

Berikut rumus untuk mencari daya beda (*Indeks Diskriminasi*):

$$D = \frac{\bar{x}_A - \bar{x}_B}{x_{maks}}$$

Keterangan:

D : Daya beda soal

\bar{x}_A : Skor rata-rata peserta didik berkemampuan tinggi

\bar{x}_B : skor rata-rata peserta didik berkemampuan rendah

X_{maks} : skor maksimum yang ditetapkan

Hasil Penelitian dan Pembahasan

1. Menentukan Tujuan Penyusunan Instrumen

Langkah pertama pengembangan instrumen asesmen ini yaitu menetapkan tujuan penyusunan instrumen. Penyusunan instrumen bertujuan untuk: (1) Menghasilkan instrumen asesmen berbasis *Open-ended Question* pada kurikulum merdeka untuk mengukur tingkat kemampuan berpikir kreatif peserta didik kelas IV, dan (2) Menganalisis kriteria instrumen asesmen berbasis *Open-ended Question* pada kurikulum merdeka untuk mengukur kemampuan tingkat berpikir kreatif peserta didik kelas IV.

2. Menentukan Ruang Lingkup Materi

Materi yang dikembangkan yaitu materi pada Bab VI “Indonesiaku Kaya Budaya” dalam mata pelajaran IPAS dengan cakupan materi kearifan lokal, keragaman budaya dan manfaat upaya pelestarian budaya. Pemilihan materi ini berdasarkan: (1) Saran dan Masukan guru kelas IV SD, hal ini disebabkan karena mata pelajaran IPAS dinilai sangat membutuhkan instrumen asesmen *Open-ended Question* bakal mengukur kemampuan berpikir kreatif peserta didik kelas IV, dan (2) materi tersebut sesuai dengan keadaan materi pembelajaran yang sedang berlangsung di kelas IV semester 2. Untuk ruang

lingkup materi lebih jelasnya ditampilkan pada Tabel 6 di bawah.

Tabel. 2 Cakupan Materi Instrumen Asesmen Open Ended Question

Ruang Lingkup Materi	
Pokok Materi	Indonesiaku Kaya Budaya
Sub Pokok Materi	<ul style="list-style-type: none"> • Kearifan Lokal • Keragaman budaya • Manfaat dan upaya pelestarian budaya
Capaian Pembelajaran	Peserta didik mendeskripsikan keragaman budaya dan kearifan lokal serta upaya pelestariannya.
Tujuan Pembelajaran	<ul style="list-style-type: none"> • Mendeskripsikan keragaman budaya dan kearifan lokal di daerahnya masing-masing. • Mengetahui manfaat dan pelestarian keragaman budaya di Indonesia

3. Mengembangkan Indikator Item Instrumen

Indikator item instrumen yang dipakai adalah indikator kemampuan berpikir kreatif. Indikator berpikir kreatif yang dipakai peneliti merupakan indikator yang dapat dilihat serta diamati dari perolehan tes percobaan instrumen. Merujuk pada Munandar dalam tulisan Qomariyah & Subekti (2021), sebagai hasilnya terdapat 4 indikator yang dipilih, antara lain: (1) Kelancaran, (2) Keluwesan, (3) Keaslian, dan (4) Kerincian. Indikator tersebut kemudian disusun kedalam kisi-kisi item instrumen supaya peredarannya lebih merata selanjutnya item instrumen disusun sesuai dengan kisi-kisi yang sudah tersedia. Dari mengembangkan indikator item instrumen ini didapatkan hasil kisi-kisi instrumen tes berpikir kreatif.

4. Mengembangkan Item Instrumen

Langkah selanjutnya yaitu mengembangkan item instrumen, item instrumen atau soal yang terdapat dalam

penelitian ini berbentuk tes uraian dengan jenis *Open Ended Question* yang berjumlah 12 item instrumen dengan alokasi waktu 90 menit. Seluruh item instrumen yang dibuat sudah tersedia kunci jawaban yang berguna untuk mempermudah guru dalam pengoreksian hasil pengerjaan peserta didik dalam menyelesaikan soal, serta sudah tersedia pedoman penskoran yang berisi tentang kriteria jawaban dan perincian skor yang diberikan kepada peserta didik untuk persoalan yang telah dikerjakan.

