

Validitas Instrumen Penilaian Berbasis Kelas untuk Mengukur Kompetensi Literasi Sains Peserta Didik pada Materi Pemanasan Global

Yeni Sherliyanti^{1#} dan Mukhayyarotin Niswati Rodliyatul Jauhariyah²

^{1,2}Jurusan Fisika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Surabaya

[#]Email: yenisherliyanti.20012@mhs.unesa.ac.id

Abstrak

Penelitian ini bertujuan untuk mendeskripsikan kualitas dari instrumen penilaian berbasis kelas yang dikembangkan. Kualitas ini berupa validitas teoritis dan validitas empiris instrumen penilaian berbasis kelas pada materi pemanasan global. Instrumen yang dikembangkan yaitu instrumen penilaian kelas pada aspek pengetahuan, keterampilan serta sikap dengan indikator literasi sains. Instrumen ini diharapkan dapat digunakan untuk mengukur kompetensi literasi sains pada peserta didik. Tahap penelitian ini mengacu pada model pengembangan ADDIE terbatas yaitu hanya pada tahap *Analysis*, *Design* dan *Development*. Teknik pengumpulan data pada penelitian ini yaitu menggunakan metode angket untuk mengetahui validitas teoritis dan metode tes untuk mengetahui validitas empiris. Validitas teoritis dilakukan oleh 3 validator ahli, sedangkan validitas empiris dilakukan melalui uji coba terbatas kepada mahasiswa. Instrumen penilaian yang dikembangkan yaitu berupa soal pilihan ganda yang berjumlah 30 butir soal. Hasil validitas teoritis instrumen pada aspek pengetahuan, keterampilan dan sikap dinyatakan layak digunakan dengan kriteria sangat valid. Validitas empiris yang meliputi validitas empiris butir soal, reliabilitas, tingkat kesukaran, dan daya pembeda didapatkan hasil yaitu 9 butir soal dari 30 soal dinyatakan tidak valid dan tidak digunakan. Sedangkan sisa soal yaitu 21 butir soal dinyatakan valid dan dapat digunakan. Butir soal tersebut nantinya dapat digunakan untuk mengukur profil literasi sains pada peserta didik.

Kata kunci: validitas teoritis, validitas empiris, literasi sains

Abstract

This research aims to describe the quality of the classroom-based assessment instruments developed. This quality encompasses the theoretical validity and empirical validity of the classroom-based assessments instruments on the subject of global warming. The instruments developed include classroom assessment tools for knowledge, skills, and attitudes with science literacy indicators. These instruments are expected to measure students' science literacy competencies. The research follows a limited ADDIE development model, focusing only on the Analysis, Design, and Development stages. Data collection techniques in this study include questionnaires to determine theoretical validity and test to determine empirical validity. Theoretical validity was assessed by three expert validators, while empirical validity was tested through limited trials with students. The developed assessment instrument consists of 30 multiple choice questions. The results showed that the theoretical validity of the instruments in the aspects of knowledge, skills, and attitudes was considered highly valid and suitable for use. For empirical validity, including item validity reliability, difficulty level, and discrimination power, it was found that 9 out of the 30 questions were invalid and thus not used. The remaining 21 questions were deemed valid and can be used. These questions can be used to measure students' science literacy profiles.

Keywords: theoretical validity, empirical validity, science literacy

PENDAHULUAN

Penilaian adalah rangkaian yang paling krusial pada keseluruhan proses pembelajaran. Penilaian dilakukan untuk mengetahui hasil belajar atau untuk mengukur pencapaian peserta didik (Marlina, 2021). Hasil penilaian akan mencerminkan mutu dari pengetahuan yang dimiliki atau keterampilan yang telah dicapai (Arifin, 2009). Salah satu jenis penilaian yaitu penilaian berbasis kelas.

Penilaian berbasis kelas mencatat pencapaian hasil belajar dan kompetensi dengan menyatakan secara jelas standar yang harus dicapai dan telah tercapai, serta dilengkapi dengan peta kemajuan belajar peserta didik (Asrul *et al.*, 2014). Penilaian berbasis kelas tidak hanya berfokus pada aspek pengetahuan saja tetapi juga melibatkan aspek sikap dan aspek keterampilan.

Namun pada kenyataan yang ditemukan di lapangan yaitu banyak pelaksanaan penilaian saat ini hanya fokus pada pengukuran aspek pengetahuan melalui tes tertulis (Mahadaniar et al., 2013). Padahal tidak semua tujuan pembelajaran dan metode pembelajaran akan efektif dinilai dengan cara ini. Sebagai contoh, proses pembelajaran sains tidak hanya mencakup aspek pengetahuan, tetapi juga aspek sikap dan keterampilan. Tetapi guru jarang memperdalam pengetahuan mengenai penilaian berbasis kelas karena alasan kesibukan dan juga karena kurangnya referensi yang tersedia di lingkungan sekolah (Arifin, 2010). Keterbatasan referensi ini menyulitkan guru dalam mengembangkan instrumen penilaian berbasis kelas.

Literasi sains merupakan kemampuan untuk mengaplikasikan keterampilan sains dalam kehidupan sehari-hari, tidak hanya mencakup pemahaman teori, tetapi juga kemampuan untuk menghadapi dan menyelesaikan masalah melalui pendekatan ilmiah serta dapat memberikan solusi yang tepat (Fery Haryadi et al., 2015). Menurut Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi (2021), literasi sains merupakan suatu komponen pendukung untuk peserta didik supaya mampu memiliki empat kompetensi hidup (4C) sebagai upaya untuk menghadapi perkembangan pendidikan dan teknologi di abad 21 ini. Kompetensi 4C yang dimaksud yaitu *critical thinking and problem solving skills* (berpikir kritis dan mampu menyelesaikan masalah), *creativity* (kreativitas), *collaboration* (kolaborasi), dan *communication* (komunikasi). Kompetensi ini dapat membantu siswa dalam menghadapi tantangan yang kompleks dan juga mendorong perkembangan karakter yang sesuai dengan Profil Pelajar Pancasila. Selain para ilmuwan, seluruh masyarakat juga diharapkan menggunakan kemampuan untuk berpikir ilmiah dan aktif terlibat dalam isu-isu perkembangan ilmu pengetahuan dan teknologi (Syamsiyah, 2016). Dengan demikian, penguasaan kompetensi literasi sains sangat krusial bagi peserta didik di semua negara.

