

Analisis Kualitas Instrumen Penilaian Berbasis Kelas untuk Mengukur Kompetensi Literasi Sains Peserta Didik SMA pada Materi Energi Terbarukan

Nadia Dewi Masithoh¹ dan Mukhayyarotin Niswati Rodliyatul Jauhariyah²

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

nadia.20075@mhs.unesa.ac.id mukhayyarotinjauhariyah@unesa.ac.id

ABSTRACT

The aim of this research is to analyze the quality of class-based instruments in the domains of knowledge, attitudes and skills to measure scientific literacy competency. This research uses a limited ADDIE research design consisting of Analyze, Design and Development stages. Data was collected using test and questionnaire methods. This research was tested on 30 Unesa Physics Education Study Program students. Theoretical validity results show that the assessment instrument prepared meets the valid category. The theoretical validity results were obtained from validation by 3 expert validators, namely 2 Unesa Physics Education lecturers and 1 MBI Amanatul Ummah Pacet Mojokerto teacher. Empirical validity was obtained from testing the assessment instrument on 30 active Unesa Physics Education students with a total of 15 questions. Then, the question item was declared valid with an account of 0.922 and an rtable of 0.361.

Keywords: *Theoretical Validity, Empirical Validity, Scientific Literacy*

ABSTRAK

Tujuan penelitian ini yaitu menganalisis kualitas instrumen berbasis kelas dalam ranah pengetahuan, sikap dan keterampilan untuk mengukur kompetensi literasi sains. Penelitian ini menggunakan desain penelitian ADDIE terbatas yang terdiri dari tahap *Analyze, Desain dan Development*. Data dikumpulkan menggunakan metode tes dan angket. Penelitian ini diuji cobakan kepada 30 mahasiswa Prodi Pendidikan Fisika Unesa. Hasil validitas teoritis menunjukkan bahwa instrumen penilaian yang disusun memenuhi kategori valid, hasil validitas teoritis ini diperoleh dari validasi oleh validator 3 ahli yaitu 2 dosen Pendidikan Fisika Unesa dan 1 Guru MBI Amanatul Ummah Pacet Mojokerto. Validitas empiris diperoleh dari uji coba instrumen penilaian pada 30 Mahasiswa aktif Pendidikan Fisika Unesa dengan jumlah butir soal yakni 15 butir soal. Kemudian, butir soal tersebut dinyatakan valid dengan r_{hitung} sebesar 0,922 dan r_{tabel} sebesar 0,361.

Kata Kunci: Validitas Teoritis, Validitas Empiris, Literasi Sains

PENDAHULUAN

Penilaian berbasis kelas adalah penilaian dalam arti “*assessment*”. Maksudnya, data dan informasi dari penilaian berbasis kelas merupakan salah satu bukti yang dapat digunakan untuk mengukur keberhasilan suatu tujuan pendidikan (Arifin, 2020). Tujuan pendidikan yang dimaksud adalah standar kompetensi, kompetensi dasar, dan indikator pencapaian hasil belajar yang terdapat pada kurikulum. Penilaian berbasis kelas dilakukan untuk memberikan keseimbangan pada tiga

domain, yaitu kognitif, afektif dan psikomotor dengan menggunakan berbagai bentuk dan model penilaian yang dilakukan secara sistematis dan sistemik, menyeluruh dan berkelanjutan. Menurut sejarah pengembangan dan penggunaan sistem penilaian di dunia pendidikan, dapat disimpulkan bahwa penilaian merupakan upaya sistematis yang dikembangkan oleh suatu institusi pendidikan yang ditujukan untuk menjamin tercapainya kualitas proses pendidikan serta kualitas kemampuan peserta didik sesuai dengan tujuan yang telah ditetapkan. Penilaian juga merupakan proses pembelajaran antara lain sebagai kegiatan menghimpun fakta-fakta dan dokumen belajar peserta didik yang dapat dipercaya untuk melakukan perbaikan program, apabila kegiatan penilaian tersebut terjadi sebagai bagian dari program kelas. Oleh karena itu penilaian berfungsi membantu guru untuk merencanakan kurikulum dan program pembelajaran, maka kegiatan penilaian membutuhkan informasi bervariasi dari setiap individu dan atau kelompok peserta didik serta guru. Guru dapat melakukan penilaian dengan cara mengumpulkan catatan yang diperoleh melalui pertemuan, observasi, portofolio, proyek, produk, ujian, serta data hasil *interview* dan *survey*. Penilaian kinerja merupakan berbagai macam tugas dan situasi di mana peserta tes diminta untuk mendemonstrasikan pemahaman dan mengaplikasikan pengetahuan yang mendalam, serta keterampilan di dalam berbagai macam konteks (Pamungkas *et al.*, 2024). Penggunaan *classroom assessment* pada prinsipnya harus berperan dalam upaya meningkatkan motivasi dan membangkitkan kembali motivasi peserta didik dalam mengikuti proses pembelajaran dan bukan sebaliknya.