5. Validitas Isi

Setelah instrumen asesmen dibuat kemudian peneliti menyerahkan produk yang dikembangkan dan lembar validasi kepada 4 ahli atau validator untuk melakukan penilaian. Penilaian ini dilakukan guna mengetahui kelayakan rancangan produk yang dikembangkan. Adapun daftar nama validator instrumen ditampilkan pada Tabel 7 sebagai berikut:

Tabel. 3 Daftar Nama Validator Instrumen Asesmen Open Ended Question

Nama Validator	Jabatan	Tanggal Validasi
Annisa Mawardini, S.ST, M.Pd.	Dosen PGSD Universitas Djuanda	31 Maret 2024
H. Muhammad Ichsan, M.Pd.	Dosen PGSD Universitas Djuanda	28 Maret 2024
Muhamad Iswandi, S.Pd.	Guru Kelas IV SD Negeri Harjasari 2	25 Maret 2024
Meriatmi Lestari, S.Pd.	Guru Kelas IV SD Negeri Harjasari 2	18 Maret 2024

Dari kegiatan penilaian validasi yang telah dilaksanakan, didapatkan nilai validitas isi dari 4 validator menggunakan rumus indeks aiken dapat dilihat pada Tabel 8 sebagai berikut:

Tabel. 4 Hasil Analisis Validasi Ahli

Skala Butir	Nilai V	Niali V Tabel	Ket
1	0,739	0,92	Tidak Valid
2	0,856	0,92	Tidak Valid

3	0,928	0,92	Valid
4	0,794	0,92	Tidak Valid
5	0,922	0,92	Valid
6	0,922	0,92	Valid
7	0,939	0,92	Valid
8	0,806	0,92	Tidak Valid
9	0,939	0,92	Valid
10	0,922	0,92	Valid
11	0,794	0,92	Tidak Valid
12	0,950	0,92	Valid

Berdasarkan perolehan nilai yang dihitung menggunakan rumus perhitungan indeks aiken menunjukkan terdapat 7 item instrumen memiliki nilai *indeks v* $\geq 0,92$ yang artinya valid, dan 5 item instrumen lainnya memiliki nilai *indeks v* $< 0,92$ yang artinya tidak valid. Dari hasil tersebut berarti item dalam instrumen asesmen layak untuk digunakan, akan tetapi perlu diperhatikan bahwa, terdapat beberapa hal yang harus diperbaiki guna memperoleh kualitas instrumen asesmen yang lebih baik.

6. Revisi

Langkah selanjutnya adalah revisi instrumen. Revisi instrumen didasarkan pada hasil validasi isi yang dilakukan oleh ahli dan percobaan instrumen skala kecil. Berikut hasil penjelasan tiap-tiap revisi:

a. Revisi Produk Setelah Validasi Isi

Selain untuk menilai produk instrumen asesmen, validasi isi juga bertujuan untuk mendapatkan komentar dan saran berdasarkan penilaian judgment dari keempat validator atau ahli dengan melihat aspek materi, bahasa, dan instrumen. Dari komentar dan saran inilah nantinya dijadikan landasan untuk melakukan revisi.

Pertama, terdapat masukan dari ahli untuk kisi-kisi instrumen tes, dimana untuk tingkat berpikir kreatif yang ada pada soal nomor 2 dan 9 seharusnya berada pada ranah kognitif C4 (menganalisis) bukan C6 (menciptakan).

Kedua, masukan diberikan validator untuk item soal tes yang ada pada instrumen asesmen. Dimana untuk item instrumen 1, 3, 4, 5, dan 10 terdapat rumusan pertanyaan kurang

relevan atau kurang jelas. Adapun masukan untuk item soal nomor 6 gambar yang digunakan usahakan harus jelas dan berfungsi. Lalu, untuk item soal nomor 2, 3, 6, 11, dan 12 penggunaan huruf kapital dan penulisan perlu diperhatikan. Untuk item soal nomor 3 dan 4 beberapa istilah sulit harus disederhanakan atau dicari padanannya. Kemudian untuk item soal nomor 7, 8, dan 9 penggunaan tanda baca perlu diperhatikan. Selanjutnya untuk item soal nomor 5, 7, dan 8 ada beberapa kata yang perlu dihilangkan atau diperbaiki. Serta untuk item soal nomor 1, 4, dan 5 terdapat kalimat pertanyaan yang harus diubah.