Namun pada kenyataannya kompetensi literasi sains peserta didik di Indonesia masih tergolong dalam kategori rendah dan mengalami penurunan. Berdasarkan survei PISA skor literasi sains yang diraih Indonesia masih berada di bawah rata-rata skor internasional. Pada tahun 2015 Indonesia mencatat skor 403 dari skor rata-rata internasional sebesar 500. Lalu pada tahun 2018 Indonesia mencatat skor 396 dari rata-rata internasional 500, sedangkan pada tahun 2022, skor Indonesia turun menjadi 383 dari rata-rata internasional yang sama. Hal ini membuktikan bahwa kompetensi literasi sains di Indonesia mengalami penurunan dari tahun 2015 hingga 2022 (Sutrisna, 2021); (Kemdikbudristek, 2023).

Sedangkan untuk peringkat berdasarkan hasil survei PISA, pada tahun 2012 Indonesia berada di posisi 64 dari 65 negara partisipan. Pada tahun 2015 Indonesia berada di posisi 62 dari 70 negara. Lalu pada tahun 2018 Indonesia berada di posisi 70 dari 78 negara. Sedangkan pada tahun 2022 Indonesia menempati posisi 64 dari 78 negara. Hal tersebut membuktikan bahwa kompetensi literasi sains peserta didik di Indonesia tergolong rendah dibandingkan dengan negara-negara lain (Lutfiah & Admoko, 2023).

Rendahnya kompetensi literasi sains di Indonesia dipengaruhi oleh beberapa faktor. Salah satu faktor penyebabnya yaitu kurangnya latihan dan pengalaman peserta didik dalam mengerjakan soal yang berbasis literasi sains. Peserta didik masih jarang mengerjakan soal-soal yang terkait dengan fenomena ilmiah dan berkaitan dengan lingkungan sekitar. Selama ini soal-soal yang dikerjakan hanya sebatas mengaplikasikan teori dan rumus saja. Guru fisika di sekolah juga mengakui bahwa mereka masih menghadapi kesulitan dalam mengembangkan instrumen penilaian berbasis literasi sains. Para guru mengemukakan bahwa instrumen penilaian yang dibuat umumnya masih kurang dikaitkan dengan masalah kontekstual dalam kehidupan sehari-hari dan lingkungan sekitar. Hal tersebut dikarenakan masih kurangnya referensi terkait dengan instrumen penilaian literasi sains. Selain itu, para guru juga masih belum memiliki banyak pengalaman dalam membuat instrumen penilaian berbasis literasi sains.

Berdasarkan penjelasan tersebut, diperlukan pengembangan instrumen penilaian yang mampu secara efektif mengukur kompetensi literasi sains peserta didik. Sebagian besar penelitian masih mengembangkan instrumen penilaian untuk aspek pengetahuan saja. Penelitian yang dilakukan oleh Wardi (2023) menghasilkan instrumen penilaian berbasis literasi sains aspek pengetahuan pada materi inti atom dan radioaktivitas yang bersifat valid dan reliabel. Agustya (2023) juga melakukan penelitian yang menyusun instrumen penilaian literasi sains pada materi keterbatasan sumber energi yang bersifat valid secara teoritis dan empiris.

Penelitian yang dilakukan tersebut hanya sebatas mengembangkan instrumen penilaian pada aspek pengetahuan saja. Padahal kompetensi literasi sains dapat diukur selama proses pembelajaran. Untuk mengukur kompetensi literasi sains dalam pembelajaran, diperlukan instrumen penilaian berbasis kelas yang dapat menilai secara holistik, baik untuk aspek pengetahuan, keterampilan maupun sikap. Dengan melibatkan ketiga aspek penilaian tersebut diharapkan proses pembelajaran di dalam kelas terintegrasi dengan indikator literasi sains.

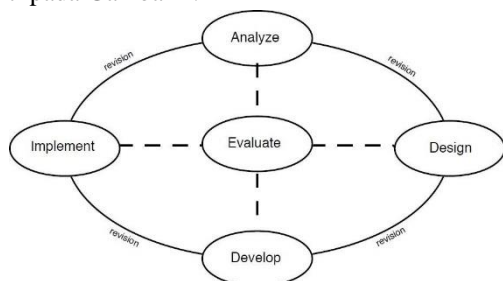
Saat ini dunia sedang gencar merencanakan program pengendalian pemanasan global. Para ahli yakin bahwa

perubahan iklim yang terjadi sejak 1950 disebabkan oleh aktivitas manusia. Pemanasan global dapat meningkatkan suhu bumi secara signifikan dan mengakibatkan kondisi sulit bagi tanaman untuk tumbuh. Hal tersebut berpotensi mengancam terjadinya krisis pangan yang serius bagi manusia. Jika tidak ditangani dengan cepat, masa depan kehidupan manusia dapat menjadi sangat sulit. (Imani *et al.*, 2016). Oleh karena itu, sangat penting untuk peserta didik memiliki pengetahuan terkait pemanasan global dan upaya mitigasi pemanasan global serta literat terhadap hal tersebut. Pemanasan global menjadi salah satu materi fisika yang cocok dengan literasi sains.