Instrumen adalah alat yang dapat digunakan untuk membantu setiap orang jika ingin mengumpulkan sebuah data. Instrumen juga dapat digunakan untuk mengumpulkan data atau informasi jika ingin mengembangkan kemampuan belajar peserta didik pada suatu materi. Penilaian adalah suatu kegiatan yang dilakukan jika seseorang ingin mengumpulkan informasi dan mendapatkan sebuah hasil dari penilaian tersebut. Penilaian terkadang dilakukan karena seseorang ingin mengetahui efektivitas ataupun kualitas dari suatu produk. Penilaian sudah tidak asing dalam kehidupan sehari-hari termasuk dalam suatu pembelajaran. Menurut Darmayanti & Setiawati (2022) *asesmen* atau penilaian dapat dihubungkan dengan kemampuan seseorang, seperti kecerdasan, keterampilan, kecepatan, dan juga ketepatan dalam menyelesaikan pekerjaan atau tugas tertentu. Keputusan penilaian dalam hasil belajar sangat bermanfaat untuk merefleksikan pengetahuan, cara belajar, dan mendorong tanggung jawab dalam belajar peserta didik. Hal ini karena kegiatan penilaian diharuskan dapat memberikan informasi kepada guru untuk meningkatkan kemampuan mengajar dan membantu peserta didik untuk mencapai perkembangan belajar secara optimal.

Literasi sains merupakan salah satu keterampilan abad 21 yang penting untuk dikuasai (Fitriana & Supahar, 2019). Literasi sains merupakan kemampuan esensial yang diperlukan untuk mengorganisasi, menganalisis serta menginterpretasikan informasi. (Gultepe & Kilic, 2015) mengungkapkan literasi sains sebagai pengetahuan ilmiah yang dijadikan dasar dalam kehidupan. (Irwan *et al.*, 2019) menyatakan literasi sains berkaitan dengan keterampilan bertindak dan berpikir saintifik untuk menyikapi isu sains serta memahami lingkungan, sosial modern, dan teknologi.

Literasi sains menumbuhkembangkan sikap kepedulian terhadap diri dan lingkungan dalam menghadapi dan menyelesaikan permasalahan di kehidupan sehari-hari. Penilaian atau *assessment* yang mengukur literasi sains siswa perlu dilakukan dalam jangka waktu tertentu. PISA (*Programme for International Student Assessment*) dan TIMSS (*Trends in International Mathematics and Science Study*) adalah contoh program penilaian yang bertujuan untuk mengukur literasi sains pada siswa secara internasional. PISA melakukan survei yang mengukur literasi sains setiap tiga tahun sekali, sedangkan TIMSS melakukannya setiap empat tahun sekali (Novanti *et al.*, 2018). PISA dan TIMSS sebagian besar mencakup penilaian tentang bagaimana siswa dapat menerapkan dan menggunakan konsep-konsep ilmiah yang dipelajari dalam kehidupan nyata. PISA melakukan penilaian literasi sains yang tidak hanya berfokus pada konten sains, tetapi juga pengetahuan sains melalui pertanyaan dan topik yang bermakna. Penilaian ini juga bertujuan untuk menilai sejauh mana individu dapat membentuk pemikiran dan pengetahuan ilmiah serta teknologi yang dapat dipertanggungjawabkan oleh bukti ilmiah dan penjelasan yang teoritis (Chasanah *et al.*, 2022).

Situasi sesungguhnya memperlihatkan fakta kalau literasi peserta didik Indonesia masih dalam kategori rendah. PISA 2022 diselenggarakan oleh OECD dan melibatkan penilaian. Setiap beberapa tahun sekali, OECD melakukan penilaian terhadap siswa internasional melalui PISA. Target RPJMN untuk literasi sains adalah 402, dan skor PISA 2022 adalah 383, yang merupakan penurunan 12 poin dari tahun 2018. rendahnya tingkatan literasi sains di tahun 2022 membuat peserta didik lebih sulit untuk memecahkan masalah dan membuat mereka lambat dalam mengambil keputusan. Hal ini juga menghambat peserta didik untuk mengembangkan dan meningkatkan kemampuan kreatif mereka dalam penerapan sains dalam kehidupan sehari-hari.