Ketiga, masukan validator berikan untuk kunci jawaban instrumen asesmen. Dimana kunci jawaban pada item soal nomor 2 dan 7 terdapat kata dan jawaban yang perlu disederhanakan.

Dari revisi produk setelah validasi ahli ini secara keseluruhan masukan ditunjukkan hanya untuk kisi-kisi, soal dan kunci jawaban saja. Namun, untuk instrumen asesmen bagian pedoman penskoran walaupun tidak dapat masukan dari validator tetap harus direvisi karena terdapat beberapa perubahan pada pertanyaan item soal dan juga kunci jawaban. Maka berdasarkan revisi produk setelah validasi isi ini didapatkan hasil produk awal pengembangan instrumen asesmen yang mempunyai kualitas baik dan layak digunakan pada percobaan skala kecil.

b. Revisi Produk Setelah Percobaan instrumen

Revisi produk setelah percobaan instrumen merupakan revisi berdasarkan hasil analisis kriteria pada percobaan skala kecil dan hasil wawancara 10 peserta didik kelas IV pada percobaan skala kecil. Berdasarkan hasil analisis percobaan skala kecil terdapat 4 item instrumen tidak memenuhi kriteria pada item

instrumen soal 1,2,5 dan 9 karena validitas item instrumen tidak valid.

Dari 4 item instrumen yang tidak memenuhi kriteria tersebut, untuk item soal 1 dan 2 diputuskan untuk dihilangkan karena selain item soal tidak valid, pada item soal nomor 1 juga memiliki indeks daya beda yang jelek, serta item soal nomor 2 juga dihilangkan karena tidak sesuai dengan proses pembelajaran peserta didik. Adapun untuk item soal 5 dan 9 diputuskan untuk direvisi karena item soal nomor 5 memiliki daya beda yang cukup dan lebih tinggi dari keempat soal tersebut, sedangkan soal nomor 9 diputuskan direvisi karena sesuai dengan proses pembelajaran peserta didik akan tetapi rumusan soal tersebut belum jelas.

Adapun revisi produk percobaan instrumen berdasarkan hasil wawancara dengan 10 peserta didik kelas IV adalah terdapat beberapa kata atau frasa yang kurang jelas dan sulit dimengerti dalam soal, akibatnya peserta didik tidak mampu memahami pertanyaan dan kebingungan dalam menjawab pertanyaan.

Maka dari itu peneliti melakukan revisi terhadap beberapa soal yang dianggap peserta didik kurang jelas dan sulit dimengerti, diantaranya yaitu sebagai berikut: Pertama, untuk soal nomor 3 peserta didik mengalami kebingungan dalam pilihan yang diberikan dalam soal karena tidak terdapat angka dalam pilihan yang disediakan. Kedua, untuk soal nomor 4 peserta didik tidak mengerti kata gawai dan budaya pop dalam soal, maka peneliti memperjelas apa yang dimaksud kata tersebut dalam teks cerita untuk menjawab soal nomor 4. Ketiga, untuk soal nomor 5 peserta didik cenderung mengabaikan ketentuan yang terdapat dalam soal maka dari itu peneliti lebih menekankan lagi ketentuan yang ada dalam soal tersebut. Keempat, untuk soal nomor 9 perlu

diperbaiki karena rumusan soal dalam pertanyaan nomor 9 belum jelas.

Berdasarkan revisi setelah percobaan instrumen skala kecil, maka didapatkan hasil instrumen asesmen dengan item soal yang berjumlah 10 baik dan layak digunakan untuk percobaan instrumen skala besar.