Instrumen penilaian dapat digunakan pada peserta didik untuk mengukur kompetensi literasi sains apabila telah dinyatakan valid. Instrumen dikatakan valid apabila telah diuji secara teoritis dan empiris. Sehingga penelitian ini dilaksanakan untuk mendeskripsikan validitas teoritis dan validitas empiris instrumen penilaian kelas berbasis literasi sains untuk mengukur kompetensi literasi sains pada materi pemanasan global. Penelitian yang dilakukan ini berbeda dengan penelitian terdahulu. Pada penelitian terdahulu instrumen penilaian yang dikembangkan hanya sebatas instrumen penilaian pada aspek pengetahuan saja. Sedangkan pada penelitian ini dikembangkan instrumen penilaian pada aspek pengetahuan, aspek keterampilan, serta aspek sikap. Instrumen yang dikembangkan disusun berdasarkan indikator kompetensi literasi sains.

METODE PENELITIAN

Jenis penelitian yang digunakan adalah penelitian pengembangan atau sering disebut *Research and Development* (R&D) oleh Sugiyono (2015). Penelitian pengembangan ini menggunakan model ADDIE. Model ADDIE terdiri dari lima tahapan yaitu Analisis (*Analyze*), Perencanaan (*Design*), Pengembangan (*Development*), Pelaksanaan (*Implementation*), dan Evaluasi (*Evaluation*) seperti pada Gambar 1.



Gambar 1. Desain Penelitian Model ADDIE (Branch, 2009)

Pada penelitian ini tahapan penelitian terbatas hanya sampai tahap development karena penelitian ini bertujuan untuk mendeskripsikan validitas teoritis dan validitas empiris instrumen penilaian yang dikembangkan. Teknik pengumpulan data yang digunakan pada penelitian ini yaitu metode angket dan tes. Metode angket dilakukan

melalui angket validasi yang diisi oleh para ahli yaitu dua dosen dari jurusan Fisika Unesa dan satu guru fisika SMA Negeri 15 Surabaya. Sedangkan metode tes, dilakukan melalui uji coba terbatas kepada 30 mahasiswa Pendidikan Fisika angkatan 2023 di Unesa.

Tahapan yang pertama yaitu *Analysis*. Tahap ini dilakukan dengan menggunakan analisis kebutuhan, analisis kurikulum, dan analisis materi. Tahapan kedua yaitu *Design* dilakukan dengan menyusun perangkat instrumen penilaian kelas yang terdiri dari kisi-kisi, naskah soal, serta rubrik penilaian. Tahapan yang terakhir yaitu *Development* yang dimana pada tahap ini akan dilakukan validasi kepada para ahli untuk mengetahui validitas teoritis. Selain itu, pada tahap ini juga dilakukan uji coba terbatas untuk mengetahui validitas empiris butir soal.

Validitas teoritis dilakukan dengan menggunakan lembar validasi yang diisi oleh validator sebagai ahli. Validitas ini berupa analisis terhadap aspek kelayakan isi, konstruk, dan bahasa. Instrumen penilaian yang divalidasi pada tahap ini yaitu instrumen penilaian aspek pengetahuan, keterampilan, dan sikap. Lembar validitas penilaian ini dianalisis berdasarkan skala Likert, seperti pada Tabel 1.

Tabel 1. Tabel Skala Likert

Skor	Kriteria
4	Sangat valid
3	Valid
2	Tidak valid
1	Sangat tidak valid

(Sugiyono, 2017)

Skor dari ketiga validator akan diambil nilai modulusnya dari setiap aspek. Penilaian validitas teoritis dilihat dari aspek isi, konstruk, dan bahasa. Kemudian, dihitung persentase kecocokan dari tiap validator menggunakan persamaan (1).

$$\text{Percentage of Agreement (PoA)} = 100\% \times \left(1 - \frac{A-B}{A+B}\right) \quad (1)$$

Setelah dinyatakan layak secara teoritis, selanjutnya instrumen penilaian tersebut akan diujicobakan secara terbatas kepada 30 mahasiswa Pendidikan Fisika Unesa. Instrumen penilaian yang diujicobakan hanya instrumen penilaian pada aspek pengetahuan saja. Data tersebut kemudian dianalisis menggunakan software AnatesV4 untuk mengetahui validitas empiris butir soal, reliabilitas, tingkat kesukaran dan daya beda.

Kriteria validitas empiris butir soal bergantung pada hasil koefisien korelasi yang langsung muncul pada software AnatesV4. Kriteria validitas item butir soal pada Tabel 2.

Tabel 2. Kriteria Validitas Butir Soal

Nilai	Kriteria
0,81 – 1,00	Sangat tinggi
0,61 – 0,80	Tinggi
0,41 – 0,60	Cukup
0,21 – 0,40	Rendah
0,00 – 0,20	Sangat rendah

Untuk reliabilitas jika diperoleh r hitung lebih besar daripada r tabel maka instrumen penilaian tersebut dapat dikatakan reliabel dan dapat digunakan. Untuk menghitung r hitung menggunakan persamaan (2).

$$r_{hitung} = \left(\frac{n}{n-1} \right) \left(1 - \frac{\sum St^2}{St^2} \right) \quad (2)$$

Tingkat kesukaran butir soal diukur menggunakan persamaan (3)

$$Tingkat\ kesukaran = \frac{mean}{skor\ maksimum\ untuk\ setiap\ butir\ soal} \quad (3)$$

Kriteria tingkat kesukaran soal didasarkan pada Tabel 3.

Tabel 3. Kriteria tingkat kesukaran soal

Nilai	Interpretasi
0,00 – 0,30	Sukar
0,31 – 0,70	Sedang
0,71 – 1,00	Mudah

Daya beda soal dapat diketahui menggunakan persamaan (4)

$$DP = \frac{XKa - XKb}{Skor\ maksimum} \quad (4)$$

Arikunto (2018) mengklarifikasi daya pembeda menjadi lima kriteria seperti pada Tabel 4.