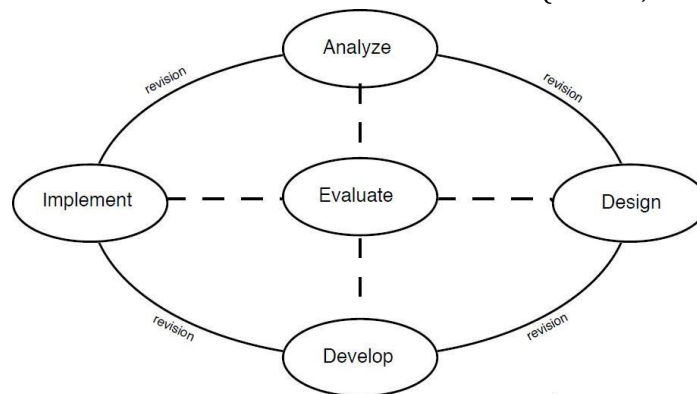
Dari hasil PISA, memberikan gambaran umum tentang literasi sains di Indonesia, tetapi diperlukan lebih banyak evaluasi literasi sains, misalnya dalam lingkup sekolah. Literasi sains belum termasuk dalam alat penilaian yang dibuat oleh para pendidik untuk dipakai sebagai alat ukur hasil belajar peserta didik (Ridwan *et al.*, 2020). Penilaian pembelajaran berbasis literasi sains dapat diterapkan dengan proses penemuan yang dihubungkan dalam kehidupan sehari-hari atau disebut pembelajaran kontekstual. Pembelajaran kontekstual ialah sebuah alat yang dipakai oleh para pendidik guna memasukkan kehidupan sehari-hari dan skenario dunia nyata ke dalam kelas. Ini bisa menginspirasi peserta didik untuk menghubungkan pengetahuan yang mereka peroleh selama pendidikan dengan aplikasi dunia nyata.

Instrumen penilaian digunakan untuk mengukur kompetensi literasi sains apabila telah dinyatakan valid. Kualitas instrumen penilaian yang baik apabila telah diuji secara teoritis dan empiris. Sehingga penelitian ini dilaksanakan untuk mendeskripsikan kualitas instrumen penilaian kelas berbasis literasi sains untuk mengukur kompetensi literasi sains pada materi energi terbarukan.

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 Gambar 1.

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



Pada tahap penelitian ini, hanya dilakukan pada tahap *development* (Pengembangan) karena bertujuan untuk menganalisis kualitas instrumen penilaian berdasarkan validitas teoritis dan empiris yang dikembangkan. Teknik pengumpulan data pada penelitian ini menggunakan metode tes dan angket. Metode tes diterapkan kepada 30 mahasiswa aktif angkatan 2023 S1 Pendidikan Fisika Unesa melalui uji coba instrumen dan metode angket dilakukan melalui hasil validasi yang diisi oleh para ahli yaitu dua dosen dari Pendidikan Fisika Unesa dan satu guru MBI Amanatul Ummah Pacet Mojokerto.

Pada desain model penelitian, di tahap awal ialah *analyze*, tahap ini diterapkan untuk memperoleh hasil analisis kebutuhan, kurikulum dan materi. Kemudian, tahap kedua ialah *design*, diterapkan untuk merancang instrumen penilaian yang berisi kisi-kisi, naskah soal dan rubrik penilaian yang akan dikembangkan. Kemudian, untuk tahap ketiga ialah *development*, tahap ini merupakan tahapan terakhir dalam penelitian ini yang diterapkan untuk memperoleh hasil validasi dari para ahli untuk mengetahui validitas teoritis instrumen. Pada tahap terakhir ini juga dilakukan uji coba untuk mengetahui validitas empiris butir soal.

Validitas teoritis diterapkan dengan menggunakan lembar validasi yang di validasi oleh 3 validator para ahli. Validasi ini berisi tentang penilaian pada aspek kelayakan isi, bahasa dan konstruk. Instrumen penilaian yang akan divalidasi terdiri dari instrumen penilaian pada ranah pengetahuan, sikap dan keterampilan. Lembar validasi penilaian kemudian dianalisis berdasarkan Skala Likert.

Tabel 1. Skala Likert

Skor	Kriteria
4	Sangat Valid
3	Valid
2	Cukup Valid
1	Tidak Valid

(Sugiyono, 2017)

Berdasarkan skor pada Tabel 1, terdapat kriteria dari tidak valid sampai sangat valid. Skor yang diperoleh dari 3 dosen ahli atau validator akan diambil nilai modus atau nilai yang sering muncul dalam validasi pada setiap ranah. Penilaian validitas teoritis ini dinilai berdasarkan aspek kelayakan isi, konstruk dan bahasa. Kemudian, akan dihitung persentase nilai dari rata-rata 3 validator tersebut. Berikut persamaan yang digunakan untuk menghitung persentase validitas teoritis, yaitu menggunakan rumus PoA.

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

Dengan:

A = Frekuensi aspek yang dinilai oleh pengamat dengan memberikan frekuensi tinggi

B = Frekuensi aspek yang dinilai oleh pengamat dengan memberikan frekuensi rendah

Instrumen dikatakan valid jika skor POA $\geq 75\%$. (Borich, 1994) Apabila, instrumen sudah dinyatakan valid berdasarkan validitas teoritis. Maka, akan dilakukan uji coba terbatas pada mahasiswa pendidikan fisika Universitas Negeri Surabaya. Hasil dari Uji coba nantinya digunakan untuk mengetahui apakah butir soal dinyatakan valid atau tidak yang mencakup dalam ranah pengetahuan. Hasil data dari uji coba ini di analisis menggunakan excel untuk mengetahui nilai validitas empiris butir soal, reliabilitas, tingkat kesukaran dan daya pembeda.