7. Percobaan Instrumen

Instrumen asesmen sudah dibuat, divalidasi, dan direvisi selanjutnya di coba dalam percobaan instrumen skala kecil dan percobaan instrumen skala besar. Tahap uji percobaan instrumen skala kecil ini berfungsi untuk mengkaji ulang hasil penilaian dari para validator dan untuk mengetahui saran maupun komentar peserta didik terhadap instrumen asesmen yang sudah dikembangkan. Percobaan uji instrumen skala kecil dilaksanakan pada hari Rabu tanggal 8 Mei 2024 kepada 10 peserta didik kelas IV SD Negeri Harjasari 2 Kota Bogor dengan peserta didik berkemampuan tinggi (4), sedang (4), dan rendah (2). Berdasarkan percobaan instrumen skala kecil ini terdapat 12 item soal yang digunakan dengan alokasi waktu 90 menit.

Dari hasil percobaan instrumen skala kecil didapatkan hasil peserta didik mampu menyelesaikan soal dengan kurun waktu 90 menit dan hasil pengukuran kemampuan berpikir kreatif 10 peserta didik, dimana 3 orang peserta didik berada dalam kategori sangat kreatif, dan 6 orang dalam kategori kreatif, serta 1 orang dalam kategori cukup kreatif.

Selanjutnya, tahap percobaan instrumen skala besar dilakukan guna mendapatkan informasi mengenai keefektifan perubahan yang sudah dilakukan berdasarkan hasil revisi percobaan uji skala kecil. Percobaan instrumen skala besar dilaksanakan pada hari Rabu tanggal 22 Mei 2024 kepada seluruh peserta didik kelas IV SD Negeri Harjasari 2 Kota Bogor yang berjumlah 30 peserta didik. Untuk

percobaan instrumen skala besar terdapat 10 item soal yang digunakan dengan alokasi waktu 90 menit.

Berdasarkan percobaan instrumen skala besar ini didapatkan hasil peserta didik mampu menyelesaikan soal dengan kurun waktu kurang lebih 90 menit dan dari 30 peserta didik terdapat 4 orang peserta didik dalam kategori sangat kreatif, 21 orang dalam kategori kreatif, 4 peserta didik cukup kreatif, dan 1 orang dalam kategori tidak kreatif.

8. Analisis Hasil Tes

Uji coba instrumen asesmen sudah dilakukan, langkah selanjutnya adalah menganalisis hasil tes uji coba. Analisis hasil tes digunakan untuk mendapatkan informasi mengenai setiap kriteria kualitas item soal. Analisis kriteria kualitas item instrumen dibantu program *Microsoft Excel 2019* dan *SPSS 16.0*. Berikut analisis hasil tes percobaan instrumen skala kecil dan percobaan instrumen skala besar.

a. Analisis Hasil Tes Percobaan Skala Kecil

Pada percobaan uji skala kecil digunakan 12 item soal yang digunakan. Dari percobaan skala kecil ini diperoleh hasil analisis tes sebagai berikut:

1) Validitas Item Soal

ketentuan validitas didasarkan pada nilai r_{hitung} (*Corrected Item-Total Correlation*) $> r_{tabel}$ sebesar 0,632, untuk $df = 10 - 2 = 8$ dengan tingkat signifikansi 0,05 maka item instrumen tersebut dapat dikatakan valid dan sebaliknya. Berikut hasil analisis data validitas item instrumen percobaan skala kecil ditampilkan pada Tabel 9 di bawah.

Tabel. 5 Hasil Analisis Percobaan Instrumen Skala Kecil

No. Item	r_{hitung}	r_{tabel}	Ket.
1	0,240	0,632	Tidak Valid
2	0,417	0,632	Tidak Valid
3	0,756	0,632	Valid
4	0,857	0,632	Valid
5	0,417	0,632	Tidak Valid
6	0,637	0,632	Valid

7	0,774	0,632	Valid
8	0,723	0,632	Valid
9	0,417	0,632	Tidak Valid
10	0,681	0,632	Valid
11	0,808	0,632	Valid
12	0,704	0,632	Valid