Tabel 4. Klasifikasi daya pembeda

Daya beda	Kriteria
$\leq 0,00$	Sangat jelek
0,00 – 0,20	Jelek
0,21 – 0,40	Cukup
0,41 – 0,70	Baik
0,71 – 1,00	Sangat baik

HASIL DAN PEMBAHASAN

Tahap Analysis (Analisis)

Tahap analisis ini dilakukan dengan menggunakan analisis kebutuhan, analisis kurikulum, dan analisis materi. Tahapan ini dilakukan dengan studi pustaka dan wawancara. Analisis kebutuhan ini dilakukan untuk mengetahui pentingnya kompetensi literasi sains bagi peserta didik melalui studi pustaka. Selain studi pustaka, juga dilakukan wawancara terhadap guru fisika dan beberapa peserta didik di SMA Negeri 15 Surabaya.

Berdasarkan hasil studi pustaka diperoleh bahwa kompetensi literasi sains adalah suatu kompetensi penting yang harus dimiliki oleh seorang peserta didik. Literasi sains berperan sebagai salah satu pendukung peserta didik untuk memiliki kompetensi hidup 4C sebagai bekal untuk menghadapi perkembangan pendidikan dan teknologi di abad 21 ini. Namun pada kenyataannya kompetensi literasi sains yang dimiliki oleh peserta didik di Indonesia masih berada pada tingkat rendah. Hal tersebut sesuai dengan survei yang telah dilaksanakan oleh PISA pada Gambar 2 bahwa peringkat Indonesia pada kompetensi literasi sains berada di peringkat bawah. Selain itu, nilai rata-rata Indonesia berada jauh dibawah dari rata-rata internasional.



Gambar 2. Perubahan Skor Literasi Sains
(Kemdikbudristek, 2023)

Menurut hasil wawancara dengan guru fisika di SMA Negeri 15 Surabaya diketahui bahwa masih kurangnya instrumen penilaian kelas berbasis literasi sains. Penilaian yang dibuat masih kurang melibatkan masalah kontekstual dalam kehidupan sehari-hari dan lingkungan sekitar. Selain itu, wawancara juga diberikan kepada peserta didik di SMA Negeri 15 Surabaya. Mereka mengaku bahwa mereka kurang terlatih dan kurang terbiasa dalam mengerjakan soal berbasis literasi sains. Selama ini soal-soal yang dikerjakan hanya sebatas mengaplikasikan teori dan rumus saja. Hasil analisis kebutuhan ini dapat membuktikan pentingnya pengembangan instrumen penilaian kelas berbasis literasi sains. Hasil penelitian terdahulu juga menunjukkan bahwa instrumen penilaian saat ini hanya mencakup pemahaman dan penerapan konsep saja, dengan soal-soal yang tidak terkait langsung dengan fenomena dalam kehidupan sehari-hari (Anisa, 2017).

Analisis kurikulum ini dilaksanakan untuk menentukan jenis kurikulum yang digunakan dalam proses pembelajaran di sekolah. Berdasarkan hasil wawancara oleh guru fisika di SMA Negeri 15 Surabaya, diketahui bahwa di sekolah tersebut menggunakan kurikulum merdeka untuk kelas X dan XI serta kurikulum 2013 untuk kelas XII. Pada kurikulum merdeka terdapat capaian pembelajaran yang harus dicapai oleh peserta didik pada setiap fasenya. Proses pembelajaran kelas X di SMA Negeri 15 mengacu pada capaian pembelajaran fisika di Fase E.

Analisis materi dilaksanakan untuk menentukan materi apa yang akan dijadikan dasar pada instrumen penilaian yang disusun. Pada tahap ini dilakukan analisis terhadap materi pemanasan global agar sesuai dengan capaian

proses pembelajaran yang akan dilakukan. Tujuannya agar dalam setiap butir soal yang dikembangkan sesuai dengan konteks materi yang dipelajari peserta didik. Materi yang akan digunakan dalam instrumen penilaian ini adalah pemanasan global. Secara garis besar, materi pemanasan global membahas mengenai fenomena pemanasan global, efek rumah kaca, penyebab dari pemanasan global, dampak yang ditimbulkan, serta upaya mitigasi untuk mengatasinya.

Tahap Design (Perencanaan)

Pada tahap perencanaan ini dilakukan dengan menyusun perangkat penelitian instrumen penilaian kelas berbasis literasi sains. Instrumen penilaian yang disusun mengacu pada kompetensi literasi sains menurut PISA 2018. Instrumen yang dirancang yaitu sebagai berikut a) Kisi-kisi instrumen penilaian aspek pengetahuan literasi sains materi pemanasan global, b) Kisi-kisi instrumen penilaian aspek keterampilan, c) Kisi-kisi instrumen penilaian aspek sikap melalui observasi, d) Kisi-kisi instrumen penilaian aspek sikap melalui angket, e) Naskah soal instrumen penilaian aspek pengetahuan literasi sains materi pemanasan global, f) Rubrik penilaian instrumen penilaian aspek keterampilan, g) Rubrik penilaian kisi instrumen penilaian aspek sikap melalui observasi, h) Instrumen penilaian aspek sikap melalui angket.

Tahap Development (Pengembangan)

Instrumen penilaian yang dikembangkan akan melalui tahap validitas teoritis oleh para ahli. Validitas teoritis dilaksan oleh para ahli untuk menilai instrumen penilaian yang dibuat dari ranah isi, konstruk, dan bahasa. Hasil validasi dari tiga ahli tersebut akan dianalisis untuk memperoleh *Percentage of Agreement* atau persentase kecocokan antar individu. Instrumen penilaian yang divalidasi yaitu instrumen penilaian aspek pengetahuan, keterampilan dan sikap. Hasil validasi disajikan sebagai berikut.