Uji validitas pada empiris ini digunakan untuk mengetahui apakah butir soal dinyatakan valid atau tidak yang mencakup dalam ranah pengetahuan. Perhitungan validitas empiris ini dilakukan dengan korelasi *product moment* (Arikunto, 2018).

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

Keterangan:

r_{xy} = Koefisien korelasi *product moment*

N = Banyaknya responden uji coba

X = Jumlah skor item

Y = Jumlah skor total

Instrumen tes berbasis literasi sains yang dikembangkan bisa dinyatakan valid apabila memenuhi kriteria cukup dengan nilai $r_{XY} \geq 0,41$. Untuk

menginterpretasikan besarnya koefisien korelasi digunakan kriteria seperti tampak pada Tabel 2.

Tabel 2. Kriteria Uji Validitas Empiris

Skor	Kategori
$0,81 \leq r_{XY} \leq 1,00$	Sangat Tinggi
$0,61 \leq r_{XY} \leq 0,80$	Tinggi
$0,41 \leq r_{XY} \leq 0,60$	Cukup
$0,21 \leq r_{XY} \leq 0,40$	Rendah
$0,0 \leq r_{XY} \leq 0,20$	Sangat Rendah

(Arikunto, 2018)

Instrumen tes mempunyai reliabilitas yang tinggi apabila tes tersebut menghasilkan hasil yang konsisten. Reliabilitas tes yang berbentuk uraian akan diukur dengan rumus koefisien Alpha Cronbach dengan persamaan 3.3.

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

Keterangan:

r_{11} = Nilai Reliabilitas

n = Banyaknya item soal

$\sum \sigma_i^2$ = Jumlah varian skor butir soal

σ_t^2 = Varian total

Pada kriteria uji reliabilitas instrumen tes yakni dari perhitungan r_{11} kemudian dibandingkan dengan $r_{product\ moment}$ dalam tabel, bisa dikatakan reliabel apabila $r_{hitung} > r_{tabel}$. Apabila ingin mengetahui tinggi rendahnya sebuah reliabilitas instrumen yang telah digunakan pada Tabel 3.

Tabel 3. Kategori Reliabilitas Instrumen

Interval Skor	Kategori
$0,800 \leq r_{11} \leq 1,00$	Sangat Tinggi
$0,600 \leq r_{11} \leq 0,800$	Tinggi
$0,400 \leq r_{11} \leq 0,600$	Cukup
$0,200 \leq r_{11} \leq 0,400$	Rendah
$0,0 \leq r_{11} \leq 0,200$	Sangat Rendah

(Arikunto, 2018)

Tingkat kesukaran menurut Arifin (2016) dapat ditentukan menggunakan persamaan 4.

$$Tingkat\ Kesukaran = \frac{\frac{\sum x}{N}}{skor\ maksimum} \quad (4)$$

Keterangan:

$\sum x$ = jumlah skor peserta didik tiap soal

N = jumlah peserta didik

Soal yang bagus adalah soal yang tidak terlalu mudah dan tidak terlalu sukar. Oleh karena itu, soal yang mudah tidak dapat memicu kemampuan peserta didik. Sedangkan, soal yang terlalu sulit akan menyebabkan peserta didik putus asa dan tidak semangat untuk mencoba kembali. Kriteria pada taraf tingkat kesukaran menjadi beberapa kriteria, seperti Tabel 4.

Tabel 4 Kriteria Taraf Tingkat Kesukaran

Nilai	Interpretasi
$0,80 \leq P \leq 0,30$	Sukar
$0,31 \leq P \leq 0,70$	Sedang
$0,71 \leq P \leq 1,00$	Mudah

(Arikunto, 2018)

Untuk mengetahui daya beda instrumen tiap soal, menurut Arifin (2016) dapat menggunakan persamaan 3.5.

$$DP = \frac{\Sigma K_A - \Sigma K_B}{\text{skor maksimum}} \quad (5)$$

Keterangan:

DP = daya pembeda

ΣK_A = rata-rata kelompok atas

ΣK_B = rata-rata kelompok bawah

Tiap instrumen soal yang dikembangkan bisa dinyatakan memiliki kemampuan daya pembeda jika nilai $D \geq 20$. Arikunto (2018) menyampaikan bahwa klasifikasi daya pembeda menjadi lima kriteria pada Tabel 5.