Berdasarkan Tabel 9 maka dapat dilihat dari 12 item instrumen tersebut, terdapat 4 item soal diantaranya tidak valid, pada soal 1, 2, 5, dan 9 karena $r_{hitung} < r_{tabel}$, jika merujuk pada hasil validasi ahli untuk soal nomor 1 dan 2 menunjukkan hasil yang sama yaitu tidak valid, sedangkan untuk item soal 5 dan 9 berbanding terbalik dengan hasil validasi ahli maka berdasarkan hal ini kelayakan butir soal perlu dipertimbangkan. Adapun hasil validitas uji coba kelompok kecil 8 soal lainnya dinyatakan valid yaitu soal nomor 3, 4, 6, 7, 8, 10, 11, dan 12 karena $r_{hitung} \geq r_{tabel}$, untuk soal nomor 3, 6, 7, 8, 10 dan 12 sejalan dengan hasil validasi ahli maka berdasarkan hal ini soal yang valid dinyatakan bagus digunakan untuk uji coba selanjutnya. Untuk soal yang valid dalam validitas butir soal dapat digunakan pada uji coba selanjutnya, sedangkan soal yang tidak valid nantinya tidak digunakan lagi atau diperbaiki berdasarkan pertimbangan analisis lain.

2) Reliabilitas Item Soal

Konsistensi jawaban item soal dilakukan dengan menggunakan uji *internal consistency* dengan teknik *Cronbach Alpha*. Berikut hasil reliabilitas item soal ditampilkan pada Tabel 10 berikut:

Tabel. 6 Hasil Reliabilitas Percobaan Instrumen Skala Kecil

Reliability Statistics	
Cronbach's Alpha	N of Items
.841	12

Nilai *Cronbach Alpha* pada Tabel 10 sebesar 0,841 sehingga disimpulkan bahwa soal yang terdapat

pada instrumen asesmen adalah reliabel, dengan kategori sangat tinggi.

3) Indeks Kesukaran Item Soal

Tingkat kesukaran item soal dihitung dengan cara membagi rata-rata (*mean*) skor tiap item soal dengan skor maksimal. Berikut perolehan analisis tingkat kesukaran item soal ditampilkan pada Tabel 11 di bawah.

Tabel. 7 Tingkat Kesukaran Percobaan Instrumen Skala Kecil

No. Item	Indeks Proporsi	Kategori
1	0,567	Sedang
2	0,700	Sedang
3	0,800	Mudah
4	0,667	Sedang
5	0,733	Sedang
6	0,900	Mudah
7	0,700	Sedang
8	0,867	Mudah
9	0,633	Sedang
10	0,800	Mudah
11	0,833	Mudah
12	0,800	Mudah

Tabel 11 menunjukkan item instrumen 1, 2, 4, 5, 7, dan 9 dalam kategori sedang. Sedangkan item instrumen 3, 6, 8, 10, 11, serta 12 dalam kategori mudah. Untuk kategori sukar pada percobaan uji skala kecil tidak ditemukan. Jika mengacu pada pendapat (Widiyanto, 2018), soal baik adalah soal yang dianggap tidak terlalu mudah atau terlalu sukar dan untuk item soal mudah atau sukar bukan berarti tidak boleh digunakan. Maka, berdasarkan hal ini soal pada kategori mudah boleh digunakan dengan beberapa soal yang harus diganti atau revisi dengan mempertimbangkan kegunaannya.

4) Indeks Diskriminasi Item Soal

Indeks Diskriminasi atau daya beda item soal didapatkan dari pengurangan rata-rata skor peserta didik berkemampuan tinggi dengan rata-rata skor peserta didik berkemampuan rendah lalu dibagi dengan skor maksimal yang sudah ditetapkan. Berikut data output indeks diskriminasi

item soal ditampilkan pada Tabel 12 di bawah.