Tabel 5. Hasil validasi instrumen penilaian aspek pengetahuan

Ranah	PoA	Kriteria
Ranah isi	92%	Sangat valid
Ranah konstruk	95%	Sangat valid
Ranah bahasa	93%	Sangat valid

Tabel 6. Hasil validasi instrumen penilaian aspek keterampilan

Ranah	PoA	Kriteria
Ranah isi	86%	Valid
Ranah konstruk	96%	Sangat valid
Ranah bahasa	91%	Sangat valid

Tabel 7. Hasil validasi instrumen penilaian aspek sikap

Ranah	PoA	Kriteria
Ranah isi	86%	Valid
Ranah konstruk	93%	Sangat valid
Ranah bahasa	96%	Sangat valid

Berdasarkan Tabel 7 dapat diketahui bahwa rata-rata nilai PoA pada aspek isi yaitu sebesar 92% dengan kategori sangat valid. Hal tersebut menunjukkan bahwa butir soal pada instrumen penilaian aspek pengetahuan yang dibuat telah sesuai dengan capaian pembelajaran dan tujuan pembelajaran, serta juga telah sesuai dengan konten dari kurikulum yaitu Kurikulum Merdeka. Selain itu, butir soal dinyatakan sangat valid juga karena isi dari butir soal sama dengan materi yang telah dipelajari dan indikator literasi sains menurut PISA tahun 2018. Validitas aspek isi ini menjadi aspek penting instrumen yang berkualitas karena butir soal yang dihasilkan harus mampu mencerminkan konsep yang akan diukur (Maulita *et al.*, 2019).

Nilai rata-rata PoA pada aspek konstruk yaitu sebesar 95% dengan kriteria sangat valid. Hal tersebut membuktikan bahwa butir soal disusun dengan tegas dan jelas, tidak memberikan petunjuk ke arah jawaban yang benar, terdapat petunjuk pengerjaan serta pedoman skor yang jelas. Selain itu, dalam aspek ini diketahui bahwa komponen pendukung soal seperti gambar, grafik, tabel, diagram jelas dan berfungsi dengan baik. Validitas aspek konstruk ini penting karena mencerminkan kualitas suatu instrumen apakah dapat mengungkapkan kemampuan berpikir yang dianalisis dengan akurat (Insani *et al.*, 2017).

Nilai PoA pada aspek bahasa yaitu sebesar 93% dengan kategori sangat valid. Hal itu berarti butir soal disusun dengan menggunakan bahasa Indonesia yang baik dan benar, dapat dipahami dan bersifat komutatif. Serta ragam kalimat dalam butir soal yang dibuat tidak menimbulkan penafsiran ganda.

Untuk menilai instrumen penilaian aspek keterampilan. Rata-rata nilai PoA pada validitas aspek isi yaitu sebesar 86% dengan kriteria valid. Hal tersebut menunjukkan bahwa instrumen penilaian aspek keterampilan yang disusun telah mencakup tahapan dan indikator penilaian untuk keterampilan literasi sains. Serta setiap item yang diamati telah sesuai dengan aspek keterampilan literasi sains.

Rata-rata PoA pada aspek konstruk yaitu sebesar 96% dengan kategori sangat valid. Hal tersebut menunjukkan bahwa instrumen penilaian aspek keterampilan yang dibuat memiliki petunjuk pengisian, kriteria yang diamati, dan rubrik penskoran yang disusun yang dinyatakan dengan jelas. Rubrik penilaian juga telah dapat mengukur keterampilan literasi sains. Sedangkan pada aspek bahasa diketahui rata-rata PoA yaitu 91% yang termasuk sangat valid. Hal tersebut berarti bahasa yang

digunakan dalam soal adalah bahasa Indonesia yang baik dan benar serta bersifat komutatif. Serta instrumen penilaiannya memiliki struktur kalimat yang sederhana.

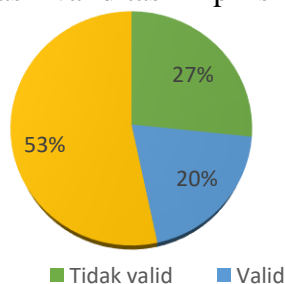
Instrumen penilaian aspek sikap juga dianalisis menggunakan validitas teoritis ini. Rata-rata nilai PoA pada validitas aspek isi yaitu sebesar 86% dengan kriteria valid. Hal tersebut menunjukkan bahwa instrumen penilaian aspek sikap yang disusun telah mencakup tahapan dan indikator penilaian untuk keterampilan literasi sains. Serta setiap item yang diamati telah sesuai dengan aspek sikap literasi sains.

Rata-rata PoA pada aspek konstruk yaitu sebesar 93% yang termasuk sangat valid. Hal ini membuktikan instrumen penilaian aspek sikap yang dibuat memiliki petunjuk pengisian, kriteria yang diamati, dan rubrik penskoran yang disusun yang dinyatakan dengan jelas. Rubrik penilaian juga telah dapat mengukur sikap literasi sains peserta didik. Sedangkan untuk aspek bahasa diketahui nilai rata-rata PoA sebesar 96% dengan kriteria sangat valid. Hal tersebut berarti bahwa bahasa yang digunakan adalah bahasa Indonesia yang baik dan benar serta bersifat komutatif. Serta instrumen penilaiannya memiliki struktur kalimat yang sederhana.

Setelah dilakukan validitas teoritis dan melalui tahap revisi, instrumen penilaian kelas berbasis literasi sains akan melalui tahap validitas empiris. analisis validitas empiris ini didapatkan dari uji coba terbatas kepada 30 mahasiswa Pendidikan Fisika di Universitas Negeri Surabaya. Skor dari hasil pengerjaan mahasiswa kemudian dianalisis untuk mengetahui validitas empiris butir soal, reliabilitas soal, tingkat kesukaran, dan daya beda soal.