Tabel 5 Klasifikasi Daya Pembeda

Interval Skor	Kategori
$DP \leq 0,00$	Sangat jelek
$0,00 \leq DP \leq 0,20$	Jelek
$0,20 \leq DP \leq 0,40$	Cukup
$0,40 \leq DP \leq 0,70$	Baik
$0,71 \leq DP \leq 1,00$	Sangat Baik

(Arikunto, 2018)

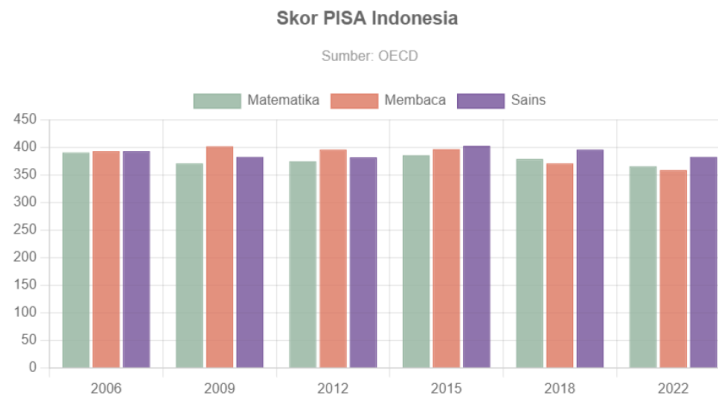
HASIL DAN PEMBAHASAN

Analyze (Analisis)

Analisis dilakukan dengan menganalisis kebutuhan di sekolah yang mencakup analisis kebutuhan, kurikulum dan materi yang digunakan. Pada analisis kebutuhan ini dilakukan dengan mengkaji literatur *review* terkait rendahnya tingkat literasi sains di Indonesia. instrumen penilaian berbasis literasi sains sangat

dibutuhkan untuk mengukur kemampuan dan kompetensi literasi sains peserta didik. Kriteria kelayakan ini menjadi aspek yang penting dalam menyusun instrumen penilaian ini yang berpedoman pada aspek literasi sains PISA (2025).

Gambar 2. Hasil Skor PISA Indonesia



Pada Gambar 2, bisa disimpulkan bahwa skor literasi sains menurut PISA peserta didik dari tahun ke tahun mengalami penurunan skor atau skor di bawah rata-rata internasional. Peserta didik belum mampu untuk mengerjakan atau memahami soal yang memuat aspek literasi sains. Nampaknya guru masih belum memahami sepenuhnya bahwa pembelajaran fisika ada kaitannya dengan pengembangan keterampilan peserta didik. Proses pembelajaran dan perangkat penilaian yang digunakan masih tradisional dan berbasis penguasaan konsep, sehingga peserta didik belum terbiasa dengan kemampuan literasi sains (Al Hakim, 2020).

Analisis ini dilakukan dengan wawancara dan literatur dari kajian. Penelitian ini dilakukan di MBI Amanatul Ummah Pacet Mojokerto yang mana sekolah ini menerapkan 2 kurikulum yaitu Kurikulum Nasional dan kurikulum internasional. Kurikulum nasional ini kombinasi antara KTSP, kurikulum 2013 dan kurikulum merdeka, serta untuk kurikulum internasional ini untuk kelas yang memprogram jenjang lanjut di Mesir. Peneliti melakukan wawancara dengan salah satu guru fisika di sekolah tersebut untuk mengetahui penggunaan instrumen penilaian berbasis kelas pada literasi sains. Kemudian, penerapan instrumen penilaian berbasis literasi sains di sekolah tersebut sudah pernah dilakukan tetapi hanya beberapa guru yang melakukan dan untuk aspek literasi sains pun belum sepenuhnya diterapkan oleh guru kepada peserta didik.

Kemudian, pada analisis kurikulum dilakukan analisis terkait kurikulum pembelajaran dari tujuan pembelajaran dan capaian pembelajaran peserta didik kelas X yang digunakan untuk menentukan materi pada instrumen penilaian yang akan dikembangkan. Tujuan pembelajaran dan capaian pembelajaran ini digunakan juga sebagai bahan pertimbangan untuk menentukan teknik penilaian instrumen yang akan dikembangkan. Kemudian, pada analisis materi, digunakan untuk mengetahui konsep-konsep utama dari materi energi terbarukan yang akan

diterapkan kepada peserta didik dalam proses pembelajaran berlangsung, tiap butir soal yang dikembangkan ini memuat konteks materi yang akan digunakan oleh peserta didik. Materi energi terbarukan memiliki beberapa sub materi yaitu bentuk energi, sumber energi, kebutuhan energi dan dampak eksplorasi energi.

Design (Perencanaan)

Pada tahap ini dilakukan perencanaan instrumen penilaian yang mencakup dari kisi-kisi soal yang berisi sub materi, indikator soal dan aspek kompetensi literasi sains. Tahap ini juga merancang sebuah instrumen angket respon peserta didik sebagai bahan evaluasi pada pengembangan penelitian ini.