Tabel. 8 Hasil Daya Beda Percobaan Instrumen Skala Kecil

No. Item	Indeks Diskriminasi	Kategori
1	0,080	Jelek
2	0,286	Cukup
3	0,669	Baik
4	0,815	Sangat Baik
5	0,306	Cukup
6	0,575	Baik
7	0,723	Sangat Baik
8	0,646	Baik
9	0,286	Cukup
10	0,620	Baik
11	0,750	Sangat Baik
12	0,603	Baik

Tabel 12 menunjukkan bahwa soal nomor 1 memiliki kategori jelek, untuk item nomor 2, 5 dan 9 memiliki kategori cukup, dan untuk item nomor 3, 6, 8, 10, dan 12 memiliki kategori baik, serta untuk soal nomor 4, 7, dan 11 dalam kategori sangat baik. Merujuk pada pendapat Fatimah (2019), soal yang berada dalam kategori jelek atau $\leq 0,20$ dapat diperbaiki atau dibuang. Namun, dalam hal ini peneliti memilih untuk membuang soal tersebut selain memiliki daya beda yang jelek soal nomor 1 juga tidak valid.

b. Analisis Hasil Tes Percobaan Skala Besar

Pada percobaan uji skala besar terdapat 10 item soal yang digunakan. Dari percobaan skala besar ini didapatkan hasil analisis tes sebagai berikut:

1) Validitas Item Soal

Pengambilan keputusan validitas didasarkan pada nilai $r_{hitung} > r_{tabel}$ sebesar 0,361, untuk $df = 30 - 2 = 28$ dengan tingkat signifikansi 0,05 maka item soal tersebut dinyatakan valid dan sebaliknya. Berikut hasil analisis data validitas item soal percobaan skala kecil ditampilkan dalam Tabel 13 di bawah:

Tabel. 9 Hasil Validitas Percobaan Instrumen Skala Besar

No. Item	r_{hitung}	r_{tabel}	Ket.
----------	--------------	-------------	------

No. Item	r_{hitung}	r_{tabel}	Ket.
1	0,378	0,361	Valid
2	0,435	0,361	Valid
3	0,678	0,361	Valid
4	0,771	0,361	Valid
5	0,595	0,361	Valid
6	0,790	0,361	Valid
7	0,680	0,361	Valid
8	0,803	0,361	Valid
9	0,555	0,361	Valid
10	0,572	0,361	Valid

Tabel 13 menunjukkan bahwa seluruh item soal memiliki status valid, hal ini sejalan dengan beberapa hasil validasi ahli maka artinya butir soal yang terdapat dalam instrumen asesmen bagus. Dari hasil perhitungan validitas butir soal maka dapat ditarik kesimpulan bahwa seluruh butir soal yang ada pada instrumen asesmen uji coba kelompok besar dinyatakan valid.

2) Reliabilitas Item Soal

Reliabilitas item soal tes instrumen asesmen *Open-ended Question* untuk mengukur kemampuan berpikir kreatif peserta didik dibantu menggunakan program *SPSS 16.0* dengan perolehan data output sebagai berikut:

Tabel. 10 Hasil Reliabilitas Percobaan Instrumen Skala Besar

Reliability Statistics	
Cronbach's Alpha	N of Items
.822	10

Pada Tabel 15 diperoleh hasil *Alpha Cronbach* senilai 0,822 sehingga dapat disimpulkan bahwa soal yang terdapat pada instrumen asesmen adalah reliabel, dengan kategori sangat tinggi.

3) Indeks Kesukaran Item Soal

Indek kesukaran item soal dianalisis dengan bantuan program *Microsoft Excel 2019* berikut hasil outputnya ditampilkan pada Tabel 15 di bawah.

Tabel. 11 Hasil Tingkat Kesukaran Percobaan Instrumen Skala Besar

No. Item	Indeks Proporsi	Kategori
1	0,722	Sedang
2	0,533	Sedang
3	0,578	Sedang

4	0,778	Sedang
5	0,800	Mudah
6	0,776	Sedang
7	0,433	Sedang
8	0,700	Mudah
9	0,833	Sedang
10	0,478	Sedang

Berdasarkan Tabel 15 terlihat bahwa tingkat kesukaran item instrumen yang terdapat pada 10 item soal diketahui 2 soal berada dalam kategori mudah pada item soal nomor 5 dan 9, sedangkan 8 item lainnya dalam kategori sedang yaitu item instrumen nomor 1, 2, 3, 4, 6, 7, 8, dan 10, adapun untuk kategori sukar dalam analisis percobaan skala besar ini hasilnya sama dengan percobaan uji skala kecil yaitu tidak ditemukan. Dari data output analisis tingkat kesukaran dapat ditarik kesimpulan bahwa soal pada instrumen asesmen dinyatakan layak digunakan dengan pertimbangan urgensi atau keberuntungan peneliti.