Hasil validitas empiris butir soal digunakan untuk menentukan suatu soal valid atau tidak valid. Hasil analisis validitas empiris butir soal ditunjukkan pada Gambar 3.

Hasil Validitas Empiris



Gambar 3. Hasil Validitas Empiris

Pada Gambar 3 diketahui bahwa soal yang termasuk tidak valid yaitu 27% atau 9 butir soal. Soal tersebut yaitu nomor 1, 6, 8, 12, 14, 19, 20, 23, 24. Sehingga 21 butir soal yang lain masuk dalam kriteria valid dan sangat valid. Butir soal dengan kategori valid dapat digunakan dan diimplementasikan kepada peserta didik. Sedangkan soal dengan kategori tidak valid tidak dapat digunakan karena

tidak memiliki kemampuan untuk mengukur kompetensi yang akan diukur (Arikunto, 2018).

Terdapat beberapa soal yang dikembangkan masuk ke dalam kategori tidak valid. Hal tersebut dapat disebabkan karena beberapa alasan. Butir soal yang dikembangkan dianggap sulit oleh peserta didik karena mereka masih belum mengetahui materi tersebut atau merasa kesulitan saat menjawab soal tersebut. Dan sebaliknya, butir soal dapat tidak valid karena terlalu mudah dan kurang memiliki kemampuan untuk membedakan antara peserta yang memiliki tingkat kemampuan yang berbeda. Seperti yang dikatakan oleh Helendra *et al.*, (2021) yaitu faktor yang mempengaruhi validitas suatu hasil tes yaitu karena soal yang dianggap mudah sehingga dapat dijawab oleh peserta didik sehingga korelasinya rendah atau juga dikarenakan suasana kelas yang kurang kondusif sehingga peserta didik menjadi kurang konsentrasi dalam mengerjakan soal.

Reliabilitas mengacu pada tingkat kestabilan atau konsistensi hasil pengukuran dari hal yang sama. Reliabilitas juga dapat dikatakan seberapa stabil nilai yang didapatkan oleh peserta yang sama ketika diuji menggunakan tes yang identik dalam waktu yang berbeda (Depdikbud, 1997). Instrumen yang dikatakan baik akan memiliki reliabilitas yang baik pula. Hasil reliabilitas dapat ditunjukkan pada Tabel 8.

Tabel 8. Hasil Uji Reliabilitas

r_{tabel}	r_{hitung}	Kategori
0,349	0,890	Reliabel

Nilai reliabilitas diperoleh dari perhitungan *Alpha Cronbach*. Hasil perhitungan diperoleh r_{hitung} sebesar 0,890. Nilai tersebut lebih besar daripada r_{tabel} yaitu sebesar 0,349. Sehingga instrumen penilaian yang dikembangkan bersifat reliabel dan dapat digunakan. Artinya jika instrumen tersebut diujikan pada peserta didik maka akan mendapatkan hasil yang relatif sama. Hal tersebut juga dikemukakan oleh Helendra & Sari. D. R., (2021) yaitu bahwa sebuah instrumen memenuhi kriteria reliabel ketika instrumen tersebut memberikan hasil yang tetap walaupun diujikan pada waktu yang berlainan.

Tingkat kesukaran butir soal digunakan untuk mengukur kualitas suatu butir soal dan dapat diklasifikasikan ke dalam kategori sukar, sedang atau mudah. Penentuan mudah atau sukar nya butir soal ditentukan oleh besarnya persentase jawaban benar peserta didik pada soal tersebut (Arikunto, 2018). Hasil analisis jawaban uji coba menunjukkan bahwa tingkat kesukaran butir soal yang dikembangkan terdiri atas mudah, sedang dan sukar. Gambar 4.2 merupakan hasil analisis tingkat

kesukaran butir soal dari instrumen penilaian yang telah dikembangkan. Dapat diketahui bahwa 2 soal (7%) berkategori sukar, 6 soal (20%) berkategori sedang, 14 soal (46%) berkategori mudah dan 8 soal (27%) berkategori sangat mudah.

Daya beda merupakan kemampuan butir soal untuk mengklasifikasikan antara peserta didik dengan kemampuan tinggi dan kemampuan rendah. Peserta didik dengan kemampuan tinggi disebut kelompok atas, sementara dengan kemampuan rendah disebut kelompok bawah. Semakin besar nilai koefisien daya beda, maka butir soal tersebut lebih baik dalam membedakan kemampuan peserta didik (Arifin, 2012). Pada penelitian ini instrumen penilaian aspek pengetahuan diujicobakan

pada 30 mahasiswa sehingga jumlah kelompok atas dan bawah masing-masing adalah 15 orang. Gambar 5 merupakan hasil analisis daya beda butir soal Terdapat 4 soal (13%) termasuk dalam kategori jelek, 6 soal (20%) dalam kategori cukup, 18 soal (60%) dalam kategori baik dan 2 soal (7%) dalam kategori sangat baik.

Berdasarkan validitas teoritis, instrumen penilaian apabila memenuhi kriteria cukup dengan persentase $\geq 41\%$ maka dapat dinyatakan valid. Sedangkan setiap butir soal dinyatakan valid secara empiris jika memenuhi kriteria cukup pada setiap aspeknya. Butir soal yang dinyatakan valid dan dapat digunakan secara keseluruhan disajikan pada Tabel 9.