Development (Pengembangan)

Pada tahap ini, instrumen penilaian yang telah dirancang pada tahap sebelumnya akan ditelaah oleh dosen ahli Fisika Universitas Negeri Surabaya dan guru MBI Amanatul Ummah Pacet Mojokerto yang akan dipakai untuk penelitian. Instrumen penilaian ini ditelaah berdasarkan 3 aspek yakni materi, konstruk dan bahasa. Selain itu, pada tahap pengembangan ini dilakukan uji coba terbatas instrumen penilaian terhadap 30 mahasiswa Prodi Pendidikan Fisika. Instrumen ini bisa digunakan apabila memenuhi kriteria valid secara teoritis dan empiris. Sehingga, di tahap ini didapatkan instrumen penilaian yang digunakan untuk mengukur profil peserta didik pada kompetensi literasi sains. Hasil yang diperoleh pada tahap ini terdiri dari validitas teoritis dan validitas empiris. Berikut hasil yang diperoleh pada validitas teoritis.

Tabel 6 Hasil Validasi Instrumen Penilaian Ranah Pengetahuan

No	Aspek yang Dinilai	Modus	Rerata Skor	PoA	Kriteria
1	Ranah isi	4	3,75	89%	Valid
2	Ranah Konstruksi	4	4	93%	Sangat Valid
3	Ranah Bahasa	4	3,5	89%	Valid

Tabel 7 Hasil Validasi Instrumen Penilaian Ranah Keterampilan

No	Aspek yang Dinilai	Modus	Rerata Skor	PoA	Kriteria
1	Ranah isi	4	3,5	86%	Valid
2	Ranah Konstruksi	4	4	90%	Sangat Valid
3	Ranah Bahasa	4	4	93%	Sangat Valid

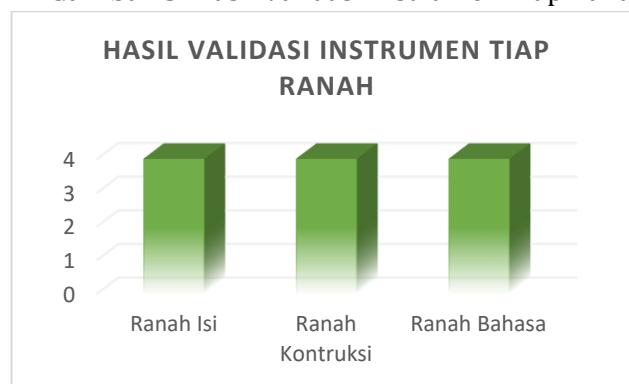
Tabel 8 Hasil Validasi Instrumen Penilaian Ranah Sikap

No	Aspek yang Dinilai	Modus	Rerata Skor	PoA	Kriteria
1	Ranah isi	4	3,75	89%	Valid

2	Ranah Konstruksi	4	4	89%	Valid
3	Ranah Bahasa	4	4	89%	Valid

Selain memberikan penilaian, validator juga memberikan saran dan masukan terhadap instrumen yang dikembangkan. Sehingga, instrumen penilaian berbasis literasi sains yang telah divalidasi selanjutnya direvisi sesuai dengan saran validator. Dari hasil data validitas teoritis, diperoleh hasil validasi keseluruhan dengan persentase PoA yaitu 90% dengan skor modus 4 dan kriteria Sangat Valid. Sehingga, instrumen penilaian berbasis literasi sains layak dikembangkan dalam kategori sangat valid.

Gambar 3. Hasil Validasi Instrumen Tiap Ranah

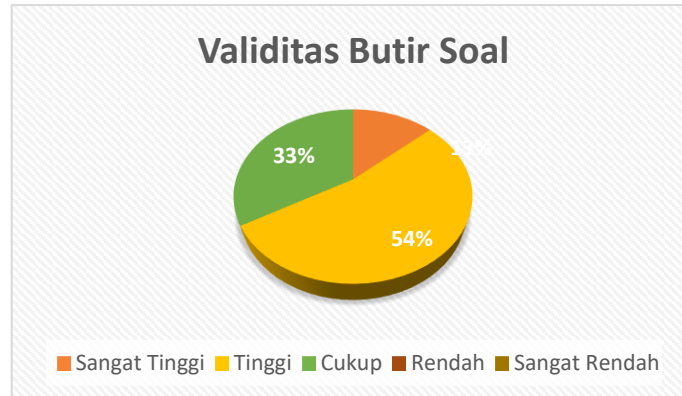


Validitas pada ranah materi menunjukkan bahwa instrumen penilaian yang disusun telah sesuai dengan kurikulum, materi, dan kemampuan yang diujikan (Cohen *et al.*, 2007). Kriteria yang diperoleh pada ranah teoritis memiliki kategori sangat valid dengan skor modus 4. Pada setiap ranah memiliki kriteria penilaian masing-masing, pada ranah konstruksi ditujukan untuk mengetahui dan menilai bagaimana individu akan bertindak dalam berbagai situasi secara spesifik yang berbeda. Pada ranah konstruksi ini memuat aspek petunjuk pengerjaan, kejelasan instrumen, rubrik penilaian dan rubrik soal. Pada ranah bahasa memuat terkait sebuah pemahaman dan kejelasan bahasa pada instrumen dalam penggunaan ejaan yang benar dan sesuai.