4) Indeks Diskriminasi Item Soal

Indeks diskriminasi item soal dianalisis menggunakan program *SPSS 16.0*. berikut hasil outputnya ditampilkan pada Tabel 16 di bawah.

Tabel. 12 Hasil Daya Beda Percobaan Instrumen Skala Besar

No. Item	Indeks Diskriminasi	Kategori
1	0,219	Cukup
2	0,287	Cukup
3	0,569	Baik
4	0,691	Baik
5	0,481	Baik
6	0,717	Sangat Baik
7	0,596	Baik
8	0,738	Sangat Baik
9	0,422	Baik
10	0,434	Baik

Tabel 16 menunjukkan bahwa item soal nomor 1 dan 2 dalam kategori cukup, untuk item instrumen 3, 4, 5, 7, 9, dan 10 dalam kategori baik, serta untuk soal nomor 6 dan 8 dalam kategori sangat baik. Merujuk pada pendapat Fatimah (2019), terkait soal

yang memiliki kategori cukup, baik dan sangat baik maka item instrumen harus dimasukkan kedalam bank soal dan dapat dipakai kembali. Dengan demikian, maka dapat disimpulkan item instrumen yang terdapat pada percobaan skala besar dinyatakan layak untuk digunakan.

9. Perakitan Instrumen

Berdasarkan percobaan instrumen skala besar didapatkan hasil instrumen asesmen dengan nilai analisis yang baik, maka pada tahap perakitan instrumen asesmen ini didapatkan hasil produk akhir dari instrumen asesmen *Open-ended Question* pada kurikulum merdeka untuk mengukur kemampuan berpikir kreatif peserta didik kelas IV SD, produk akhir instrumen asesmen ini sudah diuji baik dari segi pemakaian dan kelayakan sehingga instrumen asesmen yang telah dikembangkan baik digunakan sebagai petunjuk untuk membantu guru dalam melakukan penilaian dan pengukuran tingkat kemampuan berpikir kreatif peserta didik. Hasil perakitan instrument asesmen pembelajaran ini dapat menjadi contoh bagi guru di SD, karena tugas pendidik untuk mempunyai kemampuan serta keterampilan yang baik dalam penilaian pembelajaran, karena hal ini dapat membantu pendidik dalam memutuskan teknik asesmen mana yang akan digunakan (Prasetyo, 2017).

III. SIMPULAN

Produk instrumen asesmen *Open-ended Question* pada kurikulum merdeka untuk mengukur kemampuan berpikir kreatif peserta didik kelas IV dengan bentuk terdiri dari: (1) Kisi-kisi instrumen asesmen kemampuan berpikir kreatif, (2) Soal Tes bentuk uraian dengan jenis *Open-ended Question* yang berjumlah 10 item soal, (3) Kunci jawaban, dan 4) pedoman penskoran yang sudah terperinci.

Saran penelitian yang dapat diberikan yaitu: (1) Bagi guru, dengan adanya instrumen asesmen *Open-ended Question* pada kurikulum merdeka untuk mengukur kemampuan berpikir kreatif peserta didik kelas IV SD, diharapkan dapat digunakan oleh guru sebagai instrumen asesmen yang dapat membantu guru dalam melakukan penilaian pembelajaran untuk berpikir kreatif peserta didik. (2) Bagi penelitian yang selanjutnya, diharapkan dapat mengembangkan instrumen asesmen berbasis *Open-ended Question* pada mata pelajaran dan materi yang berbeda.