Tabel 9. Butir Soal yang dinyatakan valid berdasarkan validitas teoritis dan empiris

No Soal	VT	VE				Keterangan
		VEBS	R	TK	DP	
1	93%	0,343	0,890	Sedang	0,50	Tidak digunakan
2		0,408		Sukar	0,75	Digunakan
3		0,386		mudah	0,62	Digunakan
4		0,604		Sedang	0,75	Digunakan
5		0,652		Mudah	0,50	Digunakan
6		-0,018		Sangat mudah	0,00	Tidak digunakan
7		0,650		Sangat mudah	0,38	Digunakan
8		0,034		Sukar	0,25	Tidak digunakan
9		0,615		Sangat mudah	0,38	Digunakan
10		0,519		Mudah	0,38	Digunakan
11		0,490		Sangat mudah	0,25	Digunakan
12		-		Sangat mudah	0,00	Tidak digunakan
13		0,646		Mudah	0,50	Digunakan
14		0,277		Sangat mudah	0,25	Tidak digunakan
15		0,660		Sedang	0,50	Digunakan
16		0,737		Sangat mudah	0,50	Digunakan
17		0,761		Mudah	0,62	Digunakan
18		0,535		Mudah	0,50	Digunakan
19		0,312		Mudah	0,50	Tidak digunakan
20		0,319		Mudah	0,50	Tidak digunakan
21		0,455		Sedang	0,62	Digunakan
22		0,437		Sedang	0,38	Digunakan
23		0,406		Sangat mudah	0,12	Tidak digunakan
24		0,248		Sedang	-0,12	Tidak digunakan
25		0,593		Mudah	0,62	Digunakan
26		0,583		Mudah	0,50	Digunakan
27		0,652		Mudah	0,62	Digunakan
28		0,607		Mudah	0,50	Digunakan
29		0,637		Mudah	0,50	Digunakan
30		0,614		Mudah	0,50	Digunakan

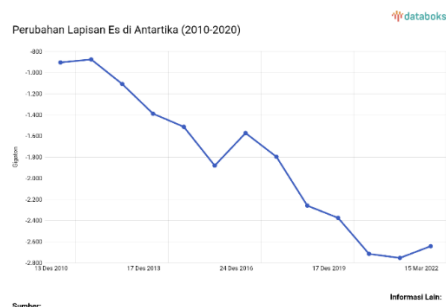
Berikut merupakan salah satu soal dari instrumen penilaian pada aspek pengetahuan yang memenuhi kriteria berdasarkan validitas teoritis dan validitas empiris, sehingga soal dapat diimplementasikan kepada peserta didik.



Pelelehan es pada gunung-gunung api yang berada di negara empat musim atau di wilayah kutub. Dr. Mirzam mencontohkan Gunung Eyjafjallajökull yang ada di Islandia. Pada 2010 lapisan es yang terdapat di gunung tersebut meleleh karena pemanasan global, dan perubahan dari musim dingin ke musim semi. Gunung Eyjafjallajökull sendiri setiap tahun esnya mencair sebanyak 11 miliar ton. “Karena es itu meleleh maka bisa dibayangkan gunung api yang tadinya tertutup es sebagai penahan tudungnya, esnya hilang tiba-tiba. Beban yang hilang tersebut membuat kekurangan tekanan yang dapat menyebabkan magma di dalam gunung tersebut sangat mudah naik ke atas sehingga gunung api kemudian meletus,” ungkapnya. Berdasarkan cuplikan berita tersebut, pemanasan global dapat mengakibatkan bencana alam lainnya, misalnya saja gunung meletus. Menurut pendapat kalian bagaimana kaitannya proses terjadinya gunung meletus dengan pemanasan global?

- Global warming membuat suhu dan volume magma yang berada di dalam gunung api meningkat sehingga gunung api meletus.
- Proses pelelehan es yang menyelimuti gunung api diakibatkan oleh global warming sehingga es yang mencair akan bercampur dengan magma
- Global warming membuat suhu rata-rata bumi tidak stabil sehingga tekanan dalam magma gunung api juga tidak stabil sehingga mengakibatkan gunung api meletus.
- Gunung api yang semula tidak aktif, dikarenakan tertutupi oleh lapisan es menjadi aktif ketika lapisan es meleleh dikarenakan peningkatan suhu rata-rata bumi
- Gunung api yang tidak aktif akan menjadi gunung api yang aktif dikarenakan lelehan es akan mengendap bersamaan dengan magma di perut gunung api.

Jawaban : A



Salah satu dampak nyata perubahan iklim adalah naiknya permukaan air laut. Kenaikan permukaan air laut itu terjadi karena lapisan es di kutub bumi mencair akibat pemanasan global. Saat kawasan es di bagian kutub mencair akibat perubahan iklim, hal tersebut dapat menimbulkan masalah bagi hewan yang bergantung pada es laut untuk berkembang biak, berburu dan mencari makan. Berdasarkan uraian tersebut, apakah dampak lebih lanjut apabila lapisan es di antartika semakin menurun?

- Penurunan suhu global secara signifikan karena es yang mencair
- Peningkatan suhu global karena es yang mencair dapat melepaskan gas-gas rumah kaca yang terperangkap di dalamnya
- Ekosistem di laut semakin berkembang karena habitatnya yang akan semakin luas
- Pola cuaca global akan semakin stabil
- Peningkatan suhu di antartika karena es yang mencair sehingga dapat dihuni oleh manusia

Jawaban : B

SIMPULAN

Berdasarkan hasil penelitian pengembangan Instrumen penilaian kelas berbasis literasi sains dinyatakan layak berdasarkan validitas teoritis. Berdasarkan validitas empiris didapatkan 9 soal dari 30 soal yang dikembangkan tidak valid dan tidak dapat digunakan. Sebanyak 21 butir soal dinyatakan valid dan dapat digunakan untuk mengukur kompetensi literasi sains peserta didik. Instrumen penilaian kelas berbasis literasi sains pada penelitian ini hanya sebatas pada satu materi yaitu pemanasan global. Oleh karena itu, diharapkan untuk penelitian selanjutnya dapat mengembangkan instrumen penilaian kelas pada materi yang lain. Butir soal yang dikembangkan pada penelitian ini tidak seimbang jumlahnya pada setiap indikator literasi sains. Diharapkan penelitian selanjutnya memperhatikan jumlah butir soal agar sama dan seimbang pada setiap indikator literasi sains. Dalam membuat instrumen penilaian berbasis literasi sains jangan memuat terlalu banyak bacaan. Fenomena sains yang ingin dimuat bisa dalam bentuk gambar, grafik, maupun tabel.