Ketika sudah didapatkan hasil dari validitas teoritis, instrumen penilaian berbasis literasi sains akan di analisis melalui tahap validitas empiris. Validitas empiris diperoleh dari uji coba instrumen penilaian kepada 30 Mahasiswa Pendidikan Fisika Unesa. Skor nilai instrumen kemudian dianalisis untuk memperoleh hasil validitas empiris, reliabilitas, tingkat kesukaran dan daya pembeda.

Hasil validitas empiris butir soal diperoleh untuk menentukan butir soal pada instrumen yang bisa dikatakan valid.

Gambar 4. Persentase Validitas Empiris



Berdasarkan Gambar diagram 4 menunjukkan bahwa dari 15 soal yang telah di kembangkan dan diujicobakan, terdapat hasil persentase validitas empiris butir soal dengan kriteria berbeda. Butir soal yang dikembangkan bisa dinyatakan valid apabila memenuhi kriteria cukup pada setiap aspek validitas empiris.

Nilai koefisien korelasi untuk menentukan reliabilitas dihitung dengan persamaan *Alpha Cronbach*. Hasil perhitungan tersebut dapat ditunjukkan pada Tabel 9.

Tabel 9 Hasil Analisis Reliabilitas Butir Soal

Subjek Uji Coba	r_{11}	r_{tabel}	Kriteria
30 Mahasiswa Prodi Pendidikan Fisika	0,922	0,361	Reliabel

Pada Tabel 9, dinyatakan bahwa instrumen penilaian literasi sains yang dikembangkan memenuhi kriteria reliabel. Tingkat konsistensi dalam instrumen yang dihasilkan dikenal sebagai reliabilitas. Dari hasil analisis jawaban instrumen diperoleh nilai reliabilitas butir soal yaitu 0,922 dari r_{tabel} yaitu 0,361 dengan kategori nilai yaitu sangat tinggi dan reliabel. Berdasarkan hasil tersebut, menunjukkan bahwa instrumen penilaian yang dikembangkan memiliki nilai reliabilitas atau konsistensi setiap soalnya sangat tinggi. Hasil dari reliabilitas tersebut ialah ketika instrumen penilaian diberikan atau diterapkan kepada peserta didik maka hasil nilai yang diperoleh rata-rata sama.

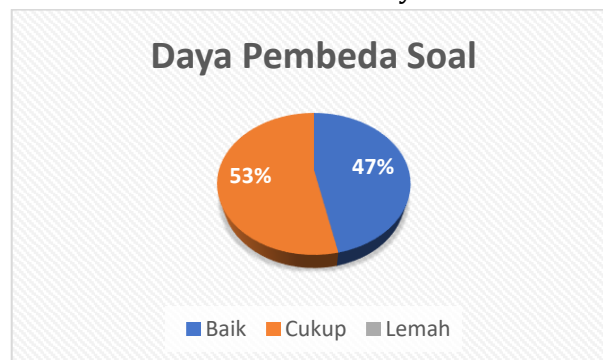
Tingkat kesukaran merupakan tingkatan kesukaran pada tiap butir soal. Untuk menentukan kriteria butir soal mudah atau sukar bergantung pada skor akhir persentase item soal seluruhnya dari jawaban instrumen penilaian (Arikunto, 2018). Pada Gambar 4.3 hasil analisis dari jawaban uji coba instrumen penilaian secara keseluruhan menunjukkan bahwa tiap butir soal yang dikembangkan memperoleh kriteria sedang dan mudah dengan rincian tujuh soal sedang dan delapan soal mudah.

Gambar 5. Persentase Tingkat Kesukaran



Daya pembeda soal berkaitan dengan kemampuan soal untuk mengkategorikan siswa sesuai dengan kemampuannya (Arikunto, 2018). Koefisien daya pembeda tiap item soal bergantung pada skor siswa yang telah dikelompokkan dalam kelas atas dan bawah. Penentuan kelompok ini mengacu pendapat Arikunto (2018) bahwa jika peserta tes kurang dari 100 maka pembagian kelompok atas dan bawah masing-masing 50%. Pada Gambar 4.4 hasil penelitian ini di uji coba kepada 30 Mahasiswa Prodi Pendidikan Fisika Unesa dengan jumlah nilai kelompok atas dan bawah yaitu 15 mahasiswa.