DAFTAR PUSTAKA

- Amalia, M. (2022). Inovasi Pembelajaran Kurikulum Merdeka Belajar Di Era Society 5.0 Untuk Revolusi Industri 4.0. *Seminar Nasional Sosial Sains, Pendidikan, Humaniora (SENASSDRA) Volume, 1(1-6)*, 1-6. <http://prosiding.unipma.ac.id/index.php/SENASSDRA>
- Apipah, I., Nindiasari, H., & Sukirwan, S. (2023). Pengembangan Instrumen Soal Literasi Numerasi pada Materi Bilangan untuk Meningkatkan Kemampuan Berpikir Tingkat Tinggi Siswa Kelas VIII MTs. *Jurnal Cendekia : Jurnal Pendidikan Matematika*, 7(3), 3083-3092. <https://doi.org/10.31004/cendekia.v7i3.2606>
- Daryanes, F., & Ririen, D. (2020). Efektivitas Penggunaan Aplikasi Kahoot Sebagai Alat Evaluasi pada Mahasiswa. 3(2), 172-186.
- Fatimah, U. L. (2019). Analisis Kesukaran Soal, Daya Pembeda dan Fungsi Distraktor. 8, 37-64. <https://doi.org/https://doi.org/10.36668/jal.v8i2.115>
- Hidayat, P. W., & Widjajanti, D. B. (2018). Analisis kemampuan berpikir kreatif dan minat belajar siswa dalam mengerjakan soal *open ended dengan pendekatan CTL An analysis of creative thinking ability and learning interest of students of junior high school in solving open ended problem with CTL app*. 13(1), 63-75.
- Panjaitan, A. H., & Surya, E. (2017). *Creative Thinking (Berpikir Kreatif) dalam Pembelajaran*. December.
- Prasetyo, T. (2017). Pengembangan Perangkat Penilaian Hasil Belajar dalam Pembelajaran Tematik Integratif Kelas V SD. *Jurnal Prima Edukasia*, 5(1), 102-111. <https://doi.org/10.21831/jpe.v5i1.7528>
- Qomariyah, D. N., & Subekti, H. (2021). Analisis Kemampuan Berpikir Kreatif Studi Eksplorasi Siswa di SMPN 62 Surabaya. *PENSA E-JURNAL: Pendidikan Sains*, 9(2), 242-246. <https://ejournal.unesa.ac.id/index.php/pensa/index>
- Restu, N., Wulandari, I., Yamtinah, S., Retno, S., Ariani, D., & Ulfa, M. (2022). Analisis Indeks Aiken untuk Mengetahui Validitas Isi Instrumen Asesmen Kompetensi Minimum Berbasis Konteks Sains Kimia. 25(2), 184-191.
- Retnawati, H. (2016). *Analisis Kuantitatif Instrumen Penelitian* (1st ed.). Parama Publishing.
- Samaratungga, F. M., & Kusuma, D. (2023). Pengembangan Butir Soal Berbasis Open-Ended pada Materi Bangun Datar untuk Meningkatkan Kemampuan Berpikir Kritis Peserta Didik di Sekolah Dasar. *Journal of Engineering Research*, 08(2), 1433-1442.
- Slamet, R., & Wahyuningsih, S. (2014). Validitas dan reliabilitas terhadap instrumen kepuasan kerja. 17(2), 51-58. <https://doi.org/https://doi.org/10.46975/aliansi.17i2.428>
- Son, A. L. (2019). *Instrumentasi Kemampuan Pemecahan Masalah*

- Matematis Analisis Reliabilitas, Validitas, Tingkat Kesukaran dan Daya Beda Butir Soal. 10(1), 41–52.*
- Uloli, R. (2021). *Berpikir Kreatif dalam Penyelesaian Masalah Tantangan Pembelajaran Abad 21* (N. Ellina (ed.)). RFM PRAMEDIA JEMBER.
- Wahyuni, D., & Palupi, B. S. (2022). *Analisis Kemampuan Berpikir Kreatif Matematis Siswa Kelas V Sekolah Dasar Melalui Soal Open-Ended. 1(April), 76–83.*
- Widiyanto, J. (2018). *Evaluasi Pembelajaran* (A. Musandi (ed.)). UNIPMA PRESS.
- Yohana, A., & Rachmadiyah, P. (2023). Pengembangan Media Pembelajaran Animasi Berbasis Pendekatan Open_Ended Question Untuk Meningkatkan Keterampilan Berpikir Kritis Ips Kelas 5 Sekolah Dasar. *Jpgsd, 11(7), 1436–1446.*