DAFTAR PUSTAKA

- Agustya, E.N., Jauhariyah. M. N. R., (2023) Analysis of the Science Literacy Competency Profile of High School on Limited Energy Sources. *Kappa Journal*, 7(3), 363-367
- Arifin, Z. (2009). *Evaluasi Pembelajaran*. Remaja Rosdakarya.
- Arifin. (2010). *Metode Penelitian Kualitatif, Kuantitatif, dan R&D*. Bandung. Alfabeta
- Arifin, Z. (2012). *Evaluasi Pembelajaran*. Jakarta: Direktorat Jenderal Pendidikan Islam Kementrian Agama
- Arikunto, S. (2015). *Dasar-Dasar Evaluasi Pendidikan*. Bumi Aksara.
- Arikunto, S. (2018). *Dasar-Dasar Evaluasi Pendidikan, Edisi 3*. Bumi Aksara.
- Asrul, Ananda, R., & Rosnita. (2014). *Evaluasi Pembelajaran*. Citapustaka Media.
- Branch, R. M. (2009). *Instructional Design-The ADDIE Approach*. New York: Springer
- Fery Haryadi, E. S., Priyono, A. B., & Retnoningsih, A. (2015). Desain Pembelajaran Literasi Sains Berbasis Problem Based Learning Dalam Membentuk Keterampilan Berpikir Kritis Siswa Info Artikel. In *JISE* (Vol. 4, Issue 2). [Http://Journal.Unnes.Ac.Id/Sju/Index.Php/Jise](http://Journal.Unnes.Ac.Id/Sju/Index.Php/Jise)
- Helendra, & Sari. D. R. (2021). Pengembangan Instrumen Asesmen Berbasis Literasi Sains Tentang Materi Sistem Ekskresi Dan Sistem Pernapasan. *Jurnal Ilmiah Profesi Guru*, 4(1), 17–25.
- Imani, H. A., Sati, I. M., & Purwanto. (2016). Profil Literasi Sains Siswa SMP Di Kota Bandung Terkait Tema Pemanasan Global. *Seminar Nasional Pendidikan IPA Pascasarana Universitas Negeri Malang*, 242248.
- Insani, N., F, & Sunarti, T. (N.D.). Keterlaksanaan Model Pembelajaran Sains Teknologi Masyarakat Untuk Meningkatkan Literasi Sains Dalam Pembelajaran Fisika. *Inovasi Pendidikan Fisika*, 7(2), 149153.
- Karira, F. N., Sunarti, T., Niswati, M., Setyarsih, W. (2023). Validitas Instrumen Tes Berbasis Literasi Sains Untuk Mengukur Keterampilan berpikir Kritis Siswa SMA pada Materi Energi Terbarukan. *Inovasi Pendidikan Fisika*. 12(2). 118-125
- Kemdikbudristek. (2021). *Pengembangan Instrumen Penilaian Berbasis Literasi Sains*.
- Kemdikbudristek. (2023). *PISA 2020 Dan Pemulihan Pembelajaran Di Indonesia*.
- Lestari, D., Setyarsih, W. (2020). Kelayakan Instrumen Penilaian Formatif Berbasis Literasi Sains Peserta Didik Pada Materi Pemanasan Global. *IPF : Inovasi Pendidikan Fisika*, 09(03), 561-570.
- Lutfiah, K., & Admoko, S. (2023). Validitas Lembar Kerja Berbasis Masalah Untuk Meningkatkan Kemampuan Literasi Sains Pada Materi Pemanasan Global. *Inovasi Pendidikan Fisika*, 12(2).
- Mahadaniar, A., Siswaningsih, W., & Nahadi. (2013). Pengembangan Instrumen Penilaian Berbasis Kelas Pada Pembelajaran Hidrokarbon. *Jurnal Pengajaran MIPA*, 18(2), 245–252.
- Mardhiyyah, L. A., Rusilowati, A., & Linuwih, S. (2016). Pengembangan Instrumen Asesmen Literasi Sains Tema Energi. *Journal of Primary Education*.
- OECD. (2019). *PISA 2018 Assessment And Analytical Framework, PISA*.
- Ratna Sari, D. (2021). Pengembangan Instrumen Asesmen Berbasis Literasi Sains Tentang Materi Sistem Eksresi Dan Sistem Pernapasan. 4, 17-25. <https://doi.org/10.23887/Jippg.V4i1>
- Sugiyono. (2015). *Metode dan Penelitian Dan Pengembangan*. Bandung: Alfabeta
- Sugiyono. (2017). *Statistik Untuk Penelitian*. Alfabeta.
- Sutrisna, N. (2021). Analisis Kemampuan Literasi Sains Peserta Didik SMA Di Kota Sungai Penuh. *Jurnal Inovasi Penelitian*, 1(12).
- Syamsiyah, S. (2016). Ualitis Instrumen Penilaian Literasi Sains Siswa Kelas VII Pada Materi Interaksi Antar Makhluk Hidup. *Pensa: Jurnal Pendidikan Sains*, 4(03).
- Wardi, L. Z. (2023). Analisis Profil Kompetensi Literasi Sains Siswa SMA Pada Materi Inti Atom dan Radioaktivitas. *Inovasi Pendidikan Fisika*, 12(2), 74-80)