Gambar 6. Persentase Daya Pembeda



KESIMPULAN DAN SARAN

Kualitas instrumen penilaian berbasis kelas untuk mengukur kompetensi literasi sains yang dikembangkan dinyatakan valid berdasarkan validitas teoritis dan empiris. Instrumen penilaian literasi sains pada aspek pengetahuan, keterampilan dan sikap dinyatakan valid berdasarkan validitas teoritis. Kemudian, instrumen penilaian pada aspek pengetahuan yang dilakukan setelah uji coba pada mahasiswa fisika menghasilkan validitas empiris dengan 15 butir soal dinyatakan valid, serta diperoleh skor nilai reliabilitas dengan r_{hitung} sebesar 0,922 dan r_{tabel} sebesar 0,361. Sehingga, instrumen penilaian yang disusun bisa dinyatakan reliabel dan dapat digunakan.

DAFTAR PUSTAKA

- Al Hakim, R. R. (2020). Model Energi Indonesia, Tinjauan Potensi Energy Terbarukan Untuk Ketahanan Energi Di Indonesia : Literatur Review. *Andasih Jurnal Pengabdian Kepada Masyarakat*, 1(1), 1-11.
- Ardianti, R., Surahman, E., & Sujarwanto, E. (2022). Pengembangan Instrumen Penilaian Literasi Sains pada Bahasan Usaha dan Energi di Madrasah Aliyah. *Journal for Physics Education and Applied Physics*, 4(1). <http://jurnal.unsil.ac.id/index.php/Diffraction>
- Chasanah, N., Widodo, W., & Suprpto, N. (2022). Pengembangan Instrumen Assessment Literasi Sains Untuk Mendeskripsikan Profil Peserta Didik. *PENDIPA Journal of Science Education*, 6(2), 474-483. <https://doi.org/10.33369/pendipa.6.2.474-483>
- Drs. Zainal Arifin, M. P. (2020). Evaluasi Pembelajaran. In *Evaluasi Pembelajaran*.
- Dwi Indrawati, M., & Sunarti, T. (2018). Pengembangan Instrumen Penilaian Literasi Sains Fisika Peserta Didik Pada Bahasan Gelombang Bunyi Di SMA Negeri 1 Gedangan Sidoarjo. *Jurnal Inovasi Pendidikan Fisika (JIPF)*, 07(01), 14-20.
- Fitriana, F., Kurniawati, Y., & Utami, L. (2019). Analisis Keterampilan Proses Sains Peserta Didik Pada Materi Laju Reaksi Melalui Model Pembelajaran Bounded Inquiry Laboratory. *JTK (Jurnal Tadris Kimiya)*, 4(2), 226-236. <https://doi.org/10.15575/jtk.v4i2.5669>
- Gulpepe, N., & Kilic, Z. (2015). Effect of scientific argumentation on the development of scientific process skills in the context of teaching chemistry. *International Journal of Environmental and Science Education*, 10(1), 111-132. <https://doi.org/10.12973/ijese.2015.234a>
- Irwan, A. P., Usman, & Amin, B. D. (2019). Analisis Kemampuan Literasi Sains Peserta Didik. *Jurnal Sains Dan Pendidikan Fisika (JSPF)*, 15(3), 17-24.
- Kurniawati, R. D., & Wijastuti, A. Y. (2021). Pengembangan Instrumen Assesmen Berbasis Literasi Sains Materi Gaya Dan Gerak Bagi Siswa SMPLB Tunarungu. *Jurnal GRAB KIDS*, 1.
- Nofiana, M. (2017). Profil Kemampuan Literasi Sains Siswa SMP di Kota Purwokerto Ditinjau dari Aspek Konten, Proses, dan Konteks Sains. *JSSH (Jurnal Sains Sosial Dan Humaniora)*, 1(2), 77. <https://doi.org/10.30595/jssh.v1i2.1682>
- OECD. (2022). *Programme For International Student Assesment (PISA) Results From PISA 2022*.
- Pamungkas, I., Sumaryoto, S., & Sriyono, H. (2024). Analisis Muatan Nilai-Nilai Pendidikan Karakter Pada Buku Teks Pelajaran IPS SMP Kurikulum 2013. *Herodotus: Jurnal Pendidikan IPS*, 6(3), 380. <https://doi.org/10.30998/herodotus.v6i3.16980>
- Ridwan, I., Anggraeni, S., & Supriyatno, B. (2020). Analisis Lembar Kerja Peserta Didik Sekolah Menengah Atas Pada Praktikum Uji Urin. *Biodik*, 6(3), 319-327. <https://doi.org/10.22437/bio.v6i3.9462>
- Rusilowati, A., Kurniawati, L., Nugroho, S. E., & Widiyatmoko, A. (2016). Developing an instrument of scientific literacy asesment on the cycle theme.

EduInovasi: Journal of Basic Educational Studies

Vol 4 No 3 (2024) 1658 - 1671 P-ISSN 2774-5058 E-ISSN 2775-7269

DOI: 47467/eduinovasi.v4i3.4348

International Journal of Environmental and Science Education, 11(12), 5718-5727